World Journal of Pharmaceutical and Life Sciences
WJPLS
www.wjpls.org
SJIF Impact Factor: 7.409

DATA INTEGRATION AND ARTIFICIAL INTELLIGENCE: TRANSFORMING NEW DRUG DEVELOPMENT

Shaikh Zainab, Mohd. Arfat*, Shaikh Awais and Syed Ayaz Ali

Department of Pharmacology, Y.B Chavan College of Pharmacy, Aurangabad, (431001) Maharashtra, India.

***Corresponding Author: Mohd. Arfat**

Department of Pharmacology, Y.B Chavan College of Pharmacy, Aurangabad, (431001) Maharashtra, India.

Article Received on 09//04/2025 **Article Revised on 30/04/2025** **Article Accepted on 21/05/2025**

ABSTRACT

The review signifies and is based on the development of new drug, which is a highly complex and resource-intensive process that has been revolutionized by the integration of artificial intelligence (AI) and data integration techniques. The article focuses on the integration of AI and data promises that is to transform the drug development process, making it more efficient, cost-effective, and tailored to patient needs. As these technologies evolve, they have the potential to deliver innovative therapies that improve patient outcomes and redefine healthcare innovation. Data integration involves harmonizing information from diverse sources, such as genetic, proteomic, chemical, and clinical datasets, to provide a comprehensive understanding of patient characteristics, drug-target interactions, and disease mechanisms. This allows researchers to make informed decisions at every stage of drug development, from target identification to clinical trials. AI, particularly machine learning (ML) and deep learning (DL), plays a pivotal role in analyzing this integrated data. These technologies identify patterns in large-scale datasets, predict drug efficacy, optimize therapeutic strategies, and enhance drug discovery. For example, deep learning models have been used to discover novel compounds like halicin, an antibiotic effective against drug-resistant bacteria. AI also accelerates clinical trials by predicting patient recruitment outcomes and optimizing trial designs. Moreover, AI-driven process optimization improves pharmaceutical manufacturing by ensuring quality control and minimizing production disruptions. Despite these advancements, challenges such as data standardization, and ethical concerns like data privacy remain significant.

KEYWORDS: Data integration, Artificial intelligence, Drug development, Drug discovery.

Highlights

- Data integration is the process of merging data from several source systems to produce cohesive sets of information for analytical and operational purpose.
- Artificial intelligence (AI) has the potential to revolutionize the drug discovery process, offering improved efficiency, accuracy, and speed.

1. INTRODUCTION

In medical science, there are a number of applications, methods, and protocols that are used in the effective and efficient development of a drug. The process of drug development is a complex and highly regulated process, which requires years of research, theoretical and practical work, and constant development of technologies. The main aim and focus of the study is to study the technological improvements and use in the development of a new drug and the technical role in the process. The introduction is divided into two parts, one that explains the Data Integration and the other is about the role of Artificial Intelligence in the development of the drug process.

Data integration and artificial intelligence have a close relationship because both fields work together to improve drug research and healthcare. It is the most basic form, healthcare data integration is the deliberate act of combining several data sources from the healthcare sector into a single, cohesive dataset. Integrating data from multiple sources is a part of this process. These sources include wearable technology, other relevant data sources, electronic health records (EHRs), which concentrate on individual patient records, integrating healthcare data, and more. Its function has changed throughout time, turning it from merely an important technology project to a key element in improving patient care, increasing operational effectiveness, and completely changing the healthcare industry. Its importance in providing efficient, superior patient care will only increase going forward. Data is the cornerstone of the healthcare industry, offering an all-encompassing view of patient health and the provision of care. The correctness and accessibility of this data are essential for every diagnosis, course of therapy, and engagement with patients.[1]

1.1 What Is New Drug Development?

Drug development encompasses all the activities involved in transforming a compound from a drug candidate (the end-product of the discovery phase) to a product approved for marketing by the appropriate regulatory authorities. Efficiency in drug development is critical for commercial success, for two main reasons:

- Development accounts for roughly two-thirds of total R&D costs. The cost per project is significantly higher throughout the development phase and drastically increases as the project progresses into the final stages of clinical development. Keeping these costs under control is a top priority for management. Failure of a compound late in development entails a significant financial loss.
- Speed of development is a significant aspect in determining sales income since time invested in development reduces the period of patent protection once the drug enters the market. As soon as the patent expires, generic competition significantly reduces sales income.[1]

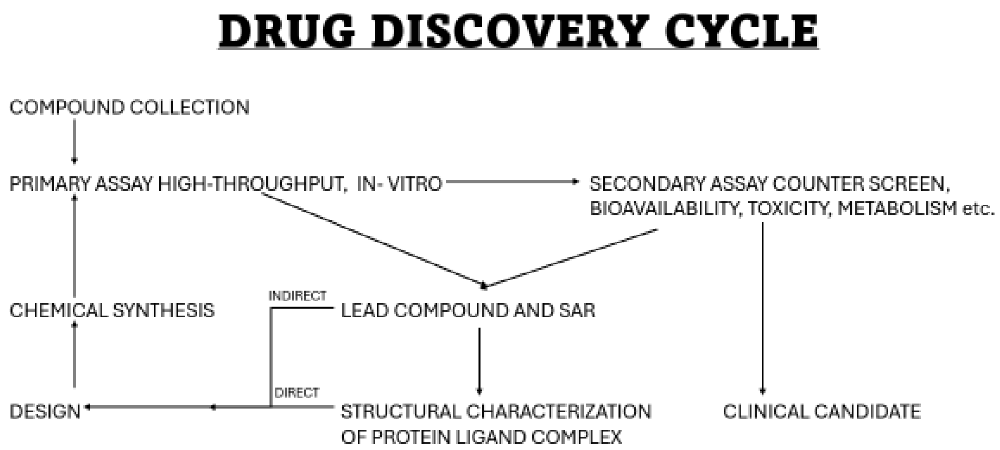


Fig. No. 01: Schematic Diagram of Drug Discovery Cycle.

1.1.1 Discovery and Development: Researchers must find or "discover" a particular molecule, usually a protein, metabolite, RNA molecule, or DNA sequence, that is essential to a disease state and that a drug can target to produce positive and therapeutic effects long before any work on drug development and manufacturing can begin. Researchers can next look for a substance or compounds that interact with the target molecule and could be used as drug candidates after making that discovery. In order to evaluate each compound's performance and viability as the ultimate, most effective therapeutic material, researchers must execute a number of experiments on the many compounds that are typically recognized as prospective candidates. They usually evaluate things like administration, side effects, absorption, and possible interactions.[2]

1.1.2 Preclinical Investigations: During the discovery phase, promising chemicals are thoroughly tested in a laboratory environment. This involves testing the drug's safety, effectiveness, and possible negative effects in vivo on animals as well as in vitro on cells.[2]

1.1.3 Development of Clinical Practices: In this stage, the medication is tested on humans through a sequence of clinical studies, usually split into three stages:

- ➢ Phase I : In a small sample of healthy volunteers, safety and dose are the main concerns.
- ➢ Phase II : Assesses the medication's efficacy in treating the intended ailment in a bigger patient population.
- ➢ Phase III : Verifies the medication's effectiveness and assesses its impact on current therapies on a significantly bigger patient base.[3]

1.1.4 FDA Exam: An extensive application for drug approval is submitted to regulatory agencies such as the FDA upon the successful completion of clinical trials. This entails a thorough examination of every piece of information gathered during the development phase.[3]

1.1.5 Following the Market: Drugs are constantly observed to track their long-term safety and effectiveness in the actual world, even after approval. This guarantees patient safety and assists in identifying any unanticipated side effects.

