SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF PHARMACY

UNIT – I COMPUTER APPLICATIONS IN PHARMACY – BP205T

UNIT – I

Number system: Binary number system, Decimal number system, Octal number system, Hexadecimal number systems, conversion decimal to binary, binary to decimal, octal to binary etc, binary addition, binary subtraction – One's complement ,Two's complement method, binary multiplication, binary division Concept of Information Systems and Software : Information gathering, requirement and feasibility analysis, data flow diagrams, process specifications, input/output design, process life cycle, planning and managing the project

Number Systems

The number system is a way to represent or express numbers. You have heard of various types of number systems such as the whole numbers and the real numbers. But in the context of computers, we define other types of number systems. They are:

- The decimal number system
- The binary number system
- The octal number system and
- The hexadecimal number system

Decimal Number System (Base 10)

In this number system, the digits 0 to 9 represents numbers. As it uses 10 digits to represent a number, it is also called the base 10 number system. Each digit has a value based on its position called place value. The value of the position increases by 10 times as we move from right to left in the number.
For example, the value of 786 is
$= 7 \times 10^2 + 8 \times 10^1 + 6 \times 10^0$
$= 700 + 80 + 6$

Binary Number System (Base 2)

A computer can understand only the "on" and "off" state of a switch. These two states are represented by 1 and 0. The combination of 1 and 0 form binary numbers. These numbers represent various data. As two digits are used to represent numbers, it is called a binary or base 2 number system.
The binary number system uses positional notation. But in this case, each digit is multiplied by the appropriate power of two based on its position.
For example, $(101101)_2$ in decimal is
$= 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$
$= 1 \times 32 + 0 \times 16 + 1 \times 8 + 1 \times 4 + 0 \times 2 + 1 \times 1$

$= 32 + 8 + 4 + 1$
$= (45)_{10}$

Octal Number System (Base 8)

This system uses digits 0 to 7 (i.e. 8 digits) to represent a number and the numbers are as a base of 8.
For example, $(24)_8$ in decimal is
$= 2\times8^1+4\times8^0$
$= (20)_{10}$

Hexadecimal Number System (Base 16)

In this system, 16 digits used to represent a given number. Thus it is also known as the base 16 number system. Each digit position represents a power of 16. As the base is greater than 10, the number system is supplemented by letters. Following are the hexadecimal symbols: 0, 1, 2, 3, 4, 5, 6,
7, 8, 9, A, B, C, D, E, F
To take A, B, C, D, E, and F as part of the number system is conventional and has no logical or deductive reason.

Number System Chart

Name	Base	Symbols	Example
Decimal	10	0,1,2,3,4,5,6,7,8,9	$(2795)_{10}$
Binary	2	0,1	111000010
Octal	8	0,1,2,3,4,5,6,7	$(1576)_8$
Hexadecimal	16	0,1,2,3,4,5,6,7,8,9,A,B,C,D,E,F	3DB

Information system

Information systems (**IS**) are formal, sociotechnical, organizational systems designed to collect, process, store, and distribute information. In a sociotechnical perspective, information systems are composed by four components: task, people, structure (or roles), and technology.

The six components that must come together in order to produce an information system are: (Information systems are organizational procedures and do not need a

computer or software, this data is erroneous)

1. Hardware: The term hardware refers to machinery. This category includes the computer itself, which is often referred to as the central processing unit (CPU), and all of its support equipment. Among the support, equipment are input and output devices, storage devices and communications devices.
2. Software: The term software refers to computer programs and the manuals (if any) that support them. Computer programs are machine-readable instructions that direct the circuitry within the hardware parts of the system to function in ways that produce useful information from data. Programs are generally stored on some input/output medium, often a disk or tape.
3. Data: Data are facts that are used by programs to produce useful information. Like programs, data are generally stored in machine-readable form on disk or tape until the computer needs them.
4. Procedures: Procedures are the policies that govern the operation of a computer system. "Procedures are to people what software is to hardware" is a common analogy that is used to illustrate the role of procedures in a system.
5. People: Every system needs people if it is to be useful. Often the most overlooked element of the system are the people, probably the component that most influence the success or failure of information systems. This includes "not only the users, but those who operate and service the computers, those who maintain the data, and those who support the network of computers."
6. Feedback: it is another component of the IS, that defines that an IS may be provided with a feedback

Data is the bridge between hardware and people. This means that the data we collect is only data until we involve people. At that point, data is now information.

Types of information system

Some examples of such systems are:

- data warehouses
- enterprise resource planning
- enterprise systems
- expert systems
- search engines
- geographic information system
- global information system
- office automation.

Systems Development Life Cycle

An effective System Development Life Cycle (SDLC) should result in a high quality system that meets customer expectations, reaches completion within time and cost

evaluations, and works effectively and efficiently in the current and planned Information Technology infrastructure.

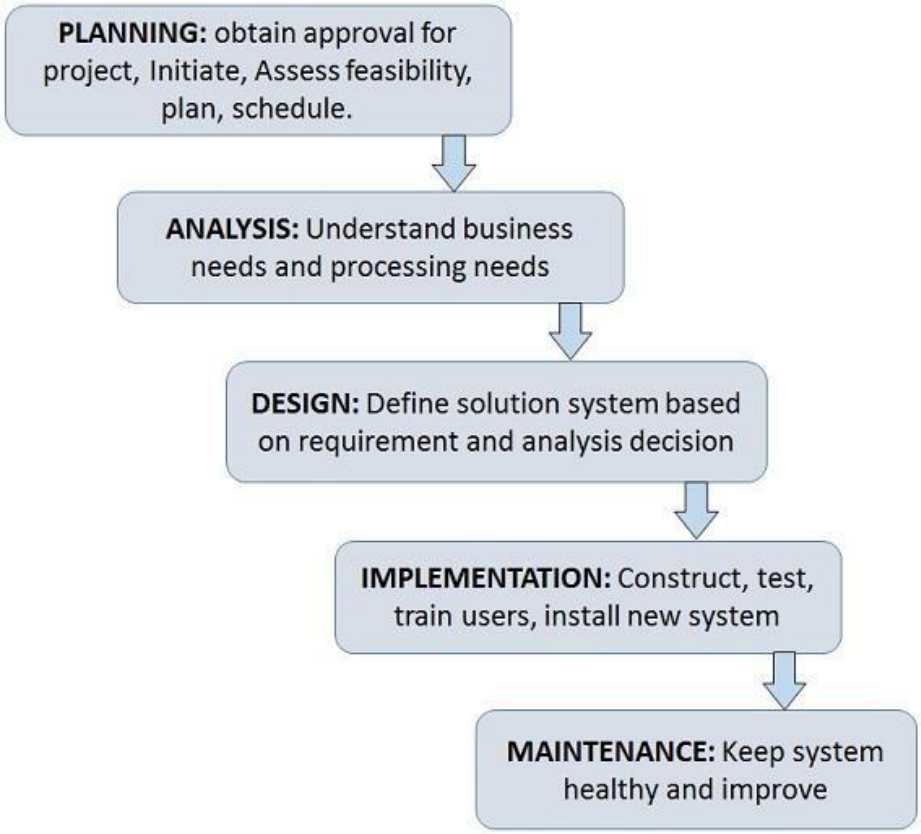
System Development Life Cycle (SDLC) is a conceptual model which includes policies and procedures for developing or altering systems throughout their life cycles.

SDLC is used by analysts to develop an information system. SDLC includes the following activities –

- requirements
- design
- implementation
- testing
- deployment
- operations
- maintenance

Phases of SDLC

Systems Development Life Cycle is a systematic approach which explicitly breaks down the work into phases that are required to implement either new or modified Information System.



****

Binary to Decimal Conversion

Method 1
Multiply bits with powers of two

1. Write the powers of two (8 bit representation)

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0

2. Calculate the powers of two

128	64	32	16	8	4	2	1

3. Write the corresponding bit of the binary number under each power of two

0	0	1	1	1	0	0	1

4. Multiply each bit with its corresponding power of two. Add all products.

$0 \cdot 128 + 0 \cdot 64 + 1 \cdot 32 + 1 \cdot 16 + 1 \cdot 8 + 0 \cdot 4 + 0 \cdot 2 + 1 \cdot 1$
$= 32 + 16 + 8 + 1 = 57$

5. The result of the sum is the decimal number

$$111001_2 = 57_{10}$$

Decimal to Binary Conversion

Method 1
Descending Powers of Two and Subtraction

1. Write the power of two

$2^0 = 1$
$2^1 = 2$
$2^2 = 4$
$2^3 = 8$
$2^4 = 16$
$2^5 = 32$
$2^6 = 64$
$2^7 = 128$
$2^8 = 256$

2. Write the number as a sum of powers of two

$57 = 32 + 25$
$= 32 + 16 + 9$
$= 32 + 16 + 8 + 1$
$= 1 \times 32 + 1 \times 16 + 1 \times 8 + 1 \times 1$
$= 1 \times 2^5 + 1 \times 2^4 + 1 \times 2^3 + 1 \times 2^0$
$= 1 \times 2^5 + 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$

3. Extract the coefficients of the powers of two

$$57_{10} = 111001_2$$

Octal to Binary

Octal number is one of the number systems which has value of base is 8, that means there only 8 symbols: 0, 1, 2, 3, 4, 5, 6, and 7. Whereas **Binary** number is most familiar number system to the digital systems, networking, and computer professionals. It is base 2 which has only 2 symbols: 0 and 1, these digits can be represented by off and on respectively.

Conversion from Octal to Binary number system

There are various direct or indirect methods to convert a octal number into binary number. In an indirect method, you need to convert an octal number into other number system (e.g., decimal or hexadecimal), then you can convert into binary number by converting each digit into binary number from hexadecimal system and using conversion system from decimal to binary number.

There is a simple direct method to convert an octal number to binary number. Since there are only 8 symbols (i.e., 0, 1, 2, 3, 4, 5, 6, and 7) in octal representation system and its base (i.e., 8) is equivalent of 23=8. So, you can represent each digit of octal in group of 3 bits in binary number.

Octal Symbol	Binary equivalent
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

This method is simple and also works as reverse of Binary to Octal Conversion. The algorithm is explained as following below.

- Take Octal number as input
- Convert each digit of octal into binary.
- That will be output as binary number.

Example-1 Convert octal number 540 into binary

number. According to above algorithm, equivalent

binary number will be,

$= (540)_8$
$= (101\ 100\ 000)_2$
$= (101100000)_2$

This is very simple conversion, you can use for mixed (integer with fractional) octal number as well.

Example-2 − Convert octal number 352.563 into binary number.