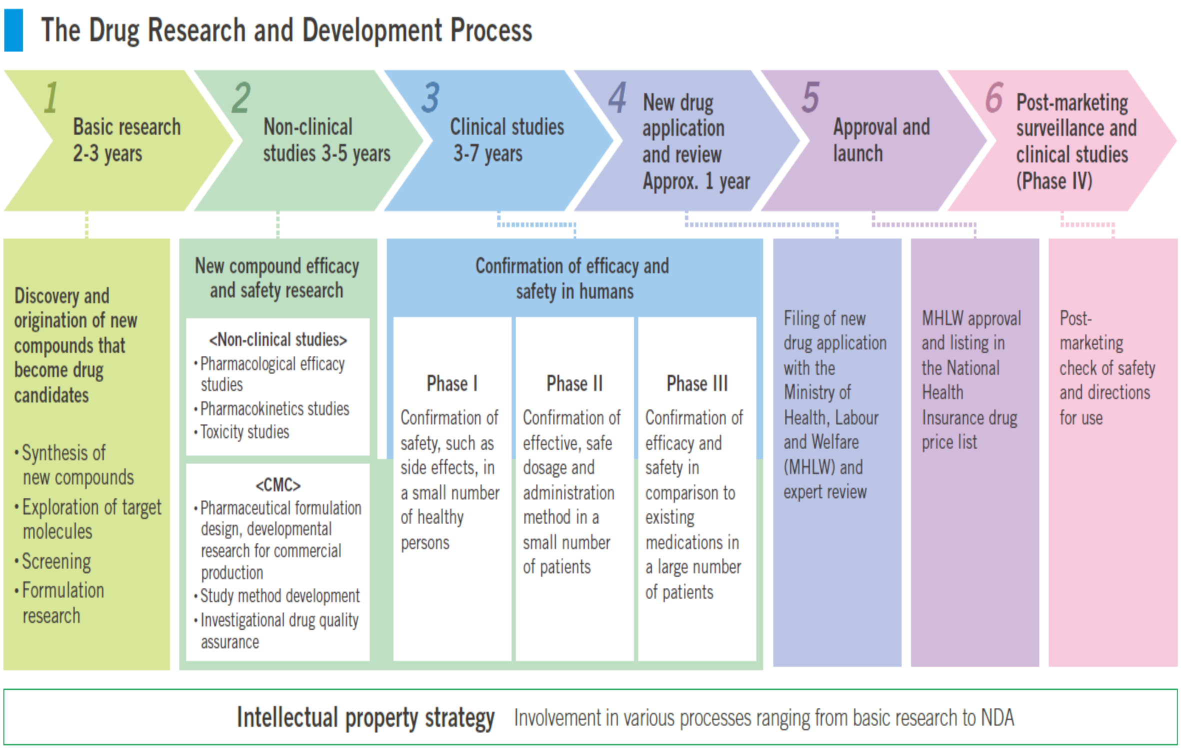


Fig. No. 02: The Drug Research and Development Process.[2]

1.2 What Is Data Integration?

The process of combining data from various sources into one cohesive location is known as data integration. Depending on the datasets, pharmaceutical data integration can take many different forms, but in general, it entails combining information from gene data, clinical trials, various healthcare systems, and the drug development process. Healthcare providers, insurance companies, pharmaceutical firms, and other associated entities are some of the many sources from which this information is frequently obtained. Enhancing the effectiveness of the healthcare system, particularly in the treatment of life-threatening illnesses, the integrated data offers a platform for predictive analytics to estimate patient outcomes.[3]

Increasing the action ability of your data is the fundamental idea behind this. Technology is continually evolving and becoming more sophisticated. The new emerging technologies like cloud storage and other big data technologies have overcome the data manipulation techniques that were employed ten to twenty years ago.[4] Although data integration is not a new process, its importance has grown due to the volume of data being generated today. For businesses handling massive volumes of data, it has changed from being a standard procedure to a vital task. However, despite how straightforward it may seem, there are numerous obstacles to overcome in the process of successfully managing and utilizing data.[5]

1.2.1 History of Data Integration

The 1980s saw the invention of the concept of data integration. For the Integrated Public Use of Microdata Series (IPUMS), the University of Minnesota developed the first data integration system driven by structured metadata in 1991.[6] However, this data integration method came with a number of complexities and infrastructure-related problems. Data integration solutions therefore required improvement, just like any other technology. Self-service and automation are used by today's data integration technologies to quickly, securely, and less laboriously aggregate data and create connections.[7] Enterprises started to generate more data and demand more from it over time. As a result, data integration solutions have improved in terms of both usability and granularity. Administrators have the option of migrating data in batches or transactions, transforming and filtering data in real time, and even creating scenarios where failures are automatically handled by the data integration platform. Extract, transform, and load (ETL) tools were first created with the simple idea of taking data from one source, transforming it into a format that another application can understand, and then loading it into that application. This approach gave rise to data integration. Taking the raw financial data from several receipts, putting it into a spreadsheet, and then loading it into an accounting programme would be a simple example.[8]

1.2.2 Modernization In Data Integration

Pharmaceutical companies need data integration from a business standpoint. Cost- effective medications can be made with the help of data gathered from patient records.

Clinical studies can be optimized with the use of the current information available. They can also readily give new formulas and increase industrial efficiencies.

Patients can receive personalized marketing communications made with the data. The pharmaceutical corporations' benefits from all of these as a competitive advantage. The data that is provided is in an unorganized format and has a vast field to cover. Such a amount of data needs to be organized, analyzed, complied and interpreted appropriately to benefit the whole organization in a specified and actionable manner.

Fig. No: 03: Data Modernization Benefits.[15]

Development and modernization in data integration needs to maintain data governance standards, quality of data, and confidentiality of data, that is accessed by the healthcare industry or pharma sector. Once the integration is done correctly, different data sets can be used by pharmaceutical industries multiple times, which provides opportunities for setting up and accomplishing new goals. Data integration is empowered by artificial intelligence (AI) and machine learning to achieve and fulfill various algorithms.