According to above algorithm, equivalent binary number will be,

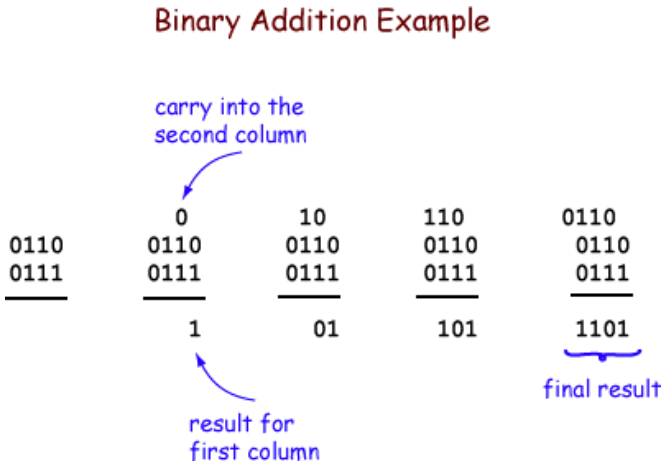
$= (352.563)_8$
$= (011\ 101\ 010\ .\ 101\ 110\ 011)_2$

$= (011101010.101110011)_2$

Binary addition

Binary Addition

A	B	Sum	Carry
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1



One's Complement and Two's Complement

One's complement and two's complement are two important binary concepts. Two's complement is especially important because it allows us to represent signed numbers in binary, and one's

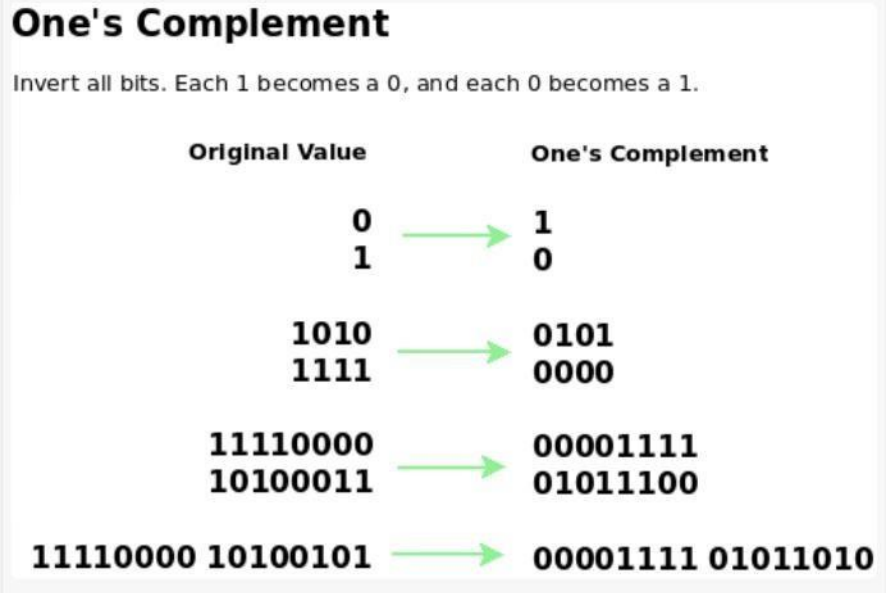
complement is the interim step to finding the two's complement.

Two's complement also provides an easier way to subtract numbers using addition instead of using the longer.

One's Complement

If all bits in a byte are inverted by changing each 1 to 0 and each 0 to 1, we have formed the one's complement of the number.

```
Original       One's Complement
-------------------------------
10011001  -->  01100110
10000001  -->  01111110
11110000  -->  00001111
11111111  -->  00000000
00000000  -->  11111111
```



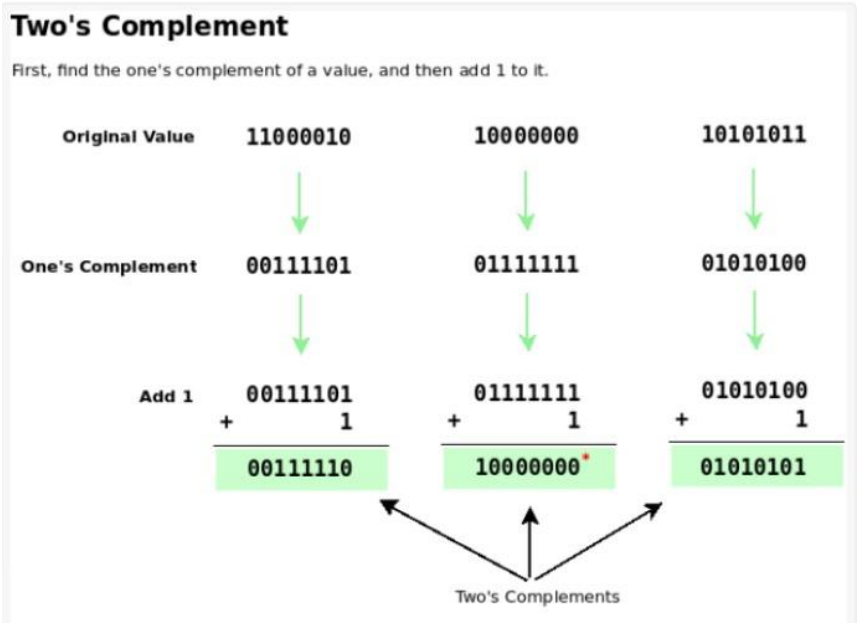
One's complement is useful for forming the two's complement of a number.

Two's Complement (Binary Additive Inverse)

The two's complement is a method for representing positive and negative integer values in binary. The useful part of two's complement is that it automatically includes the sign bit.

Rule: To form the two's complement, add 1 to the one's complement.

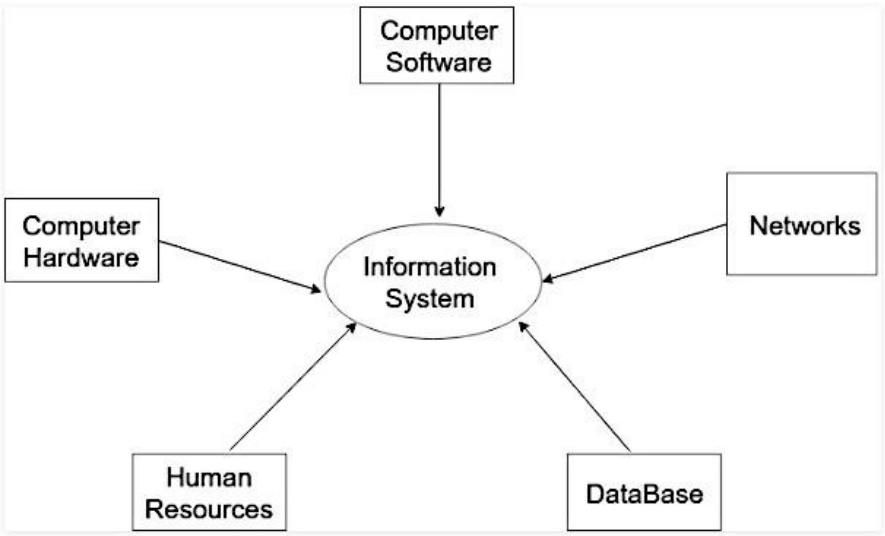
```
Step 1: Begin with the original binary value

    10011001  Original byte

Step 2: Find the one's complement

    01100110  One's complement

Step 3: Add 1  to the one's complement

    01100110  One's complement
  +        1  Add 1
 -----------
    01100111  <---  Two's complement
```



****

Components Of Information System

An **Information system** is a combination of hardware and software and telecommunication networks that people build to collect, create and distribute useful data, typically in an organisational, It defines the flow of information within the system. The objective of an information system is to provide appropriate information to the user, to gather the data, processing of the data and communicate information to the user of the system.



Components of the information system are as follows:

1. Computer Hardware:
Physical equipment used for input, output and processing. What hardware to use it depends upon the type and size of the organisation. It consists of input, an output device, operating system, processor, and media devices. This also includes computer peripheral devices.

2. Computer Software:
The programs/ application program used to control and coordinate the hardware components. It is used for analysing and processing of the data. These programs include a set of instruction used for processing information.
Software is further classified into 3 types:

1. System Software
2. Application Software
3. Procedures

3. Databases:
Data are the raw facts and figures that are unorganised that are and later processed to generate information. Softwares are used for organising and serving data to the user, managing physical storage of media and virtual resources. As the hardware can't work without software the same as software needs data for processing. Data are managed using Database management system.
Database software is used for efficient access for required data, and to manage knowledge bases.

4. Network:

- Networks resources refer to the telecommunication networks like the intranet, extranet and the internet.
- These resources facilitate the flow of information in the organisation.
- Networks consists of both the physicals devises such as networks cards, routers, hubs and cables and software such as operating systems, web servers, data servers and application servers.

- Telecommunications networks consist of computers, communications processors, and other devices interconnected by communications media and controlled by software.
- Networks include communication media, and Network Support.

5. Human Resources:

It is associated with the manpower required to run and manage the system. People are the end user of the information system, end-user use information produced for their own purpose, the main purpose of the information system is to benefit the end user. The end user can be accountants, engineers, salespersons, customers, clerks, or managers etc. People are also responsible to develop and operate information systems. They include systems analysts, computer operators, programmers, and other clerical IS personnel, and managerial techniques.

****

Project

manageme

nt

Definition

Project management is the application of processes, methods, skills, knowledge and experience to achieve specific project objectives according to the project acceptance criteria within agreed parameters.

What is a project?

A project is a unique, transient endeavour, undertaken to achieve planned objectives, which could be defined in terms of outputs, outcomes or benefits. A project is usually deemed to be a success if it achieves the objectives according to their acceptance criteria, within an agreed timescale and budget. Time, cost and quality are the building blocks of every project.

Time: scheduling is a collection of techniques used to develop and present schedules that show when work will be performed.
Cost: how are necessary funds acquired and finances managed?
Quality: how will fitness for purpose of the deliverables and management processes be assured?

The core components of project management are:

- defining the reason why a project is necessary;
- capturing project requirements, specifying quality of the deliverables, estimating resources and timescales;

- preparing a business case to justify the investment;
- securing corporate agreement and funding;
- developing and implementing a management plan for the project;
- leading and motivating the project delivery team;
- managing the risks, issues and changes on the project;
- monitoring progress against plan;
- managing the project budget;
- maintaining communications with stakeholders and the project organisation;
- provider management;
- closing the project in a controlled fashion when appropriate

References

1. https://www.toppr.com/guides/computer-aptitude-and-knowledge/basics-of- computers/number-systems/

2. https://en.wikipedia.org/wiki/Information_system

3. https://www.tutorialspoint.com/system_analysis_and_design/system_analysis_and_design_devel opment_life_cycle.htm

****

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF PHARMACY

UNIT – II COMPUTER APPLICATIONS IN PHARMACY – BP205T

UNIT – II

Web technologies:Introduction to HTML, XML,CSS and Programming languages, introduction to web servers and Server Products Introduction to databases, MYSQL, MS ACCESS, Pharmacy Drug database

HTML and XML

HTML is an abbreviation for HyperText Markup Language. **XML** stands for eXtensible Markup Language. **HTML** was designed to display data with focus on how data looks. **XML** was designed to be a software and hardware independent tool used to transport and store data, with focus on what data is.

HTML: HTML (**Hyper Text Markup Language**) is used to create web pages and web applications. It is a markup language. By HTML we can create our own static page. It is used for displaying the data not to transport the data. HTML is the combination of Hypertext and Markup language. Hypertext defines the link between the web pages. A markup language is used to define the text document within tag which defines the structure of web pages. This language is used to annotate (make notes for the computer) text so that a machine can understand it and manipulate text accordingly.

Exa
mple
INP
UT

```
<!DOCTYPE html>
<html>
<head>
        <title>GeeksforGeeks</title>
</head>
<body>
        <h1>GeeksforGeeks</h1>
        <p>A Computer Science portal for geeks</p>
</body>
</html>
```

Output

XML: XML (**eXtensible Markup Language**) is also used to create web pages and web applications. It is dynamic because it is used to transport the data not for displaying the data. The design goals of XML focus on simplicity, generality, and usability across the

Internet. It is a textual data format with strong support via Unicode for different human languages. Although the design of XML focuses on documents, the language is widely used for the representation of arbitrary data structures such as those used in web services.

INPUT

```
<?xml version = "1.0"?>
<contactinfo>
        <address category = "college">
                <name>G4G</name>
                <College>Geeksforgeeks</College>
                <mobile>2345456767</mobile>
        </address>
</contactinfo>
```

Output:

G4G

Geeksforge

eks

234545676

7

Difference between HTML and XML: There are many differences between HTML and XML. These important differences are given below:

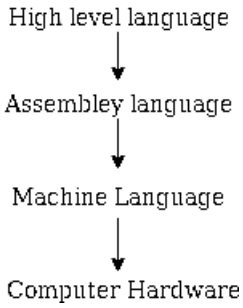
HTML	**XML**
HTML stands for **Hyper Text Markup Language.**	XML stands for **eXtensible Markup Language.**
HTML is static.	XML is dynamic.
HTML is a markup language.	XML provides framework to define markup languages.
HTML can ignore small errors.	XML does not allow errors.
HTML is not Case sensitive.	XML is Case sensitive.
HTML tags are predefined tags.	XML tags are user defined tags.
There are limited number of tags in HTML.	XML tags are extensible.

HTML does not preserve white spaces.	White space can be preserved in XML.
HTML tags are used for displaying the data.	XML tags are used for describing the data not for displaying.
In HTML, closing tags are not necessary.	In XML, closing tags are necessary.

Programming languages

A program is a set of instructions given to a computer to perform a specific operation. or computer is a computational device which is used to process the data under the control of a computer program.While executing the program, raw data is processed into a desired output format. These computer programs are written in a programming language which are high level languages. High level languages are nearly human languages which are more complex then the computer understandable language which are called machine language, or low level language.So after knowing the basics, we are ready to create a very simple and basic program. Like we have different languages to communicate with each other, likewise, we have different languages like C, C++, C#, Java, python, etc to communicate with the computers. The computer only understands binary language (the language of 0's and 1's) also called machine-understandable language or low-level language but the programs we are going to write are in a high-level language which is almost similar to human language.

Hierarchy of Computer language –



Most Popular Programming Languages –

- C
- Python
- C++
- Java
- SCALA
- C#
- R
- Ruby
- Go
- Swift

- JavaScript

Characteristics of a programming Language –

- A programming language must be simple, easy to learn and use, have good readability and human recognizable.
- Abstraction is a must-have Characteristics for a programming language in which ability to define the complex structure and then its degree of usability comes.
- A portable programming language is always preferred.
- Programming language's efficiency must be high so that it can be easily converted into a machine code and executed consumes little space in memory.
- A programming language should be well structured and documented so that it is suitable for application development.
- Necessary tools for development, debugging, testing, maintenance of a program must be provided by a programming language.
- A programming language should provide single environment known as Integrated Development Environment(IDE).
- A programming language must be consistent in terms of syntax and semantics.

Drug databases and their applications

Drug Databases

Drug databases are sites where information about drugs and medications are stored, and one of the largest (and most commonly used) drug databases is compiled by the **Food & Drug Administration (FDA)**. The FDA is a federal agency that oversees and controls all medications in the U.S., which includes:

- Over-the-counter (OTC) medications
- Prescription medications
- Dietary supplements
- Vaccines

Drug databases and web resources play a very important role in the pharmaceutical field.

Eg.**DrugBank**

The DrugBank database is a comprehensive, freely accessible, online database containing information on drugs and drug targets. As both a bioinformatics and a cheminformatics resource, DrugBank combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information.

The latest release of the database (version 5.0) contains 9591 drug entries including 2037 FDA- approved small molecule drugs, 241 FDA-approved biotech (protein/peptide) drugs, 96 nutraceuticals and over 6000 experimental drugs.[4] Additionally, 4270 non-redundant protein (i.e. drug target/enzyme/transporter/carrier) sequences are linked to

these drug entries. Each DrugCard entry (Fig. 1) contains more than 200 data fields with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data.

Four additional databases, HMDB, T3DB, SMPDB and FooDB are also part of a general suite
of metabolomic/cheminformatic databases. HMDB contains equivalent information on more than 40,000 human metabolites, T3DB contains information on 3100 common toxins and environmental pollutants, SMPDB contains pathway diagrams for nearly 700 human metabolic pathways and disease pathways, while FooDB contains equivalent information on ~28,000 food components and food additives.

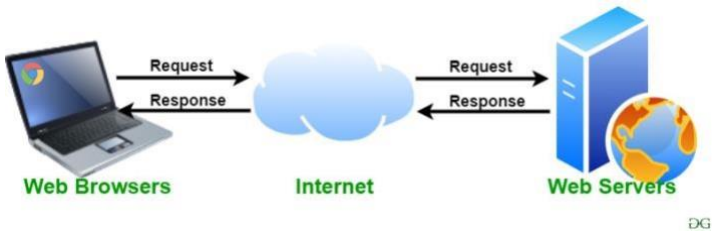
Web servers

Web Server and Its Type

Web Server: Web server is a program which processes the network requests of the users and serves them with files that create web pages. This exchange takes place using Hypertext Transfer Protocol (HTTP).

Basically, web servers are computers used to store HTTP files which makes a website and when a client requests a certain website, it delivers the requested website to the client. For example, you want to open Facebook on your laptop and enter the URL in the search bar of google. Now, the laptop will send an HTTP request to view the facebook webpage to another computer known as the webserver. This computer (webserver) contains all the files (usually in HTTP format) which make up the website like text, images, gif files, etc. After processing the request, the webserver will send the requested website-related files to your computer and then you can reach the website.

Different websites can be stored on the same or different web servers but that doesn't affect the actual website that you are seeing in your computer. The web server can be any software or hardware but is usually a software running on a computer. One web server can handle multiple users at any given time which is a necessity otherwise there had to be a web server for each user and considering the current world population, is nearly close to impossible. A web server is never disconnected from the internet because if it was, then it won't be able to receive any requests, and therefore cannot process them.



There are many web servers available in the market both free and paid.

Eg.

Apache HTTP server: It is the most popular web server and about 60 percent of the world's web server machines run this web server. The Apache HTTP web server was developed by the Apache Software Foundation. It is an open-source software which means that we can access and make changes to its code and mold it according to our

preference. The Apache Web Server can be installed and operated easily on almost all operating systems like Linux, MacOS, Windows, etc.

****

Databases and

MYSQL.

What is Database?

The **Database** is an essential part of our life. As we encounter several activities that involve our interaction with database, for example in the bank, in the railway station, in school, in a grocery store, etc. These are the places where we need to a large amount of data at one place and fetching of this data should be easy.

A database is a collection of data which is organized, which is also called as structured data. It can be accessed or stored at the computer system. It can be managed through Database management system (DBMS), which is a software which is used to manage data. Database refers to related data which is in a structured form.

In Database, data is organized into tables which consist of rows and columns and it is indexed so data gets updated, expanded and deleted easily. Computer databases typically contain file records data like transactions money in one bank account to another bank account, sales and customer details, fee details of student and product details. There are different kinds of databases, ranging from the most prevalent approach, the relational database, to a distributed database, cloud database or NoSQL database.

Types

- **Relational Database:**
  A relational database is made up of a set of tables with data that fits into a predefined category.
- **Distributed Database:**
  A distributed database is a database in which portions of the database are stored in multiple physical locations, and in which processing is dispersed or replicated among different points in a network.
- **Cloud Database:**
  A cloud database is a database that typically runs on a cloud computing platform. Database service provides access to the database. Database services make the underlying software- stack transparent to the user.

Eg. SQL

Structured **Q**uery **L**anguage or **SQL** is a standard Database language which is used to create, maintain and retrieve the data from relational databases like MySQL, Oracle, SQL Server, PostGre, etc. The recent ISO standard version of SQL is SQL:2019.

As the name suggests, it is used when we have structured data (in the form of tables). All

databases that are not relational (or do not use fixed structure tables to store data) and therefore do not use SQL, are called NoSQL databases. Examples of NoSQL are MongoDB, DynamoDB, Cassandra, etc

Microsoft access

Microsoft Access is a Database Management System (DBMS) from Microsoft that combines the relational Microsoft Jet Database Engine with a graphical user interface and softwaredevelopment tools. It is a member of the Microsoft Office suite of applications, included in the professional and higher editions.

- Microsoft Access is just one part of Microsoft's overall data management product strategy.
- It stores data in its own format based on the Access Jet Database Engine.
- Like relational databases, Microsoft Access also allows you to link related information easily. For example, customer and order data. However, Access 2013 also complements other database products because it has several powerful connectivity features.
- It can also import or link directly to data stored in other applications and databases.
- As its name implies, Access can work directly with data from other sources, including many popular PC database programs, with many SQL (Structured Query Language) databases on the desktop, on servers, on minicomputers, or on mainframes, and with data stored on Internet or intranet web servers.
- Access can also understand and use a wide variety of other data formats, including many other database file structures.
- You can export data to and import data from word processing files, spreadsheets, or database files directly.
- Access can work with most popular databases that support the Open Database Connectivity (ODBC) standard, including SQL Server, Oracle, and DB2.
- Software developers can use Microsoft Access to develop application software.

Drug databases in the practice of pharmacy

A database that provides information on drug toxicity and how specific drugs impact the environment.

DrugBank

The DrugBank database is a comprehensive, freely accessible, online database containing information on drugs and drug targets. As both a bioinformatics and a cheminformatics resource, DrugBank combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. Because of its broad scope, comprehensive referencing and unusually detailed data descriptions, DrugBank is more akin to a drug encyclopedia than a drug database. As a result, links to DrugBank are maintained for nearly all drugs listed in

Wikipedia. DrugBank is widely used by the drug industry, medicinal chemists, pharmacists, physicians, students and the general public. Its extensive drug and drug-target data has enabled the discovery and repurposing of a number of existing drugs to treat rare and newly identified illnesses.

The latest release of DrugBank (version 5.1.5, released 2020-01-03) contains 13,551 drug entries including 2,629 approved small molecule drugs, 1,372 approved biologics (proteins, peptides, vaccines, and allergenics), 131 nutraceuticals and over 6,366 experimental (discovery-phase) drugs. Additionally, 5,248 non-redundant protein (i.e. drug target/enzyme/transporter/carrier) sequences are linked to these drug entries. Each entry contains more than 200 data fields with half of the information being devoted to drug/chemical data and the other half devoted to drug target or protein data.

DrugBank is offered to the public as a freely available resource. Use and re-distribution of the data, in whole or in part, for commercial purposes (including internal use) requires a license. We ask that users who download significant portions of the database cite the DrugBank paper in any resulting publications.

References

1. https://study.com/academy/lesson/pharmacy-drug-databases-web-resources.html
2. https://www.drugbank.ca/about

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF PHARMACY

UNIT – III COMPUTER APPLICATIONS IN PHARMACY – BP205T

PHARMACY - UNIT – III

Application of computers in Pharmacy – Drug information storage and retrieval, Pharmacokinetics, Mathematical model in Drug design, Hospital and Clinical Pharmacy, Electronic Prescribing and discharge (EP) systems, barcode medicine identification and automated dispensing of drugs, mobile technology and adherence monitoring Diagnostic System, Lab-diagnostic System, Patient Monitoring System, Pharma Information System

Pharmacokinetics

Pharmacokinetics, sometimes described as what the body does to a drug, refers to the movement of drug into, through, and out of the body—the time course of its absorption, bioavailability, distribution, metabolism, and excretion.

Pharmacodynamics, described as what a drug does to the body, involves receptor binding, postreceptor effects, and chemical interactions. Drug pharmacokinetics determines the onset, duration, and intensity of a drug's effect. Formulas relating these processes summarize the pharmacokinetic behavior of most drugs.

Pharmacokinetics of a drug depends on patient-related factors as well as on the drug's chemical properties. Some patient-related factors (eg, renal function, genetic makeup, sex, age) can be used to predict the pharmacokinetic parameters in populations. For example, the half-life of some
drugs, especially those that require both metabolism and excretion, may be remarkably long in the elderly. In fact, physiologic changes with aging affect many aspects of pharmacokinetics.