1.3 Artificial Intelligence

The ability of computers and other devices to mimic human learning, comprehension, problem-solving, decision-making, creativity, and autonomy is known as artificial intelligence (AI). Artificial intelligence (AI) is finding increased application in a variety of sectors, including banking, healthcare, and transportation. Large-scale dataset analysis is the foundation of artificial intelligence, necessitating a steam of high-quality data. Nonetheless, there are difficulties with utilizing data for AI. The difficulties of using data for AI are thoroughly reviewed and critically examined in this study, along with issues related to data amount, quality, privacy, security, bias, and fairness, as well as technological know-how and skills. This essay looks closely at these issues and makes suggestions for how businesses and organizations should deal with them. Organizations can use AI to make better judgments and obtain a competitive edge in the digital era by comprehending and tackling these issues. The ability of machines to replicate human intelligence and carry out tasks that normally require human intelligence, such as learning, problem-solving, decision-making, and interpreting natural language, is known as artificial intelligence (AI).[9]

1.3.1 History of Artificial Intelligence

The concept of "a machine that thinks" (AI) has been around since ancient times, but with the introduction of electronic computers significant occasions and turning points in the development of AI have included the following:

- **1956:** The phrase "artificial intelligence" was originally used by John McCarthy in during the inaugural AI conference held at Dartmouth College. The first AI computer program to operate was created later that year by Allen Newell, J.C. Shaw, and Herbert Simon. It was called Logic Theorist.
- **1967:** Mark 1 Perceptron, the first computer built on a neural network that "learned" through trial and error, was built by Frank Rosenblatt. A year later, Marvin Minsky and Seymour Papert publish Perceptrons, a book that becomes both a seminal work on neural networks and, for a while at least, a critique of subsequent neural network research projects.
- **1995:** One of the most important textbooks in artificial intelligence is Artificial Intelligence: A Modern Approach, written by Stuart Russell and Peter Norvig. In order to distinguish computer systems according to their reasoning and thinking as opposed to their action, they explore four possible objectives or definitions of artificial intelligence.
- **2004:** A paper titled "What Is Artificial Intelligence?" Is written by John McCarthy. It offers

a definition of AI that is frequently used. At this point, the big data and cloud computing era has begun, allowing businesses to handle ever-increasing data estates that will eventually be utilized to train artificial intelligence models.

- **2022:** The performance of AI and its capacity to provide enterprise value is drastically altered by the proliferation of large language models, or LLMs, like OpenAI's ChatGPT. Large volumes of data can be used to pretrain deep-learning models using these new generative AI techniques.
- **2024:** A continuing AI renaissance is indicated by the most recent trends in AI. Richer and more reliable experiences are being offered by multimodal models that can accept various kinds of data as input. These models combine NLP voice recognition with computer vision picture recognition. In a time when enormous models with many parameters are becoming less effective, smaller models are also making progress.[9]

1.3.2 Modernization in Artificial Intelligence

The domain responsible for the well-being and survival of humans is healthcare. In order to provide necessary treatment, those working in the healthcare industry must be vigilant. The healthcare system has advanced significantly in recent years as a result of its integration with artificial intelligence (AI).[9]

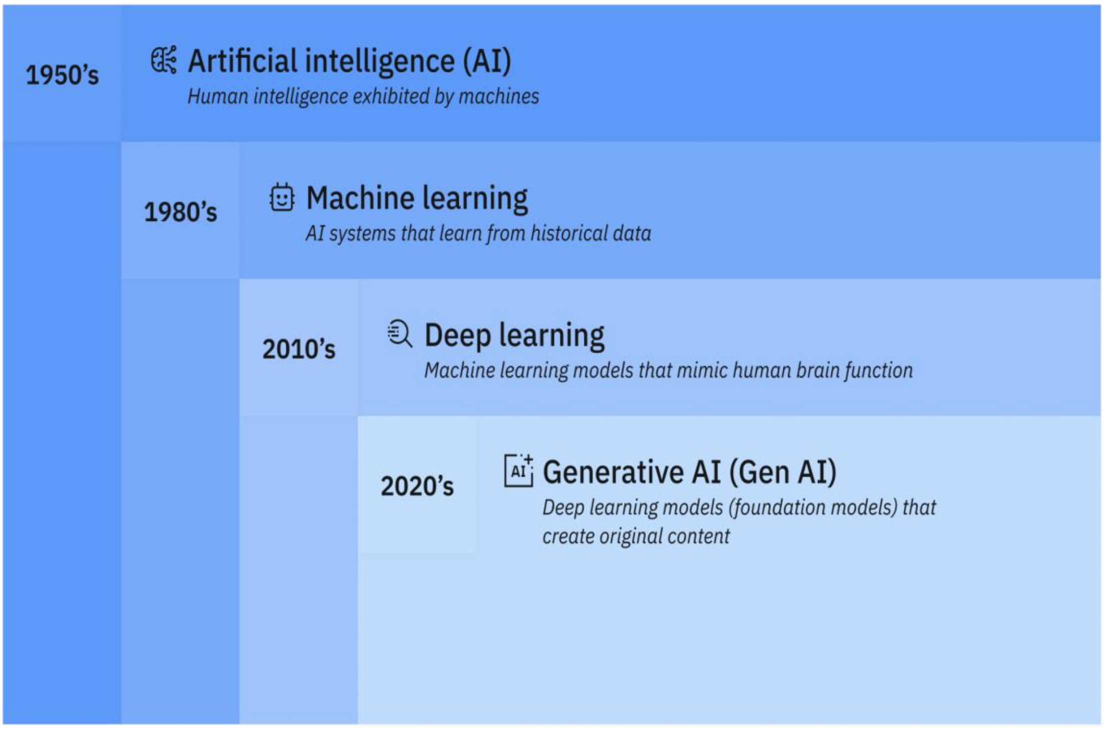


Fig. No: 04.

Progression of technological advancement over time in AI[19]

- Artificial Intelligence (AI): The ability of machines to imitate human behavior is the broad definition of artificial intelligence. It includes a wide range of methods and strategies meant to give machines the ability to see, think, learn, and decide. AI can involve machine learning algorithms, statistics, or rules. Artificial intelligence gave rise to machine learning, deep learning, and generative AI.
- Machine Learning (ML): Machine learning (ML) refers to the process by which machines learn from data without explicit programming. Based on previously collected data, machine learning algorithms use statistical methods to automatically identify patterns and generate predictions or judgments. The term was created to highlight the significance of data-driven learning and the capacity of computers to enhance their performance through exposure to pertinent data, even if machine learning is a subset of artificial intelligence.
- Deep learning (DL): As a distinct subfield of artificial intelligence (AI), deep learning is crucial because of its special powers and developments. Deep learning is a type of machine learning that uses data inspired by people to train the computer to learn. DL uses multi-layered deep neural networks to learn data representations in a hierarchical fashion. Manual feature engineering is eliminated, and pertinent features are automatically extracted. DL is more capable of managing intricate jobs and extensive datasets.
- Generative AI (Gen AI): The goal of generative AI, a subfield of deep learning and artificial intelligence, is to build models that can produce new content that closely resembles preexisting data. The information produced by these models is intended to be identical to that which may be produced by people. Popular

generative AI models that employ deep neural networks to produce lifelike text, images, and even music are called Generative Adversarial Networks (GANs).[10]

Types of Data Integration

Drug discovery and development

A key component of the drug development process is data integration. To properly develop and commercialize a novel treatment, this procedure depends on precisely and successfully integrating a wide range of data, from preclinical investigations to clinical trials. A comprehensive perspective of the process is also made possible by integrated data, which enhances effective decision-making and cross-functional communication. Data integration has broad applications in health care and drug discovery. Data integration makes it possible to identify possible drug candidates more quickly, which speeds up the drug development process in the drug discovery phase. It improves the safety profile of medications under development by strengthening the prediction capacity with respect to drug toxicity. Personalized therapy for patients can result from the integration of genomic data, clinical trial data, and medical information. In addition to improving therapy efficacy, these "patient-centric" approaches lower expenses and the possibility of side effects. Data integration thus contributes to the development of a better healthcare system.[11]