Other factors are related to individual physiology. The effects of some individual factors (eg, renal failure, obesity, hepatic failure, dehydration) can be reasonably predicted, but other factors are idiosyncratic and thus have unpredictable effects. Because of individual differences, drug administration must be based on each patient's needs—traditionally, by empirically adjusting dosage until the therapeutic objective is met. This approach is frequently inadequate because it can delay optimal response or result in adverse effects.

Knowledge of pharmacokinetic principles helps prescribers adjust dosage more accurately and rapidly. Application of pharmacokinetic principles to individualize pharmacotherapy is termed therapeutic drug monitoring.

Drug Absorption

Drug absorption is determined by the drug's physicochemical properties, formulation, and route of administration. Dosage forms (eg, tablets, capsules, solutions), consisting of the drug plus other ingredients, are formulated to be given by various routes (eg, oral, buccal, sublingual, rectal, parenteral, topical,

inhalational). Regardless of the route of administration, drugs must be in solution to be absorbed. Thus, solid forms (eg, tablets) must be able to disintegrate and deaggregate.

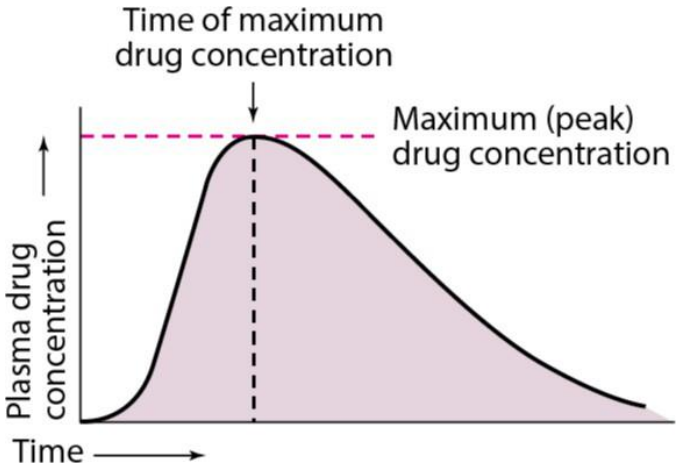
Unless given IV, a drug must cross several semipermeable cell membranes before it reaches the systemic circulation. Cell membranes are biologic barriers that selectively inhibit passage of drug molecules. The membranes are composed primarily of a bimolecular lipid matrix, which determines membrane permeability characteristics. Drugs may cross cell membranes by

- Passive diffusion
- Facilitated passive diffusion
- Active transport
- Pinocytosis

Drug Bioavailability

Bioavailability refers to the extent and rate at which the active moiety (drug or metabolite) enters systemic circulation, thereby accessing the site of action.

Bioavailability of a drug is largely determined by the properties of the dosage form, which depend partly on its design and manufacture. Differences in bioavailability among formulations of a given drug can have clinical significance; thus, knowing whether drug formulations are equivalent is essential.



Plasma drug concentration increases with extent of absorption; the maximum (peak) plasma concentration is reached when drug elimination rate equals absorption rate. Bioavailability determinations based on the peak plasma concentration can be misleading because drug elimination begins as soon as the drug enters the bloodstream. Peak time (when maximum plasma drug concentration occurs) is the most widely used general index of absorption rate; the slower the absorption, the

later the peak time.

Drug Distribution to Tissues

After a drug enters the systemic circulation, it is distributed to the body's tissues. Distribution is generally uneven because of differences in blood perfusion, tissue binding (eg, because of lipid content), regional pH, and permeability of cell membranes.

The entry rate of a drug into a tissue depends on the rate of blood flow to the tissue, tissue mass, and partition characteristics between blood and tissue. Distribution equilibrium (when entry and exit rates are the same) between blood and tissue is reached more rapidly in richly vascularized areas, unless diffusion across cell membranes is the rate-limiting step. After equilibrium, drug concentrations in tissues and in extracellular fluids are reflected by the plasma concentration. Metabolism and excretion occur simultaneously with distribution, making the process dynamic and complex.

After a drug has entered tissues, drug distribution to the interstitial fluid is determined primarily by perfusion. For poorly perfused tissues (eg, muscle, fat), distribution is very slow, especially if the tissue has a high affinity for the drug.

Drug Metabolism

The liver is the principal site of drug metabolism. Although metabolism typically inactivates drugs, some drug metabolites are pharmacologically active—sometimes even more so than the parent compound. An inactive or weakly active substance that has an active metabolite is called a prodrug, especially if designed to deliver the active moiety more effectively.

Drugs can be metabolized by oxidation, reduction, hydrolysis, hydration, conjugation, condensation, or isomerization; whatever the process, the goal is to make the drug easier to excrete. The enzymes involved in metabolism are present in many tissues but generally are more concentrated in the liver. Drug metabolism rates vary among patients. Some patients metabolize a drug so rapidly that therapeutically effective blood and tissue concentrations are not reached; in others, metabolism may be so slow that usual doses have toxic effects. Individual drug metabolism rates are influenced by genetic factors, coexisting disorders (particularly chronic liver disorders and advanced heart failure), and drug interactions (especially those involving induction or inhibition of metabolism).

Drug Excretion

The kidneys are the principal organs for excreting water-soluble substances. The biliary system contributes to excretion to the degree that drug is not reabsorbed from the GI tract. Generally, the contribution of intestine, saliva, sweat, breast milk, and lungs to excretion is small, except for exhalation of volatile anesthetics. Excretion via breast milk may affect the breastfeeding infant

Hepatic metabolism often increases drug polarity and water solubility. The resulting metabolites are then more readily excreted.

Discuss the various applications of computers in pharmacy.

Computers in pharmacy are used for the information of drug data, records and files, drug management (creating, modifying, adding and deleting data in patient files to generate reports), business details. The field of pharmacy is awe fully benefitted by use of computers getting and comparing the information to yield an accurate study. In field of operation like new drug discovery, drug design analysis, and manufacturing of drugs and in hospital pharmacy computers are widely used. The drug discovery, designing, manufacturing and analysis have become virtually possible only through the development of upcoming various hard wares and soft wares. Receiving the details, storing it and processing it and its dissemination is the main role of computers and this continuous flow of information shows effective functioning of any system.

Applications of Computers in Pharmacy

1. Usage of computers in the retail pharmacy
2. Computer aided design of drugs (CADD)
3. Use of Computers in Hospital Pharmacy
4. Data storage and retrieval
5. Information system in Pharmaceutical Industry
6. Diagnostic laboratories
7. Computer aided learning
8. Clinical trial management
9. Adverse drug events control
10. Computers in pharmaceutical formulations
11. Computers in Toxicology and Risk Assessment
12. Computational modeling of drug disposition
13. Recent development in bio computation of drug development
14. In Research Publication
15. Digital Libraries

Usage of computers in the retail pharmacy

- Providing a receipt for the patient
- Record of transaction of money
- Ordering low quantity of products via electronic transitions
- Generation of multiple analysis for day, week, month for number of prescription handles and amounts of cash
- Estimation of profits and financial rational analysis

• Printing of billing and payment details
• Inventory control purpose
• Whenever the drugs or medicaments are added to the stock or else removed from stock; the position of stock gets updated instantaneously
• Records of various drug data, i.e., drug data information
• Computers are useful for getting the complete drug information which is used to satisfy the queries
by patients about toxicology, adverse drug reactions, and drug-drug and drug-food interactions.
• Drug Bank Data Base gives complete and detailed description of drug (pharmacological and
pharmaceutical action) and also involves bioinformatics and cheminformation.

Computer aided design of drugs (CADD)
• CADD is referred as a distinct and advanced drug designing process
• It is a process of pronouncement of new medications
• With a base of the refined graphics software existing or feed data the medicinal chemist have a
scope to design the new molecules and improve their efficiency of the action

Use of computers in hospital pharmacy
• In receiving and allotment of drugs
• Storing the details of every individual
• Professional supplies
• Records of dispensed drugs to inpatient and outpatient
• Information of patients records
• Patient monitoring (blood pressure, pulse rate, temperature)

The other applications include -

Data storage and retrieval
Information system in pharmaceutical industry Pharmacoinformatics
Diagnostic laboratories
Computer aided learning Clinical trial management
Computers in pharmaceutical formulations
Computers in toxicology and risk assessment Computational modeling of drug disposition

****

Discuss the phases in drug design and development

Any drug development process must proceed through several stages in order to produce a product that is safe, efficacious, and has passed all regulatory requirements.

Detailed Stages of Drug Development

1. Discovery
2. Product Characterization
3. Formulation, Delivery, Packaging Development
4. Pharmacokinetics And Drug Disposition
5. Preclinical Toxicology Testing And IND Application
6. Bioanalytical Testing
7. Clinical Trials

Discovery

Discovery often begins with target identification – choosing a biochemical mechanism involved in a disease condition. Drug candidates, discovered in academic and pharmaceutical/biotech research labs, are tested for their interaction with the drug target. Up to 5,000 to 10,000 molecules for each potential drug candidate are subjected to a rigorous screening process which can include functional genomics and/or proteomics as well as other screening methods. Once scientists confirm interaction with the drug target, they typically validate that target by checking for activity versus the disease condition for which the drug is being developed. After careful review, one or more lead compounds are chosen.

Product Characterization

When the candidate molecule shows promise as a therapeutic, it must be characterized—the molecule's size, shape, strengths and weaknesses, preferred conditions for maintaining function, toxicity, bioactivity, and bioavailability must be determined. Characterization studies will
undergo analytical method development and validation. Early stage pharmacology studies help to characterize the underlying mechanism of action of the compound.

Formulation, Delivery, Packaging Development

Drug developers must devise a formulation that ensures the proper drug delivery parameters. It is critical to begin looking ahead to clinical trials at this phase of the drug development process. Drug formulation and delivery may be refined continuously until, and even after, the drug's final approval. Scientists determine the drug's stability—in the formulation itself, and for all the parameters involved with storage and shipment, such as heat, light, and time. The formulation must remain potent and sterile; and it must also remain safe (nontoxic). It may also be necessary to
perform leachables and extractables studies on containers or packaging.

Pharmacokinetics And Drug Disposition

Pharmacokinetic (PK) and ADME (Absorption/Distribution/Metabolism/Excretion) studies provide useful feedback for formulation scientists. PK studies yield parameters such as AUC (area under the curve), Cmax (maximum concentration of the drug in blood), and Tmax (time at which Cmax is reached). Later on, this data from animal PK studies is compared to data from early stage clinical trials to check the predictive power of animal models.

Preclinical Toxicology Testing and IND Application

Preclinical testing analyzes the bioactivity, safety, and efficacy of the formulated drug product. This testing is critical to a drug's eventual success and, as such, is scrutinized by many regulatory entities. During the preclinical stage of the development process, plans for clinical trials and an Investigative New Drug (IND) application are prepared. Studies taking place during the preclinical stage should be designed to support the clinical studies that will follow.

Bioanalytical Testing

Bioanalytical laboratory work and bioanalytical method development supports most of the other activities in the drug development process. The bioanalytical work is key to proper characterization of the molecule, assay development, developing optimal methods for cell culture or fermentation, determining process yields, and providing quality assurance and quality control for the entire development process. It is also critical for supporting preclinical toxicology/pharmacology testing and clinical trials.

Clinical Trials

Clinical trials are research investigations in which people volunteer to test new treatments, interventions or tests as a **means** to prevent, detect, treat or manage various diseases or medical conditions. Some investigations look at how people respond to a new intervention* and what side effects might occur

Drug information

It is called drug information, medication information, or drug informatics. It's really the discovery, use, and management of information in the use of medications. Drug information covers the gamut from identification, cost, and pharmacokinetics to dosage and adverse effects. We may also need information about the body, health, or diseases in order to better utilize the drug information.

Classification of Information Sources

Drug information sources have been traditionally classified in three different categories: primary, secondary, and tertiary

PRIMARY SOURCES

Primary literature consists of clinical research studies and reports, both published and

unpublished. Not all literature published in a journal is classified as primary literature, for example, review articles or editorials are not primary literature.

SECONDARY SOURCES

Secondary literature refers to references that either index or abstract the primary literature, with the goal of directing the user to relevant primary literature.

TERTIARY SOURCES

Tertiary sources provide information that has been summarized and distilled by the author or editor to provide a quick easy summary of a topic. Some examples of tertiary resources include textbooks, compendia, review articles in journals, and other general information, such as may be found on the Internet.

The role of a clinical pharmacy

Clinical pharmacy is the branch of pharmacy in which clinical pharmacists provide direct patient care that optimizes the use of medication and promotes health, wellness, and disease prevention. Clinical pharmacists care for patients in all health care settings but the clinical pharmacy movement initially began inside hospitals and clinics. Clinical pharmacists often work in collaboration with physicians, physician assistants, nurse practitioners, and other healthcare professionals. Clinical pharmacists can enter into a formal collaborative practice agreement with another healthcare provider, generally one or more physicians, that allows pharmacists to prescribe medications and order laboratory tests.

Within the system of health care, clinical pharmacists are experts in the therapeutic use of medications. They routinely provide medication therapy evaluations and recommendations to patients and other health care professionals. Clinical pharmacists are a primary source of scientifically valid information and advice regarding the safe, appropriate, and cost-effective use of medications. Clinical pharmacists are also making themselves more readily available to the public. In the past, access to a clinical pharmacist was limited to hospitals, clinics, or educational institutions. However, clinical pharmacists are making themselves available through a medication information hotline, and reviewing medication lists, all in an effort to prevent medication errors in the foreseeable future.

Clinical pharmacists interact directly with patients in several different ways. They use their knowledge of medication (including dosage, drug interactions, side effects, expense, effectiveness, etc.) to determine if a medication plan is appropriate for their patient. If it is not, the pharmacist will consult the primary physician to ensure that the patient is on the proper medication plan. The pharmacist also works to educate their patients on the importance of taking and finishing their medications.

The benefits of E – prescribing

Electronic prescribing (e-prescribing or e-Rx) is the computer-based electronic generation, transmission, and filling of a medical prescription, taking the place of paper and faxed prescriptions. E-prescribing allows a physician, pharmacist, nurse practitioner, or physician assistant to use digital prescription software to electronically transmit a new prescription or renewal authorization to a community or mail-order pharmacy. It outlines the ability to send error-free, accurate, and understandable prescriptions electronically from the healthcare provider to the pharmacy. E-prescribing is meant to reduce the risks associated with traditional prescription script writing. It is also one of the major reasons for the push for electronic medical records. By sharing medical prescription information, e-prescribing seeks to connect the patient's team of healthcare providers to facilitate knowledgeable decision making.

Barcode medication administration

Bar code medication administration (BCMA) is a bar code system designed by Glenna Sue Kinnick to prevent medication errors in healthcare settings and to improve the quality and safety of medication administration. The overall goals of BCMA are to improve accuracy, prevent errors, and generate online records of medication administration.

It consists of a bar code reader, a portable or desktop computer with wireless connection, a computer server, and some software. When a nurse gives medication to a patient in a healthcare setting, the nurse can scan the barcode on the patient's wristband on the patient to verify the patient's identity. The nurse can then scan the bar code on medication and use software to verify that he/she is administering the right medication to the right patient at the right dose, through the right route, and at the right time ("five rights of medication administration").Bar code medication administration was designed as an additional check to aid the nurse in administering medications; however, it cannot replace the expertise and professional judgment of the nurse. The implementation of BCMA has shown a decrease in medication administration errors in the healthcare setting.

The role of automated dispensing in healthcare

Automated dispensing is a pharmacy practice in which a device dispenses medications and fills prescriptions. The most important thing a hospital pharmacy should enforce is patient safety. Wrong drug and wrong dose errors are the most common errors associated with ADC use.

Automated dispensing machines—decentralized medication distribution systems that provide computer-controlled storage, dispensing, and tracking of medications—have been recommended as one potential mechanism to improve efficiency and patient safety, and they are now widely used in many hospitals.

Pharmacist's Role in Medication Adherence

Medication adherence, or taking **medications** correctly, is generally defined as the extent to which patients take **medication** as prescribed by their doctors. This involves factors such as getting prescriptions filled, remembering to take **medication** on time,

and understanding the directions

Pharmacists have a major role in improving medication adherence in patients. They can confirm that patients are on the correct medications and are not taking any other treatments/drugs that may undermine the effectiveness of important therapies.

The use of Mathematical Modeling In Drug Discovery And Development.

In the fields of medicine, biotechnology and pharmacology, drug discovery is the process by which new candidate medications are discovered. Drug discovery is a complex undertaking facing many challenges, not the least of which is a high attrition rate as many promising candidates prove ineffective or toxic in the clinic owing to a poor understanding of the diseases, and thus the biological systems, they target. Therefore, it is broadly agreed that to increase agreed that to increase the productivity of drug discovery one needs a far deeper understanding of the molecular mechanisms of diseases, taking into account the full biological context of the drug target and moving beyond individual genes and proteins. Mathematical methods are increasingly being used in drug discovery to enquire into biological systems, with a view to understanding the behavior in a more holistic way.

Present difficulties in drug development include an increase in cost and duration of drug development, and only few new medical entities reach approval. It takes from 10 to 15 years to bring a new drug to market — at a cost of more than $1 billion. Many new potential drugs fail because researchers lack reliable information about their behavior. That leads to problems for both pharma industry and public health. Moreover, one can observe some lack of interest of drug pharma for some disease areas due to high potential costs of research. Mathematical model based approaches also been suggested to expand the use of simulations in support of clinical drug development for predicting outcomes of planned trials.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF PHARMACY

UNIT – IV COMPUTER APPLICATIONS IN PHARMACY – BP205T

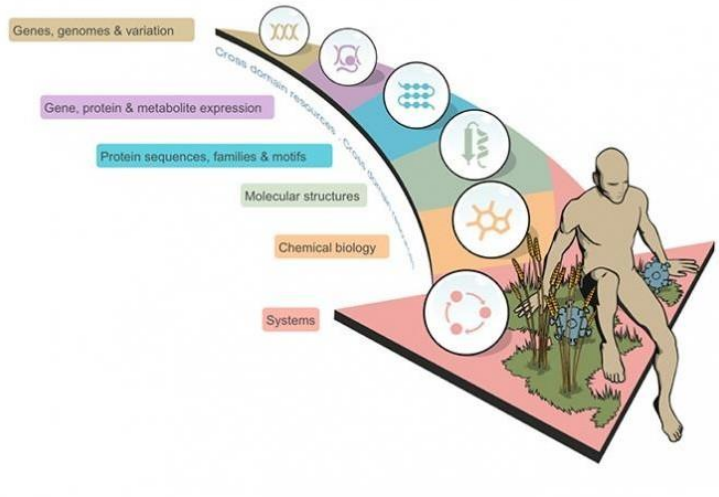
UNIT – IV

Bioinformatics: Introduction, Objective of Bioinformatics, Bioinformatics Databases, Concept of Bioinformatics, Impact of Bioinformatics in Vaccine Discovery

An overview on bioinformatics and its applications

Put simply, bioinformatics is the science of storing, retrieving and analysing large amounts of biological information. It is a highly interdisciplinary field involving many different types of specialists, including biologists, molecular life scientists, computer scientists and mathematicians.

The term bioinformatics was coined by Paulien Hogeweg and Ben Hesper to describe "the study of informatic processes in biotic systems" and it found early use when the first biological sequence data began to be shared. Whilst the initial analysis methods are still fundamental to many large-scale experiments in the molecular life sciences, nowadays bioinformatics is considered to be a much broader discipline, encompassing modelling and image analysis in addition to the classical methods used for comparison of linear sequences or three-dimensional structures.



A broad overview of the different types of data that fall within the scope of bioinformatics. Traditionally, bioinformatics was used to describe the science of storing and analysing biomolecular sequence data, but the term is now used much more broadly, encompassing computational structural biology, chemical biology and systems biology (both data integration and the modelling of systems).

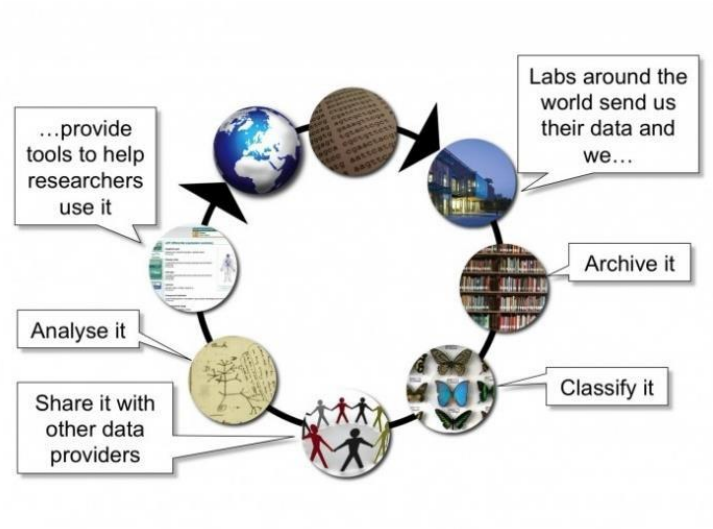
The molecular life sciences have become increasingly data driven by and reliant on data sharing through open-access databases. This is as true of the applied sciences as it is of fundamental research. Furthermore, it is not necessary to be a bioinformatician to make use of bioinformatics databases, methods and tools. However, as the generation of large data-sets becomes more and more central to biomedical research, it's becoming increasingly necessary for every molecular life scientist to understand what can (and, importantly, what cannot) be achieved using bioinformatics, and to be able to work with bioinformatics experts to design, analyse and interpret their experiments.

The role of public databases

There are a small number of bioinformatics centres of excellence worldwide that have taken on the responsibility to collect, catalogue and provide open access to published biological data (Figure 3). Among these centres are:

- The EMBL-European Bioinformatics Institute (EMBL-EBI)
- The US National Center for Biotechnology Information (NCBI)
- The National Institute of Genetics in Japan (NIG)

This work began in the early 1980s when DNA sequence data began to accumulate in the scientific literature. The EMBL Data Library (now the European Nucleotide Archive) was developed to store DNA sequences published in the scientific literature. The NCBI's GenBank and NIG's DDBJ followed.



The role of bioinformatics centres of excellence in making biological data available for the research community.

Goals of Bioinformatics

To study how normal cellular activities are altered in different disease states, the biological data must be combined to form a comprehensive picture of these activities. Therefore, the field of bioinformatics has evolved such that the most pressing task now involves the analysis and interpretation of various types of data. This includes nucleotide and amino acid sequences, protein domains, and protein structures.[16] The actual process of analyzing and interpreting data is referred to as computational biology. Important sub-disciplines within bioinformatics and computational biology include:

- Development and implementation of computer programs that enable efficient access to, management and use of, various types of information
- Development of new algorithms (mathematical formulas) and statistical measures that assess relationships among members of large data sets. For example, there are methods to locate a gene within a sequence, to predict protein structure and/or function, and to cluster protein sequences into families of related sequences.

The primary goal of bioinformatics is to increase the understanding of biological processes. What sets

it apart from other approaches, however, is its focus on developing and applying computationally intensive techniques to achieve this goal. Examples include: pattern recognition, data mining, machine learning algorithms, and visualization. Major research efforts in the field include sequence alignment, gene finding, genome assembly, drug design, drug discovery, protein structure alignment, protein structure prediction, prediction of gene expression and protein–protein interactions, genome-wide association studies, the modeling of evolution and cell division/mitosis.