Drug research and development processes are facilitated by the integration of pharmacological data from several sources, including chemical databases, biological experiments, and clinical trials.[12]

Multi- omics integration

Numerous biomedical research domains, including synthetic biology, personalised medicine, drug development, and diagnostics, have used multi-omics. Multi-omics can find new biomarkers and therapeutic targets, forecast and optimise individualised treatments, and provide fresh insights into the molecular basis of diseases and medication responses by merging several forms of omics data. The field of pharmaceutical sciences could undergo a revolution thanks to multi-omics, which could also make it possible to produce novel and efficient treatments, and Multi-omics data integration workflow shown in figure 4.[13]

Drug – drug interaction

When two medications interact with the same gene product, it might lead to drug-drug interactions, or DDIs. The scientific literature has the majority of the knowledge currently accessible regarding the relationships between genes and drugs, but it is spread throughout a great number of papers, with thousands more being published every month.[14]

Data standards and integration methods

Data representation, exchange, and processing across various systems and applications are outlined in data integration standards, which are collections of regulations and best practices. These standards are intended to guarantee data quality, compatibility, and interoperability among diverse data sources and destinations.[15]

Network pharmacology

Network pharmacology (NP) is a recently developed field that aims to comprehend pharmacological effects and interactions with various targets.[16] It makes systematic use of computer power to catalogue a drug molecule's molecular interactions within a living cell. NP emerged as a key instrument in deciphering the intricate linkages that underlie the relationship between the botanical formula and the body as a whole.[17] By permitting an objective examination of possible target spaces, it also seeks to identify novel therapeutic leads and targets as well as to repurpose already-existing pharmacological molecules for various therapeutic situations.

Pharmacovigilance

The research and practices around the identification, evaluation, comprehension, and avoidance of side effects or any other issue pertaining to medications or vaccinations are known as pharmacovigilance.

It seeks to accomplish the following goals by educating the public on how physicians, chemists, patients, and other healthcare professionals may support pharmacovigilance:

- Urge individuals to be aware of and report any negative consequences.
- Teach people how they can all contribute to reporting medication adverse effects.
- Raise awareness of national reporting systems and pharmacovigilance.[18]

2.1 Collaborative research and data sharing

Research, like business, is evolving into a cooperative endeavour that unites several industries, organisations, and, more and more, nations. This is never truer than in the natural sciences, where the phenomena being studied and the questions being posed transcend the confines of a single subject, organisation, nation, or continent. The need to gather, preserve, and exchange vast and varied structured data resources that no single research team or institution has the means or ability to gather, make accessible, and maintain is both a justification for and the goal of collaboration in the sciences.[19]

Types of Artificial Intelligence

Artificial Intelligence in healthcare industry is a technique that is developed to perform variety of roles that mimics human cognitive functions and makes the beneficial use of all kinds of structured and unstructured healthcare data. There are a number of AI used for this purpose. Different types of AI techniques that are used are enlisted below.

National Language Processing (NLP)

Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that uses algorithms and methods like machine learning, deep learning, and text analytics to interpret and analyze natural language content from documents, audio recordings, images, and other sources. NLP allows machines to understand and communicate in natural language just like humans do. The adoption of natural language processing (NLP) in healthcare is growing due to its ability to analyze and interpret vast patient datasets. By leveraging advanced algorithms and machine learning, NLP can extract valuable insights from unstructured clinical notes, transforming previously buried data into actionable information. This technology enhances understanding, improves methods, and benefits patient outcomes. While not a universal solution, NLP aids healthcare professionals in making faster, more accurate decisions. Its ability to uncover hidden patterns in medical data enables improved predictive analytics and insights into diseases, leading to better treatments and diagnostic methods. The accuracy of NLP improves with increased clinical documentation and usage, as it continuously learns and can be customized to meet specific needs. Many vendors even offer systems that can be tailored to particular medical groups, optimizing performance and outcomes.[20]

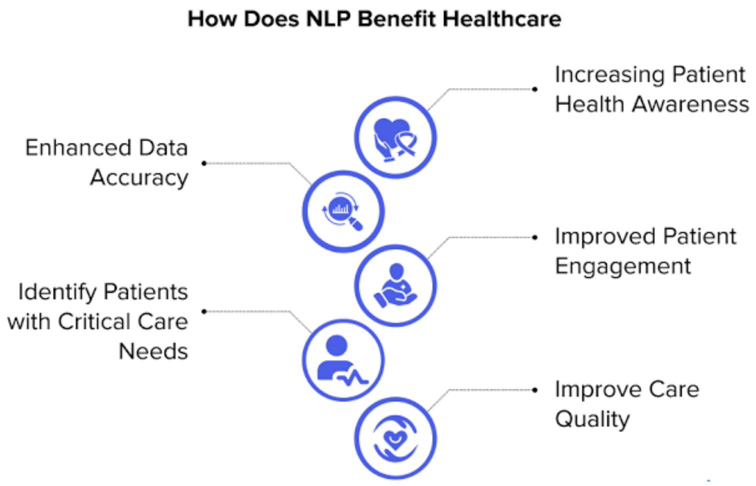


Fig. No: 05.

Different ways in which NLP helps in Benefitting Healthcare[31]

3.2 Machine Learning (ML)

Technology advancements have historically been embraced early by the healthcare industry. These days, the creation of novel medical techniques, the management of patient data, and the management of chronic illnesses are all made possible by machine learning, a branch of artificial intelligence. Numerous healthcare use cases use machine learning, and its potential is largely based on its capacity to manage complicated data. Healthcare Use of Machine Learning includes the following:

- Identifying and managing illness
- delivering medical diagnostics and
- imaging Finding and creating new medications
- Medical record organization

AI and machine learning have also affected pharmaceutical companies' efforts to find and develop new drugs. therapy manufacturers expect machine learning (ML) can anticipate how patients will react to different medications and determine which individuals have the best probability of benefiting from the therapy. The technology has already aided clinical studies in the central nervous system.

Table No: 01: Applied applications of EHR: Electronic health records, SVM: Support Vector Machine, LSTM: Long Short Term Memory Neural Network, CNN: Convolutional Neural Network, with list of references.[29]

Healthcare Area	Type of Machine Learning Model	Description	Applied or Experiment
EHRs	SVM, DT	Using EHRs for predicting diagnoses	Applied
-	RNN	Predicting post-stroke pneumonia using deep neural network approaches	Experiment
-	LSTM, CNN	Deep EHR: Chronic Disease Prediction Using Medical Notes	Experiment
-	ML	**SRML-Mortality Predictor:** A hybrid machine learning framework to predict mortality in paralytic ileus patients using Electronic Health Records (EHRs)	Experiment

3.3 Medical Imaging

Diagnostic imaging is one of the most potential clinical uses of AI, and increasing focus is being placed on developing and optimizing its functionality to enable the detection and measurement of a broad range of clinical diseases. Studies using computer-aided diagnostics have demonstrated exceptional sensitivity, specificity, and accuracy in identifying minor radiographic abnormalities, which could lead to improvements in public health. A lot of research is being done on the application of artificial intelligence (AI) in diagnostic medical imaging. AI promises to improve tissue-based detection and characterization and has demonstrated remarkable sensitivity and accuracy in identifying imaging abnormalities. However, a significant disadvantage of increased sensitivity is the ability to notice minute changes of unclear significance. An investigation of screening mammography, for instance, revealed that while artificial neural networks consistently had a better sensitivity for abnormal results, particularly for small lesions, they are no more reliable than radiologists at detecting cancer.

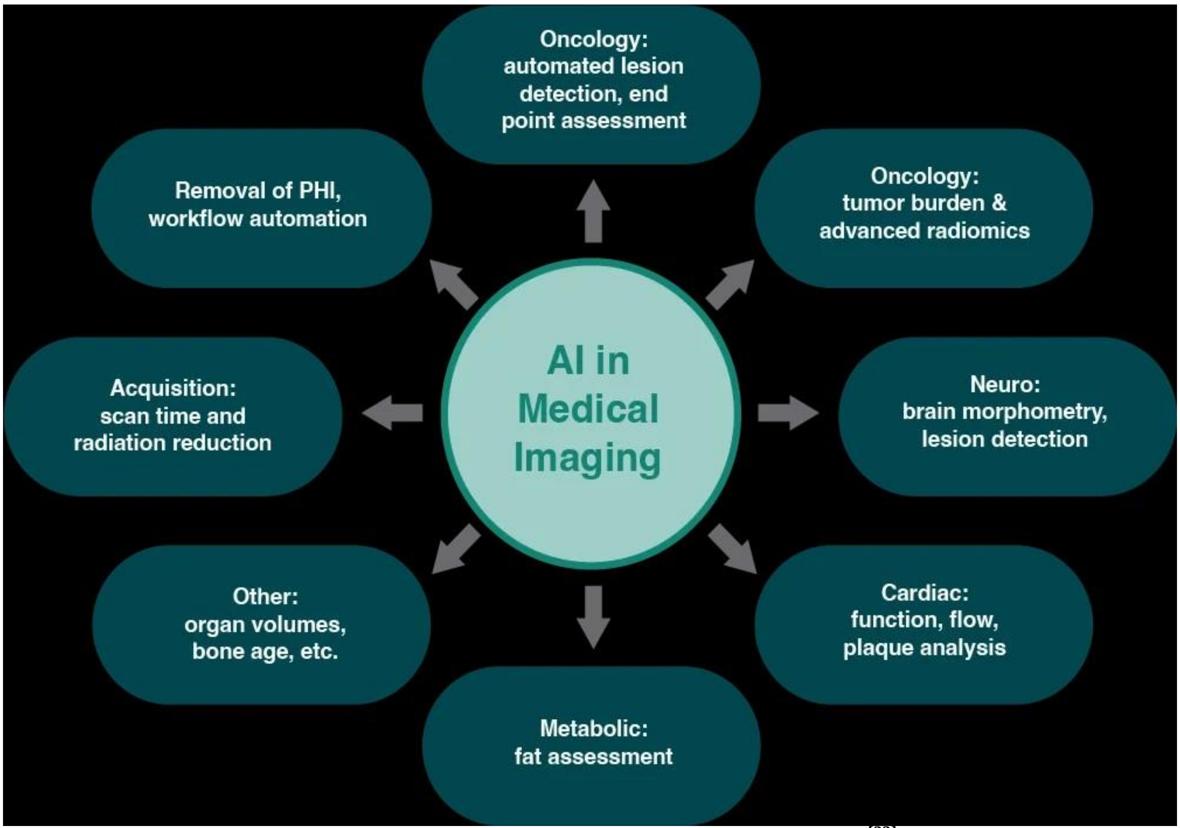


Fig. No: 06: Applications of AI in Medical Imaging.[33]

The medical community must foresee the possible unknowns of this technology at the outset of an AI-assisted diagnostic imaging revolution in order to assure its safe and efficient integration into clinical practice. Establishing AI's place in clinical medicine requires a careful evaluation of its possible risks in light of its special powers; striking a balance between improved detection and over diagnosis won't be simple. Consistent utilization of well-defined cohorts and out-of-sample external validation to improve the calibre and interpretability of AI research are essential components of this evaluation.

3.4 Genetic Engineering & Genomics

AI algorithms are inspired by human intelligence, but their applications in clinical genomics typically focus on tasks that are difficult to complete with human intelligence and prone to errors when handled with conventional statistical methods. To address the different steps involved in clinical genomic analysis, such as variant calling, genome annotation, variant classification, and phenotype-to-genotype correspondence, many of the techniques mentioned above have been modified. In the future, they might also be used for genotype-to-phenotype predictions. Here, we outline the main categories of clinical genomics issues that AI has tackled.[21]