Bioinformatics now entails the creation and advancement of databases, algorithms, computational and statistical techniques, and theory to solve formal and practical problems arising from the management and analysis of biological data.

****

Biological databases and their uses

Biological databases emerged as a response to the huge data generated by low-cost DNA sequencing technologies. One of the first databases to emerge was GenBank, which is a collection of all available protein and DNA sequences. It is maintained by the National Institutes of Health (NIH) and the National Center for Biotechnology Information (NCBI). GenBank paved the way for the Human Genome Project (HGP). The HGP allowed complete sequencing and reading of the genetic blueprint. The data stored in biological databases is organized for optimal analysis and consists of two types: raw and curated (or annotated). Biological databases are complex, heterogeneous, dynamic, and yet inconsistent.

Why are these Important?

Earlier, databases and databanks were considered quite different. However, over the time, database became a preferable term. Data is submitted directly to biological databases for indexing, organization, and data optimization. They help researchers find relevant biological data by making it available in a format that is readable on a computer. All biological information is readily accessible through data mining tools that save time and resources. Biological databases can be broadly classified as sequence and structure databases. Structure databases are for protein structures, while sequence databases are for nucleic acid and protein sequences.

Kinds of Biological Databases

Biological databases can be further classified as primary, secondary, and composite databases. Primary databases contain information for sequence or structure only. Examples of primary biological databases include:

- Swiss-Prot and PIR for protein sequences
- GenBank and DDBJ for genome sequences
- Protein Databank for protein structures

Secondary databases contain information derived from primary databases. Secondary databases store information such as conserved sequences, active site residues, and signature sequences. Protein Databank data is stored in secondary databases. Examples include:

- SCOP at Cambridge University

    - CATH at the University College of London
    - PROSITE of the Swiss Institute of Bioinformatics
- eMOTIF at Stanford Composite databases contain a variety of primary databases, which eliminates the need to search each one separately. Each composite database has different search algorithms and data structures. The NCBI hosts these databases, where links to the Online Mendelian Inheritance in Man (OMIM) is found.

The Future

Because of high-performance computational platforms, these databases have become important in providing the infrastructure needed for biological research, from data preparation to data extraction. The simulation of biological systems also requires computational platforms, which further underscores the need for biological databases. The future of biological databases looks bright, in part due to the digital world.

In terms of research, bioinformatics tools should be streamlined for analyzing the growing amount of data generated from genomics, metabolomics, proteomics, and metagenomics. Another future trend will be the annotation of existing data and better integration of databases.

With a large number of biological databases available, the need for integration, advancements, and improvements in bioinformatics is paramount. Bioinformatics will steadily advance when problems about nomenclature and standardization are addressed. The growth of biological databases will pave the way for further studies on proteins and nucleic acids, impacting therapeutics, biomedical, and related fields.

****

The role of bioinformatics in drug and vaccine development.

Vaccines are the pharmaceutical products that offer the best cost-benefit ratio in the prevention or treatment of diseases. In that a vaccine is a pharmaceutical product, vaccine development and production are costly and it takes years for this to be accomplished. Several approaches have been applied to reduce the times and costs of vaccine development, mainly focusing on the selection of appropriate antigens or antigenic structures, carriers, and adjuvants.

One of these approaches is the incorporation of bioinformatics methods and analyses into vaccine development. This chapter provides an overview of the application of bioinformatics strategies in vaccine design and development, supplying some successful examples of vaccines in which bioinformatics has furnished a cutting edge in their development. Reverse vaccinology, immunoinformatics, and structural vaccinology are described and addressed in the design and development of specific vaccines against infectious diseases caused by bacteria, viruses, and parasites.

These include some emerging or re-emerging infectious diseases, as well as therapeutic vaccines to fight cancer, allergies, and substance abuse, which have been facilitated and improved by using bioinformatics tools or which are under development based on bioinformatics strategies.

The success of vaccination is reflected in its worldwide impact by improving human and veterinary health and life expectancy. It has been asserted that vaccination, as well as clean water, has had such a major effect on mortality reduction and population growth. In addition to the invaluable role of

traditional vaccines to prevent diseases, the society has observed remarkable scientific and technological progress since the last century in the improvement of these vaccines and the generation of new ones.

This has been possible by the fusion of computational technologies with the application of recombinant DNA technology, the fast growth of biological and genomic information in database banks, and the possibility of accelerated and massive sequencing of complete genomes. This has aided in expanding the concept and application of vaccines beyond their traditional immunoprophylactic function of preventing infectious diseases, and also serving as therapeutic products capable of modifying the evolution of a disease and even cure it.

Vaccines are the pharmaceutical products that offer the best cost-benefit ratio in the prevention or treatment of diseases. In that it is a pharmaceutical product, a vaccine development and production are costly and it takes years for this to be accomplished. Several approaches have been applied to reduce the times and costs of their development, mainly focusing on the selection of appropriate antigens or antigenic structures, carriers, and adjuvants. One of these approaches is the incorporation of bioinformatics methods and analyses into vaccine development.

At present, there are many alternative strategies to design and develop effective and safe new-generation vaccines, based on bioinformatics approaches through reverse vaccinology, immunoinformatics, and structural vaccinology.

Reverse vaccinology

Reverse vaccinology is a methodology that uses bioinformatics tools for the identification of structures from bacteria, virus, parasites, cancer cells, or allergens that could induce an immune response capable of protecting against a specific disease

Immunoinformatics

The immunological system can be classified as cellular or humoral and, depending on the disease, it can be induced the expected immune response. If a vaccine that induces a cellular response is needed, for example a tuberculosis vaccine or a parasite vaccine against leishmaniasis [23], the software must search for antigens that can be recognized by the major histocompatibility complex (MHC) molecules present in T lymphocytes. Software for this purpose include TEpredict, CTLPred, nHLAPred, ProPred-I, MAPPP, SVMHC, GPS-MBA, PREDIVAC, NetMHC, NetCTL, MHC2 Pred, IEDB,
BIMAS, SVMHC, POPI, Epitopemap, iVAX, FRED2, Rankpep, BIMAS, PickPocket, KISS, and MHC2MIL.

Structural vaccinology

Structural vaccinology focuses on the conformational features of macromolecules, mainly proteins that make them good candidate antigens. This approach to vaccine design has been used mainly to select or design peptide-based vaccines or cross-reactive antigens with the capability of generating immunity against different antigenically divergent pathogens.

****

A brief timeline of the major events in the history and the origins of bioinformatics.

A Chronological Hlstory of Blolnformatlcs

- 19S3 - Watson & Crick pmposed the double helix model for DNA based x-ray data obtained by Franklin & Wilkins.
- 19S4 - Perutz's group develop heavy atom methods to solve the phase problem in protein crystallography.
- 19S5 - The sequence of the first pmtein to be analysed, bovine insulin, is announed by

- 19G9 - The ARPANET is created by linking computers at Standford and UCLA.
- **1970** - The details of the Needleman-Wiinsch algorithm for sequence comparison are published.
- 1972 - The first recombinant DNA molecule is created by Paul Berg and his gmup.
- 1973 - The Brookhaven Pmtein DataBank is annoiineced (Acta.Ctyst.B,1973,29:1764).
  Roberi Metcalfe receives his Ph.D finn Harvard University. His thesis describes Ethernet.
- 1974 - Vint Cerf and Robert Khan develop the concept of connecting networks of computers into an "internet" and develop the Transmission Control Protocol {TCP}.
- **1975** - Microsofi Corporation is founded by Bill Gates and Paul Allen.
  Two-dimensional electrophoresis, where separation of proteins on SDS polyacrylamide gel is combined with separation according to isoelectric points, is announced by P.H.OTarre1.
- 1988 - The National Centre for Biotechnology Information {NCBI) is established at the National Cancer Institute.
  The Human Genome Intiative is started (commission on Life **Sciences,** National Research council. Mapping and sequencing the Human Genome, National Academy Press: wahington, D.C.), 1988.
  The FASTA algorith for sequence comparison is published by Pearson and Lupman. A new pmgram, an Internet computer vinis defined by a student, infects 6,000 military computers in the US.

- 1989 - The genetics Computer Group (GCG) becomes a privatae company.
  Oxford Molecular Gmup,Ltd.{OMG) founded, UK by Anthony Marchigton, David Ricketts, James Hiddleston, Anthony Rsss, and W.Graham Richards. Primary products: Anaconds, Asp, Cameleon and others (molecular modeling, drug design, protein design).
- **1990** - The BLAST program (Altschul,st.al.) is implemented.
  Molecular applications group is founded in California by Michael Levitt and Chris Lee. Their primary pmducts are Look and SegMod which are used for molecular modeling and protein deisign.
  InforMax is founded in Bsthesda, MD. The company's products address sequence analysis, database and data management, searching, publication graphics, clone oonsbuction, trapping and primsr design.
- **1991** - The research institute in Geneva (CERN} announces the creation of the protocols which make -up the World Wide Web.
  The creation and use of expressed sequence tags (ESTs) is described.
  Incyte Pharmaceuticals, a genomics company headqiiartered in Palo Alto California, is

formed.
Myriad Genetics, Inc. is founded in Utah. The company's goal is to lead in the discovery of major common human disease genes and their related pathways. The company has discovered and sequenced, with its academic cnllaboratnrs, the

following major genes: BRCA1, BRACA1 , CHD1, MMAC1, MMSC1, MMSC2, CtIP, p16, p19 and MTS2.

- **1993** - CuraGen Corporation is formed in New Haven, CT.
  Affymetrix begins independent operations in Santa Clara, California.
- **1994** - Netscape Communications Corporation founded and releases Naviagator, the commerical version of NCSA's Mozilla.
  Gene Logic is formed in Maryland.
  The PRINTS database of protein motifs is published by Attwood and Beck.
  Oxford Molecular Group acquires IntelliGenetics.
- **1995** - The Haemophilus influenzea genome (1.8) is sequenced.
  The Mycoplasma genitalium genome is sequenced.
- **1996** - The genome for Saccharomyces cerevisiae (baker's yeadt, 12.1 Mb) is sequenced.
  The prosite database is reported by Bairoch, et.al.
  Affymetrix produces the first commerical DNA chips.
- **1997** - The genome for E.coli (4.7 Mbp) is published.Oxford Molecualr Group acquires the Genetics Computer Group. LION bioscience AG founded as an intergrated genomics company with strong focus on bioinformatics. The company is built from IP out of the European Molecualr Biology Laboratory (EMBL), the European Bioinformtics Institute (EBI), the GErman Cancer Research Center (DKFZ), and the University of Heidelberg.paradigm Genetics Inc., a company focussed on the application of genomic technologies to enhance worldwide food and fiber production, is founded in Research Triangle Park, NC. deCode genetics publishes a paper that described the location of the FET1 gene, which is responsible for familial essential tremor, on chromosome 13 (Nature Genetics).

- **1998** - The genomes for Caenorhabitis elegans and baker's yeast are published.The Swiss Institute of Bioinformatics is established as a non-profit foundation.Craig Venter forms Celera in Rockville, Maryland. PE Informatics was formed as a center of Excellence within PE Biosystems. This center brings together and leverges the complementary expertise of PE Nelson and Molecualr Informatics, to further complement the genetic instrumention expertise of Applied Biosystems.Inpharmatica, a new Genomics and Bioinformatics company, is established by University College London, the Wolfson Institute for Biomedical Research, five leading scientists from major British academic centres and Unibio Limited. GeneFormatics, a company dedicated to the analysis and predication of protein structure and function, is formed in San Diego.Molecualr Simulations Inc. is acquired by Pharmacopeia.
- **1999** - deCode genetics maps the gene linked to pre-eclampsia as a locus on chromosome 2p13.
- **2000** - The genome for Pseudomonas aeruginosa (6.3 Mbp) is published. The Athaliana genome (100 Mb) is secquenced.The D.melanogaster genome (180 Mb) is sequenced.Pharmacopeia acquires Oxoford Molecular Group.
- **2001** - The huam genome (3,000 Mbp) is published.