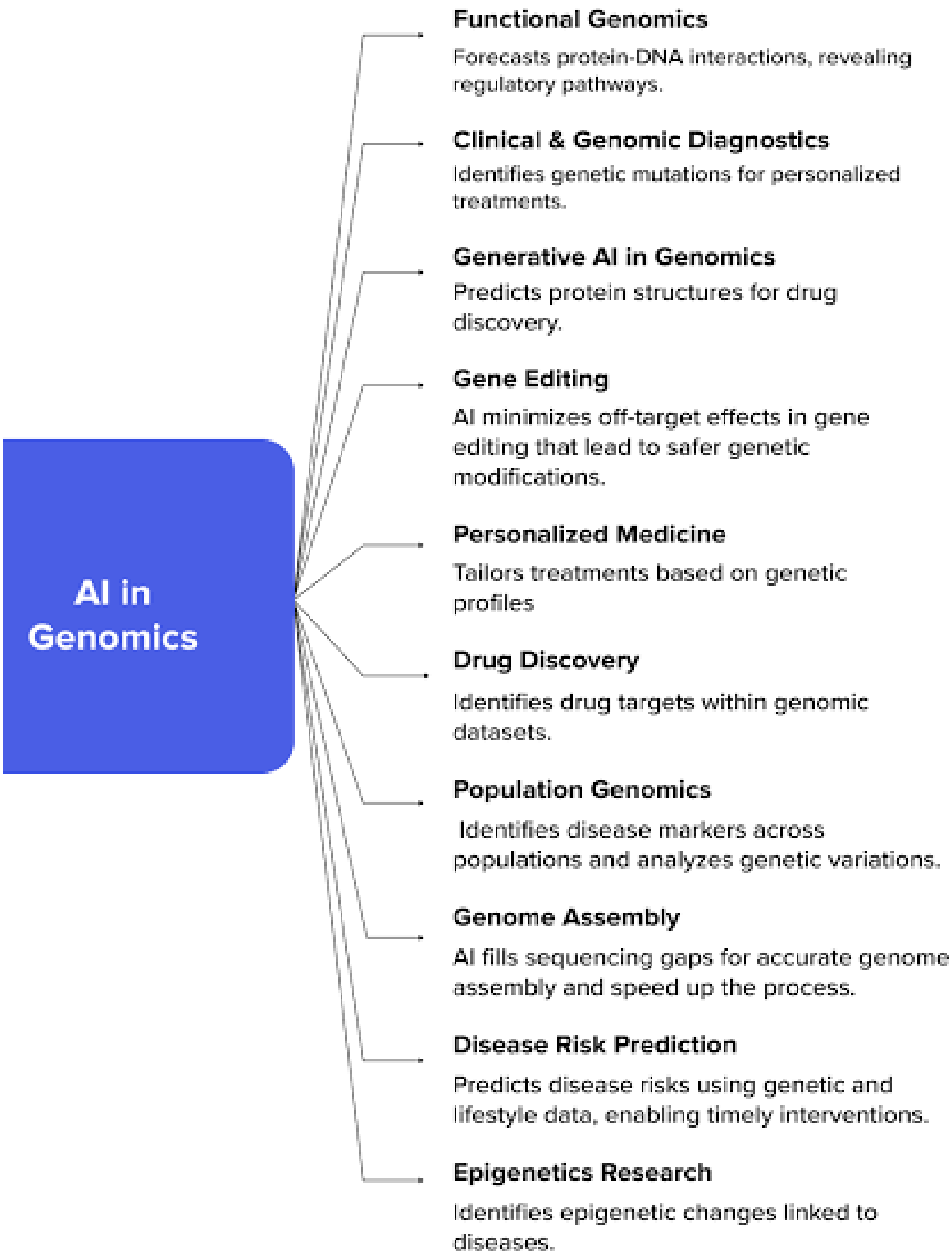


Fig. No: 07: Applications of AI in Genomics.[30]

3.5 Drug Development

Existing data from multiple sources, including high-throughput compound and fragment screening, computer modeling, and information found in the literature, serve as the foundation for the feedback-driven drug development process. Induction and deduction are switched between in this process. Hit and lead compounds are ultimately optimized as a result of this inductive–deductive cycle. Drug development is more efficient when some cycle steps are automated, which lowers errors and randomness. Finding effective new medications is a challenging task and, for the most part, the most challenging aspect of drug development. The reason for this is the enormous extent of what is referred to as chemical space, which is thought to be around 1060 molecules. AI-infused technologies have evolved into multipurpose instruments that may be used widely in many phases of drug research, including identifying and validating pharmacological targets, creating novel medications, repurposing existing ones, and enhancing R&D.[30]

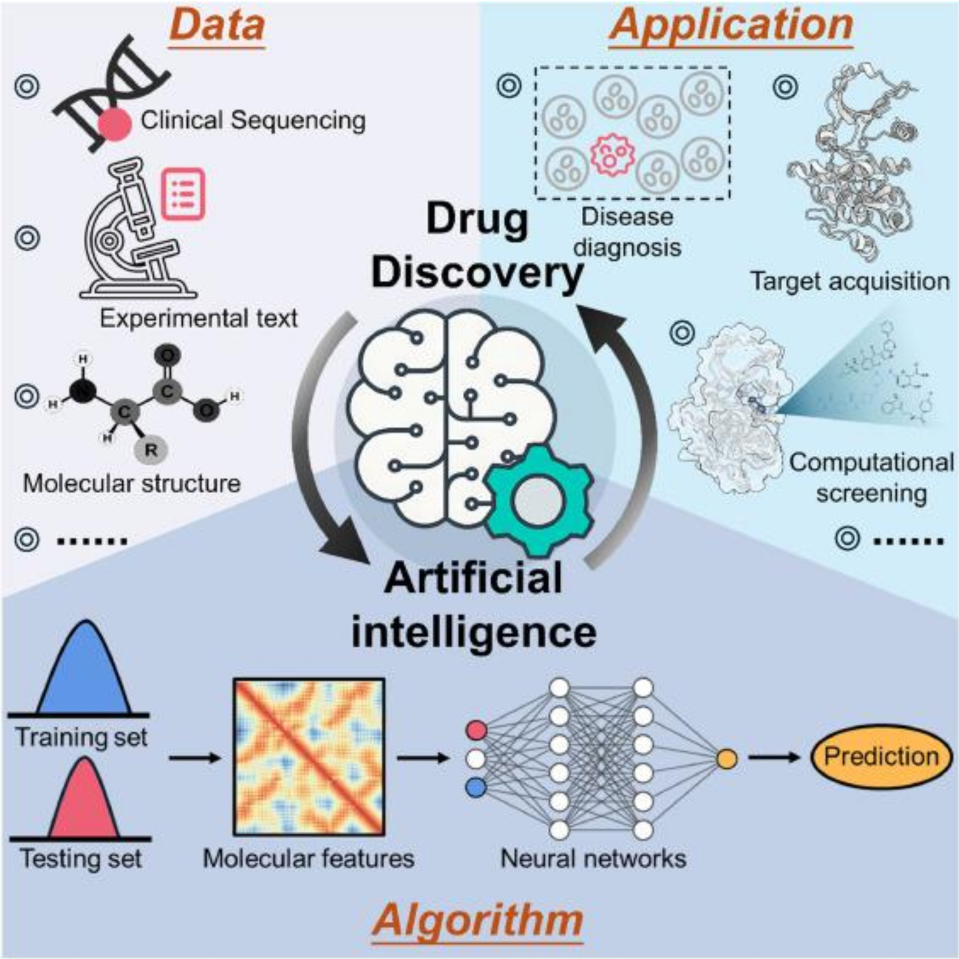


Fig. No. 08: Schematic representation of the integration of artificial intelligence (AI) in the drug discovery process.[31]

2. Data Integration in Drug Development

In a healthcare facility, data integration enables the rapid sharing of information throughout departments. Even non-technical members can access and share data using data integration solutions because to their extensive and user-friendly interface. In the highly fragmented healthcare sector, this makes data interoperability easier. A unified view of patient data is another benefit of data integration. Better decision-making is made possible for healthcare workers by the ability to access all information integrated into a single dashboard. Additionally, it conforms to privacy regulations that ensure the total privacy of patient information and shield it from unauthorized access and alteration.[31]

4.1Sources of Data In Drug Development

4.1.1 Preclinical Data

Preclinical drug discovery entails thorough safety assessments of the possible therapeutic intervention, frequently utilizing animal and cell models of disease. Going through a large number of histopathological samples is a regular part of these investigations. This makes the drug development process labor-intensive and time-consuming for contract research and pharmaceutical companies. Analytical activities that must be repeatedly performed can be quite challenging in preclinical research. For instance, needing to accurately and precisely identify extremely tiny changes is a challenge. In addition to being laborious and time-consuming, manually quantifying minor changes or counting individual cells is frequently expensive. The use of AI in preclinical research is exemplified by Lundbeck Pharmaceuticals' neuroscientific histopathological examination of alpha-synuclein, a neuronal protein associated with Parkinson's disease. Using traditional approaches to quantify alpha-synuclein-positive neurons or the entire diseased area is notoriously difficult and sluggish; however, the Aiforia Create AI solution made this possible with great speed and precision. The AI-based quantification was shown to be just as accurate as the manual method in assessing α-synuclein pathology, as evidenced by the substantial correlation (p-value <0.0001) between the manual and AI cell counts.[31]

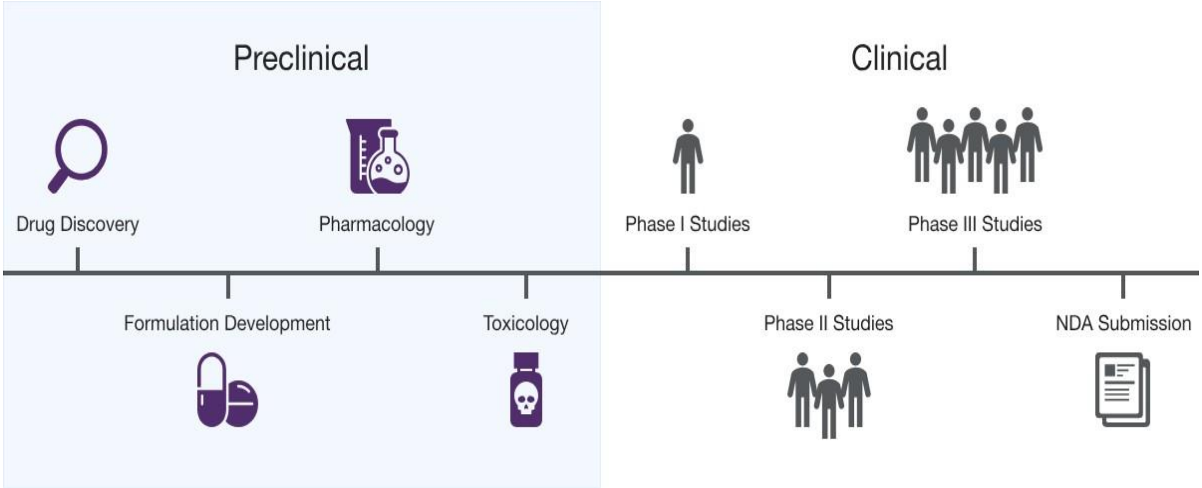


Fig. No: 09: Methods Involved from Preclinical to Clinical Process.[30]

4.1.2 Clinical Data

The majority of health and medical research relies heavily on clinical data, which can be gathered as part of a formal clinical trial program or as part of ongoing patient care. There are six main categories of clinical data:

- **Electronic health records:** The purest form of electronic clinical data that is collected at the point of care in a clinic, hospital, medical facility, or practice. The electronic medical record (EMR), as it is commonly known, is typically inaccessible to outside researchers. The information gathered includes hospitalization, patient insurance, laboratory testing, prescription medication, diagnosis, treatment, administrative and demographic data, and physiologic monitoring data.
- **Administrative data:** These are mostly hospital discharge statistics that are submitted to a government entity such as AHRQ and are frequently linked to electronic health records. Project on Healthcare Cost and Utilization (H-CUP) Based on information from the Healthcare Cost and Utilization Project (HCUP), HCUPnet is a free online query tool. Health statistics, hospital inpatient data, and ED utilization data are all accessible through it. Several datasets and sample studies are included in the project and are indicated by the information icon. It is possible to purchase datasets.
- **Claims data:** The billable interactions (insurance claims) between insured patients and the healthcare delivery system are described by claims data. Inpatient, outpatient, pharmacy, and enrollment are the four broad categories into which claims data can be divided. The government (such as Medicare) and/or private health companies (such as United HealthCare) are the sources of claims data. Medicare Basic Stand Alone (BSA) Claims Files for Public Use (PUFs) For Medicare claims, these are the Basic Stand Alone (BSA) Public Use Files (PUF). Each record in this claim-level file represents a claim made by a 5% sample of Medicare beneficiaries. Prescription medications, DME, SNF, hospice, inpatient and outpatient care, and more are all covered by claims. Each PUF includes a few demographic and claim-related characteristics.
- **Patient / Disease registries:** Clinical information systems known as disease registries monitor specific critical data for chronic disorders like Alzheimer's disease, cancer, diabetes, heart disease, and asthma. Registries frequently offer vital data for patient condition management.
- **Health surveys:** Prevalence estimates are often obtained through nationwide surveys of the most prevalent chronic illnesses in order to provide an accurate assessment of population health. One of the rare forms of data gathered especially for research is national surveys, which makes it more readily available.
- **Clinical trials data:** There may be national or discipline-specific organizations that provide clinical research data. Access is probably limited, but it is accessible through the appropriate channels. Individual contracts with private businesses may also provide access to proprietary research data.

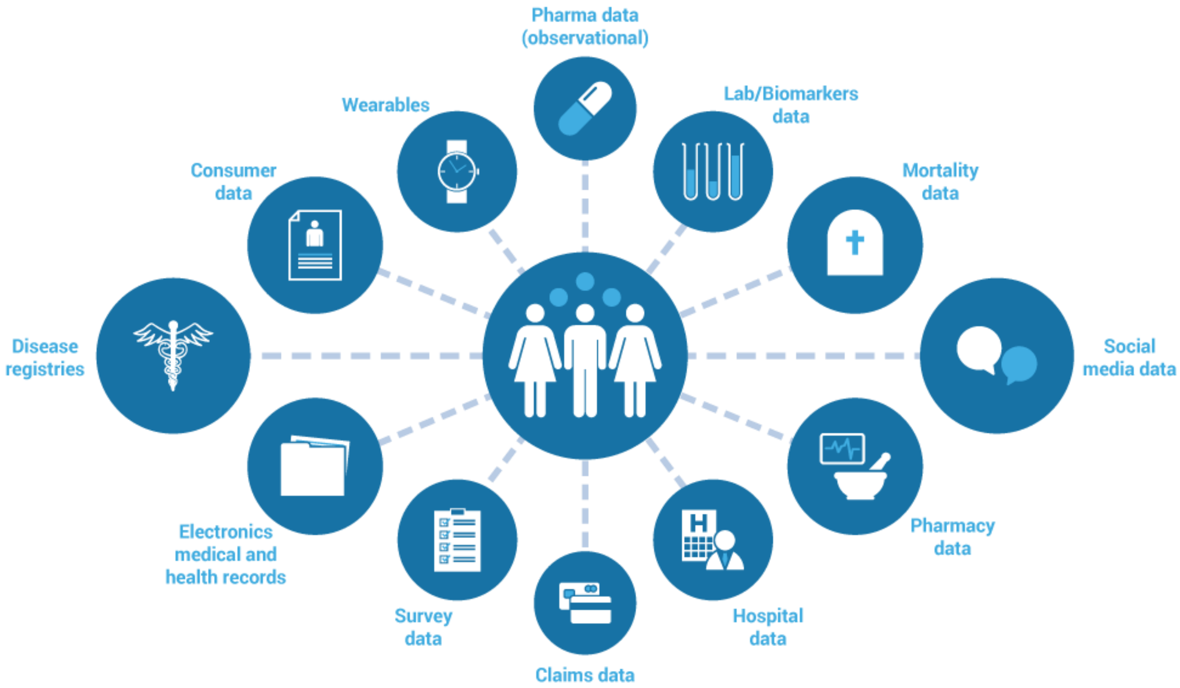


Fig. No: 10: Data types related to the health status of a natural person (e.g., diagnosis, treatment, laboratory test results).[32]