****

Nucleic acid and protein databases with an example.

The **Nucleic Acid Database** (NDB) (http://ndbserver.rutgers.edu) is a web portal providing access to information about 3D **nucleic acid** structures and their complexes.

Protein sequence databases Introduction: The **Protein database** is a collection of **sequences** from several sources, including translations from annotated coding regions in GenBank, RefSeq and TPA, as well as records from SwissProt, PIR, PRF, and PDB.

DNA databases

Primary databases

International Nucleotide Sequence Database (INSD) consists of the following databases.

- DNA Data Bank of Japan (National Institute of Genetics)

- EMBL (European Bioinformatics Institute)
- GenBank (National Center for Biotechnology Information)

DDBJ (Japan), GenBank (USA) and European Nucleotide Archive (Europe) are repositories for nucleotide sequence data from all organisms. All three accept nucleotide sequence submissions, and then exchange new and updated data on a daily basis to achieve optimal synchronisation between them. These three databases are primary databases, as they house original sequence data. They collaborate with Sequence Read Archive (SRA), which archives raw reads from high-throughput sequencing instruments.

Secondary databases

- 23andMe's database
- HapMap
- OMIM (Online Mendelian Inheritance in Man): inherited diseases
- RefSeq
- 1000 Genomes Project: launched in January 2008. The genomes of more than a thousand anonymous participants from a number of different ethnic groups were analyzed and made publicly available.
- EggNOG Database: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. It provides multiple sequence alignments and maximum-likelihood trees, as well as broad functional annotation.[

RNA databases

- miRBase: the microRNA database
- Rfam: a database of RNA families

Amino acid / protein databases

Protein sequence databases

- Database of Interacting Proteins (Univ. of California)
- DisProt: database of experimental evidences of disorder in proteins (Indiana University School of Medicine, Temple University, University of Padua)
- InterPro: classifies proteins into families and predicts the presence of domains and sites
- MobiDB: database of intrinsic protein disorder annotation (University of Padua)
- neXtProt: a human protein-centric knowledge resource
- Pfam: protein families database of alignments and HMMs (Sanger Institute)
- PRINTS: a compendium of protein fingerprints from (Manchester University)
- PROSITE: database of protein families and domains
- Protein Information Resource (Georgetown University Medical Center [GUMC])
- SUPERFAMILY: library of HMMs representing superfamilies and database of (superfamily and family) annotations for all completely sequenced organisms
- Swiss-Prot: protein knowledgebase (Swiss Institute of Bioinformatics)
- NCBI: protein sequence and knowledgebase (National Center for Biotechnology Information)

Protein structure databases

- Protein Data Bank (PDB), comprising:
    - Protein DataBank in Europe (PDBe)
    - ProteinDatabank in Japan (PDBj)
    - Research Collaboratory for Structural Bioinformatics (RCSB)
- Structural Classification of Proteins (SCOP)

****

Genome annotation and its importance

DNA annotation or genome annotation is the process of identifying the locations of genes and all of the coding regions in a genome and determining what those genes do. An annotation (irrespective of the context) is a note added by way of explanation or commentary. Once a genome is sequenced, it needs to be annotated to make sense of it.

Process

Genome annotation consists of three main steps:.

1. identifying portions of the genome that do not code for proteins
2. identifying elements on the genome, a process called gene prediction
3. attaching biological information to these elements

Automatic annotation tools attempt to perform these steps via computer analysis, as opposed to manual annotation (a.k.a. curation) which involves human expertise. Ideally, these approaches co-exist and complement each other in the same annotation pipeline.

A simple method of gene annotation relies on homology based search tools, like BLAST, to search for homologous genes in specific databases, the resulting information is then used to annotate genes and genomes. However, as information is added to the annotation platform, manual annotators become capable of deconvoluting discrepancies between genes that are given the same annotation. Some databases use genome context information, similarity scores, experimental data, and integrations of other resources to provide genome annotations through their Subsystems approach. Other databases (e.g. Ensembl) rely on curated data sources as well as a range of different software tools in their automated genome annotation pipeline

Bioinformatics in understanding molecular evolution

Molecular evolution is the process of change in the sequence composition of cellular molecules such as DNA, RNA, and proteins across generations. The field of molecular evolution uses principles of evolutionary biology and population genetics to explain patterns in these changes.

Molecular systematics is the product of the traditional fields of systematics and molecular genetics. It uses DNA, RNA, or protein sequences to resolve questions in systematics, i.e. about their correct scientific classification or taxonomy from the point of view of evolutionary biology.

Molecular systematics has been made possible by the availability of techniques for DNA sequencing, which allow the determination of the exact sequence of nucleotides or *bases* in either DNA or RNA. At present it is still a long and expensive process to sequence the entire genome of an organism, and this has been done for only a few species. However, it is quite feasible to determine the sequence of a defined area of a particular chromosome. Typical molecular systematic

analyses require the sequencing of around 1000 base pairs.

****

Bioinformatics help in understanding gene regulation

Gene regulation is the complex orchestration of events by which a signal, potentially an extracellular signal such as a hormone, eventually leads to an increase or decrease in the activity of one or more proteins. Bioinformatics techniques have been applied to explore various steps in this process.

For example, gene expression can be regulated by nearby elements in the genome. Promoter analysis involves the identification and study of sequence motifs in the DNA surrounding the coding region of a gene. These motifs influence the extent to which that region is transcribed into mRNA. Enhancer elements far away from the promoter can also regulate gene expression, through three-dimensional looping interactions. These interactions can be determined by bioinformatic analysis of chromosome conformation capture experiments.

Expression data can be used to infer gene regulation: one might compare microarray data from a wide variety of states of an organism to form hypotheses about the genes involved in each state. In a single-cell organism, one might compare stages of the cell cycle, along with various stress conditions (heat shock, starvation, etc.). One can then apply clustering algorithms to that expression data to determine which genes are co-expressed. For example, the upstream regions (promoters) of co-expressed genes can be searched for over-represented regulatory elements. Examples of clustering algorithms applied in gene clustering are k-means clustering, self-organizing maps (SOMs), hierarchical clustering, and consensus clustering methods.

****

OMIM (Online Mendelian Inheritance in Man)

Online Mendelian Inheritance in Man (OMIM) is a continuously updated catalog of human genes and genetic disorders and traits, with a particular focus on the gene-phenotype relationship. As of 28 June 2019, approximately 9,000 of the over 25,000 entries in OMIM represented phenotypes; the rest represented genes, many of which were related to known phenotypes.

OMIM is the online continuation of Dr. Victor A. McKusick's *Mendelian Inheritance in Man* (MIM), which was published in 12 editions between 1966 and 1998. Nearly all of the 1,486 entries in the first edition of MIM discussed phenotypes.

MIM/OMIM is produced and curated at the Johns Hopkins School of Medicine (JHUSOM). OMIM became available on the internet in 1987 under the direction of the Welch Medical Library at JHUSOM with financial support from the Howard Hughes Medical Institute. From 1995 to 2010, OMIM was available on the World Wide Web with informatics and financial support from the National Center for Biotechnology Information. The current OMIM website (OMIM.org), which was developed with funding from JHUSOM, is maintained by Johns Hopkins University with financial support from the National Human Genome Research Institute.

****

The importance of PUBMED

PubMed is a free search engine accessing primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics. The United States National Library of Medicine (NLM) at the National Institutes of Health maintain the database as part of the Entrez system of information retrieval.

From 1971 to 1997, online access to the MEDLINE database had been primarily through institutional facilities, such as university libraries. PubMed, first released in January 1996, ushered in the era of private, free, home- and office-based MEDLINE searching. The PubMed system was offered free to the public starting in June 1997.

Content

In addition to MEDLINE, PubMed provides access to:

- older references from the print version of *Index Medicus*, back to 1951 and earlier
- references to some journals before they were indexed in Index Medicus and MEDLINE, for instance *Science*, *BMJ*, and *Annals of Surgery*
- very recent entries to records for an article before it is indexed with Medical Subject Headings (MeSH) and added to MEDLINE
- a collection of books available full-text and other subsets of NLM records
- PMC citations
- NCBI Bookshelf

Many PubMed records contain links to full text articles, some of which are freely available, often in PubMed Central and local mirrors, such as UK PubMed Central.

Information about the journals indexed in MEDLINE, and available through PubMed, is found in the NLM Catalog.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF PHARMACY

UNIT – V COMPUTER APPLICATIONS IN PHARMACY – BP205T

UNIT – V

Computers as data analysis in Preclinical development: Chromatographic dada analysis(CDS), Laboratory Information management System (LIMS) and Text Information Management System(TIMS)

An overview of data and analysis methods using computers in healthcare.

Information has been the key to a better organization and new developments. The more information we have, the more optimally we can organize ourselves to deliver the best outcomes. That is why data collection is an important part for every organization. We can also use this data for the prediction of current trends of certain parameters and future events. As we are becoming more and more aware of this, we have started producing and collecting more data about almost everything by introducing technological developments in this direction. Today, we are facing a situation wherein we are flooded with tons of data from every aspect of our life such as social activities, science, work, health, etc. In a way, we can compare the present situation to a data deluge. The technological advances have helped us in generating more and more data, even to a level where it has become unmanageable with currently available technologies. This has led to the creation of the term 'big data' to describe data that is large and unmanageable. In order to meet our present and future social needs, we need to develop new strategies to organize this data and derive meaningful information. One such special social need is healthcare. Like every other industry, healthcare organizations are producing data at a tremendous rate that presents many advantages and challenges at the same time. In this review, we discuss about the basics of big data including its management, analysis and future prospects especially in healthcare sector.

'Big data' is massive amounts of information that can work wonders. It has become a topic of special interest for the past two decades because of a great potential that is hidden in it. Various public and private sector industries generate, store, and analyze big data with an aim to improve the services they provide. In the healthcare industry, various sources for big data include hospital records, medical records of patients, results of medical examinations, and devices that are a part of internet of things. Biomedical research also generates a significant portion of big data relevant to public healthcare.

This data requires proper management and analysis in order to derive meaningful information. Otherwise, seeking solution by analyzing big data quickly becomes comparable to finding a needle in the haystack. There are various challenges associated with each step of handling big data which can only be surpassed by using high-end computing solutions for big data analysis. That is why, to provide relevant solutions for improving public health, healthcare providers are required to be fully equipped with appropriate infrastructure to systematically generate and analyze big data.

An efficient management, analysis, and interpretation of big data can change the game by opening new avenues for modern healthcare. That is exactly why various industries, including the healthcare industry, are taking vigorous steps to convert this potential into better services and financial advantages. With a strong integration of biomedical and healthcare data, modern healthcare organizations can possibly revolutionize the medical therapies and personalized medicine.

Chromatography Data System

Chromatography is a laboratory technique for the separation of a mixture. The mixture is dissolved in a fluid called the *mobile phase,* which carries it through a structure holding another material called the *stationary phase.* The various constituents of the mixture travel at different speeds, causing them to separate. The separation is based on differential partitioning between the mobile and stationary phases. Subtle differences in a compound's partition coefficient result in differential retention on the stationary phase and thus affect the separation.

Chromatography may be preparative or analytical. The purpose of preparative chromatography is to separate the components of a mixture for later use, and is thus a form of purification. Analytical chromatography is done normally with smaller amounts of material and is for establishing the presence or measuring the relative proportions of analytes in a mixture.

Sometimes referred to as a chromatography data management system (CDMS), a chromatography data system(CDS) is a set of dedicated data-collection tools that interface and/or integrate with a laboratory's chromatography equipment. A base CDS will set up a desired methodology to be used by the chromatography equipment, acquire data from it, process the acquired data, store the information in a database, and interface with other laboratory informatics systems to import and export files and data.

A CDS may be set up for use in three primary ways:

- as a standalone system that controls two or more chromatographs
- as a standalone system that controls a single chromatograph, including LC-MS or GC-MS instruments
- as a networked system that controls multiple instruments in one or more labs

****

LIMS – Its benefits & advantages

A Laboratory Information Management System (LIMS) is software that allows you to effectively manage samples and associated data. By using a LIMS, your lab can automate workflows, integrate instruments, and manage samples and associated information.

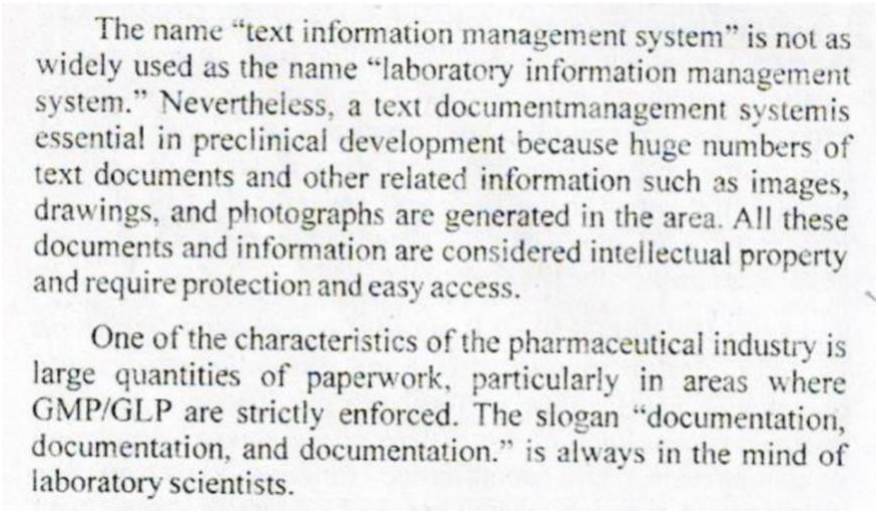
Key advantages of using a LIMS

A Laboratory Information Management System offers a multitude of benefits in terms of laboratory data management. Some of the key functional benefits of a LIMS are:

1. Sample management wherein a user can efficiently track samples through the laboratory and allocate storage locations that mimic the sample storage hierarchy.
2. Workflow automation that leads to a decrease in possible human errors by eliminating manual entry of data.
3. Configurable user interface to meet the unique requirements of different laboratories and mirror their existing workflows.
4. Secure and restricted access to the data leading to better data privacy and protection.
5. Easy data backup and data mining options, resolving data accessibility issues.
6. User-role based access distribution to mirror the real-time laboratory personnel hierarchy.
7. Ease of reporting, wherein an authorized user can quickly generate reports pertaining to (a) the various tests performed, and (b) data required for auditing and quick analysis (for example, the total number of samples logged during a particular period or from a particular region).
8. Streamlined billing process by generating invoices and integrating with the various payment portals.

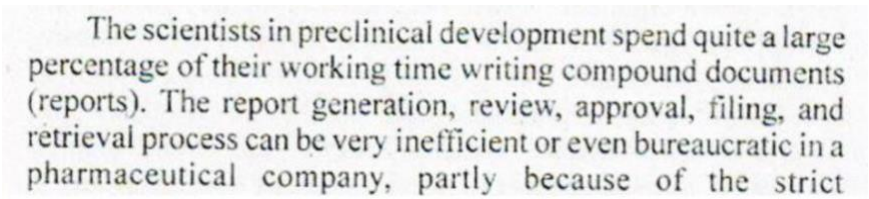
****

Text information management systems



The name "text information management system" is not as widely used as the name "laboratory information management system." Nevertheless, a text documentmanagement systemis essential in preclinical development because huge numbers of text documents and other related information such as images, drawings, and photographs are generated in the area. All these documents and information are considered intellectual property and require protection and easy access.

One of the characteristics of the pharmaceutical industry is large quantities of paperwork, particularly in areas where GMP/GLP are strictly enforced. The slogan "documentation, documentation, and documentation." is always in the mind of laboratory scientists.



The scientists in preclinical development spend quite a large percentage of their working time writing compound documents (reports). The report generation, review, approval, filing, and retrieval process can be very inefficient or even bureaucratic in a pharmaceutical company, partly because of the strict

regulations. The following scenario could be seen often as recently as the late 1980s: The scientist would prepare his report with one type or another of text and graphic software, often through multiple cut-and-paste procedures to include pictures or images. Then the scientist would make hard copies of the report for review by managers and the department head. After all the corrections were made, the scientist would print out another copy for the QA auditor for auditing (this is only done for the documents used for submission). It could take months before the report was finally ready to be filed in the company record center, where photocopies and microfilms were made and indexing took place. When an end user needed a copy of the report, he would have to make a request to the record center for a hard copy.

When TIMS is used in today's workflow, the scientist can use a report template to facilitate report writing. Some cut-and-paste procedures are still needed to include data and figures. After the draft report is completed, the scientist can send the reviewers an electronic link for the document. The reviewers can review the document and make changes and corrections with the "tracking change" function. When the review is completed, the author can choose to accept the changes or deny them. If auditing is needed, the same process can be used. The finalized document is issued within the TIMS by adding an issue date and signatures, if necessary, and converting into an unalterable PDF file. Future changes made after issuance are captured through version control. End users can also access the issued document electronically and remotely. Comparison of the new process vs. the old one has demonstrated the advantages of TIMS.

****

Chromatography and its types

The twelve types are: (1) Column Chromatography (2) Paper Chromatography (3) Thin Layer Chromatography (4) Gas Chromatography (5) High Performance Liquid Chromatography (6) Fast Protein Liquid Chromatography (7) Supercritical Fluid Chromatography (8) Affinity Chromatography
(9) Reversed Phase Chromatography (10) Two Dimensional Chromatography (11) Pyrolysis Gas Chromatography and (12) Counter Current Chromatography.

There are different kinds of chromatographic techniques and these are classified according to the shape of bed, physical state of mobile phase, separation mechanisms. Apart from these there are certain modified forms of these chromatographic techniques involving different mechanisms and are hence categorized as modified or specialized chromatographic techniques.

Laboratory information

A laboratory information system (LIS) is a software system that records, manages, and stores data for clinical laboratories. An LIS has traditionally been most adept at sending laboratory test orders to lab instruments, tracking those orders, and then recording the results, typically to a searchable database.

Components of LIMS
Components may
include:

- Electronic lab notebooks.
- Sample management programs.
- Process execution software.

- Records management software.
- Applications to interface with analytical instruments or data systems.
- Workflow tools.
- Client tracking applications.
- Best practice and compliance databases.

****

Preclinical studies in drug development

In drug development, **preclinical development**, also named **preclinical studies** and **nonclinical studies**, is a stage of research that begins before clinical trials (testing in humans) can begin, and during which important feasibility, iterative testing and drug safety data are collected.

The main goals of pre-clinical studies are to determine the safe dose for first-in-man study and assess a product's safety profile. Products may include new medical devices, drugs, gene therapy solutions and diagnostic tools. On average, only one in every 5,000 compounds that enters drug discovery to the stage of preclinical development becomes an approved drug.

Preclinical studies refer to the testing of a drug, procedure or other medical treatment in animals before trials may be carried out in humans. During preclinical drug development, the drug's toxic and pharmacological effects need to be evaluated through in vitro and in vivo laboratory animal testing.

Why are preclinical studies important?

The most important role of preclinical pharmacology studies is to identify the starting dose for Phase I clinical trials. In these studies, the safety profiles of lead compounds are evaluated through a battery of assessment assays adapted to determining side effects of new agents.

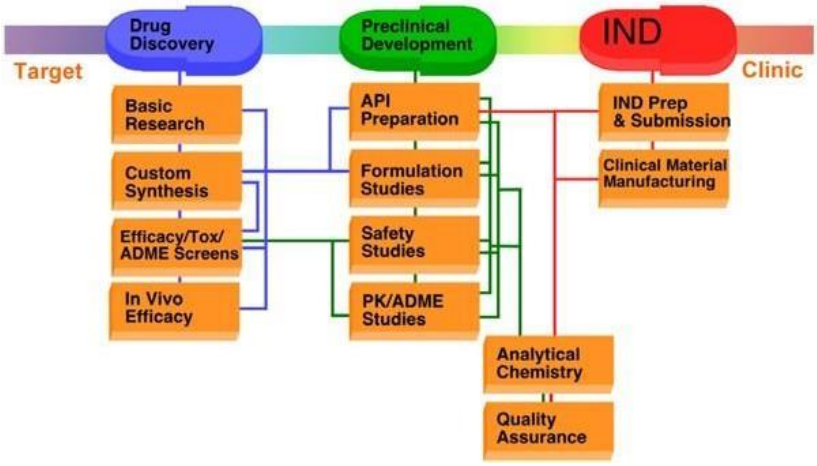
Types of data generated in a hospital environment

Health data is any data "related to health conditions, reproductive outcomes, causes of death, and quality of life"[1] for an individual or population. Health data includes clinical metrics along with environmental, socioeconomic, and behavioral information pertinent to health and wellness. A plurality of health data are collected and used when individuals interact with health care systems. This data, collected by health care providers, typically includes a record of services received, conditions of those services, and clinical outcomes or information concerning those services. Historically, most health data have been sourced from this framework. The advent of e Health and advances in health information technology, however, have expanded the collection and use of health data—but have also engendered new security, privacy, and ethical concerns. The increasing collection and use of health data by patients is a major component of digital health

Health data are classified as either structured or unstructured. Structured health data are standardized and easily transferable between health information systems. For example, a patient's name, date of birth, or a blood-test result can be recorded in a structured data format. Unstructured health data, unlike structured data, are not standardized. Emails, audio recordings, or physician notes about a patient are examples of unstructured health data.

While advances in health information technology have expanded collection and use, the complexity of health data has hindered standardization in the health care industry.

The standard operating procedures in preclinical development



Preclinical drug development stages. Following identification of a drug target and candidate compounds, several early activities, such as pharmacology, *in vivo* efficacy, and experimental toxicology, can contribute to the selection of a lead candidate for preclinical development. These preclinical activities provide the basis for an Investigational New Drug (IND) application to the FDA for permission to initiate clinical testing in humans. ADME, absorption, distribution, metabolism, and excretion; API, active pharmaceutical ingredient; PK, pharmacokinetics; Prep, preparation; Tox, toxicity.

Drug development is time consuming and costly which contains preclinical, clinical and after-market. In principle, if all the processes are straight-forward, a drug can be developed in a seven year period. In practice, drug development takes in excess of twelve years. Procedures are tightly regulated both for safety and to ensure drugs are effective. Of the many compounds studied with the potential to become a medicine, most are eliminated during the initial research phases. Clinical trials follow extensive research using in vitro and animal studies. Even so, many drugs are withdrawn or fail, never becoming approved as medicines. Common reasons include side-effects, the drug proving less effective than hoped or lacking financial viability.

References

1. https://www.wikipedia.org

2. https://www.enago.com/academy/biological-databases-an-overview-and-future-perspectives/

3. https://www.intechopen.com/books/vaccines/the-impact-of-bioinformatics-on-vaccine-design- and-development