4.1.3 Real world data

In Drug Development Process and for Pharma companies, Real World Data (RWD) is a valuable resource that can be leveraged in various ways, some of which are as follows:

- ➢ Identification of new targets for treatment: Understanding real-world patient outcomes allows researchers to identify which illnesses or biomarkers are most promising for intervention, and analyzing RWD can highlight patterns and correlations that indicate to possible new treatment targets.
- ➢ Evaluation of effectiveness and safety of Drugs: RWD makes it possible to continuously track a drug's effectiveness and safety profile across a range of demographics. This can result in more precise evaluations of a medication's advantages and disadvantages outside of the regulated setting of clinical trials.
- ➢ Support regulatory decision making: Regulatory bodies are taking RWD into account more and more when deciding whether to approve or reject drugs. RWD can assist in proving the worth of a medication by offering solid data from actual situations, which can result in more well-informed and prompt regulatory decisions.

Integrating real-world data into the drug development process enables pharmaceutical companies to make more informed decisions, improve patient outcomes, and expedite the delivery of new therapies to the market.[22]

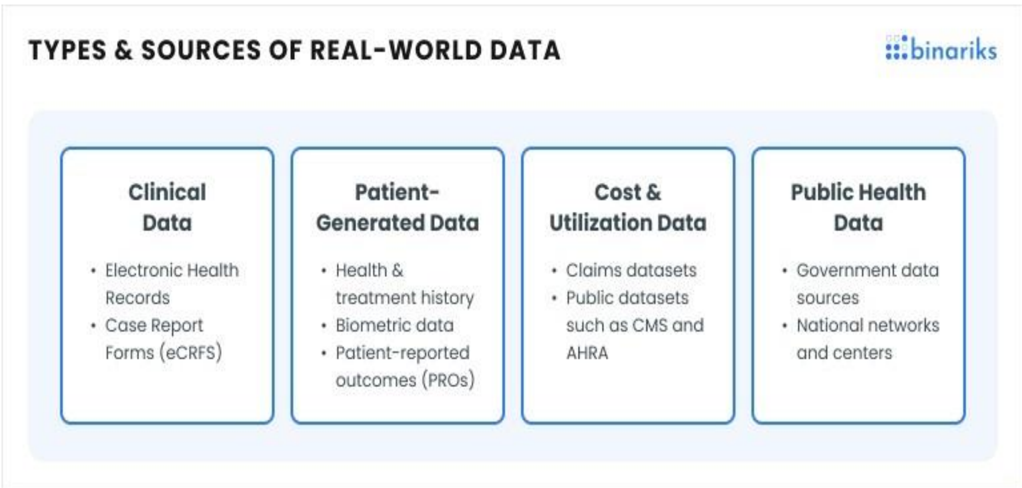


Fig. No. 11: Types and sources of Real World Data.[31]

Advantages of Real World Data in Drug Development Process

1. Enhanced patients insights
2. Faster regulatory approvals
3. Cost efficiency
4. Improved trial designs
5. Better market access strategies
6. Electronic records of data

3. Role of Artificial Intelligence (AI) in Drug Development

A significant shift in the pharmaceutical industry has resulted from the acceleration of the pharmaceutical sector's growth through the integration of artificial intelligence into medication discovery and development. We go over areas of integration, tools and methods used to enforce AI, as well as current issues and solutions. Drug discovery and development, clinical trials, disease diagnosis, various phases of pharmaceutical manufacturing, data analysis, and supply management are just a few of the ways artificial intelligence is applied in the pharmaceutical sector. Clinical trials and drug discovery account for the majority of the time and expense. Artificial intelligence can minimize human mistake in data processing, documentation, data integrity issues, and data selection throughout the journey. Prescriptive, diagnostic, predictive, and descriptive modes are all available. Several areas of pharmaceutical and medical science have already seen the application of artificial intelligence by large pharmaceutical corporations such as Pfizer, Roche, Novartis, and Johnson & Johnson.[23] AI has the potential to completely transform the way medications are found and produced since it allows us to examine enormous volumes of data, spot trends, and produce insights. Artificial Intelligence (AI) presents potential to improve drug-discovery pipeline efficiency, precision, and success rates by complementing and expediting conventional procedures.[45]

5.1 Drug Discovery and Design

Artificial Intelligence (AI) has become a pivotal force in modern drug discovery and design, offering innovative approaches to identify and optimize potential therapeutic compounds. Two significant applications of AI in this domain are de novo drug design and virtual screening.

5.1.1. AI Models for De Novo Drug Design and Molecular Property Prediction

In order to find compounds with desirable biological activity, de novo drug creation entails building new molecular structures from scratch. While traditional approaches frequently rely on[24] pre-existing chemical libraries, artificial intelligence (AI) allows the exploration of enormous chemical regions outside these constraints.

- **Generative Models:** To create novel molecules, AI uses generative models like Variational Auto-encoders (VAEs) and Generative Adversarial Networks (GANs). These models create new structures with particular characteristics by learning patterns from known substances. One study, for example, focused on the application of deep reinforcement learning to de novo drug design, in which artificial intelligence (AI) systems iteratively enhance chemical production to satisfy specific requirements.
- **Molecular Property Prediction:** AI models can analyze large datasets and predict key chemical attributes including stability, bioactivity, and solubility. This prediction capacity focuses on molecules with the best properties, which speeds up the process of finding promising therapeutic candidates. During the design process, these models evaluate and optimize molecular features, according to a survey on generative AI for de novo drug design.[33]

5.1.2. Virtual Screening of Millions of Compounds Using AI

A computational method called virtual screening is used to assess huge libraries of chemicals in order to determine which ones are most likely to bind to a therapeutic target, like a disease-associated protein. AI improves this procedure by making it faster and more accurate.

- **Structure-Based Virtual Screening:** AI systems predict binding affinities by examining the three-dimensional structures of putative ligands and target proteins. Promising candidates for additional research were identified by an AI-accelerated virtual screening tool that showed the potential to quickly screen libraries of billions of compounds.[25]
- **Machine Learning Models:** Machine learning models can prioritize candidates for experimental validation by predicting the likelihood of interactions between chemicals and targets based on training on known data. As opposed to conventional high-throughput screening techniques, this method uses less time and resources. The application of AI to drug screening and design highlights the advantages of virtual screening due to AI's capacity to manage large chemical spaces.[34]

5.2. Preclinical Studies

Preclinical research on animal and cell models, as well as human clinical trials, are all part of the process of developing and discovering new drugs. Afterward, the drug must receive regulatory approval before being put on the market. The process of finding new drugs entails identifying screening hits, optimising medicinal chemistry, and increasing the affinity, selectivity, efficacy, potency, metabolic stability, oral bioavailability, and reduction of side effects. Before conducting clinical trials, a molecule that satisfies all of these criteria will be found, and the drug development process will start.[34]

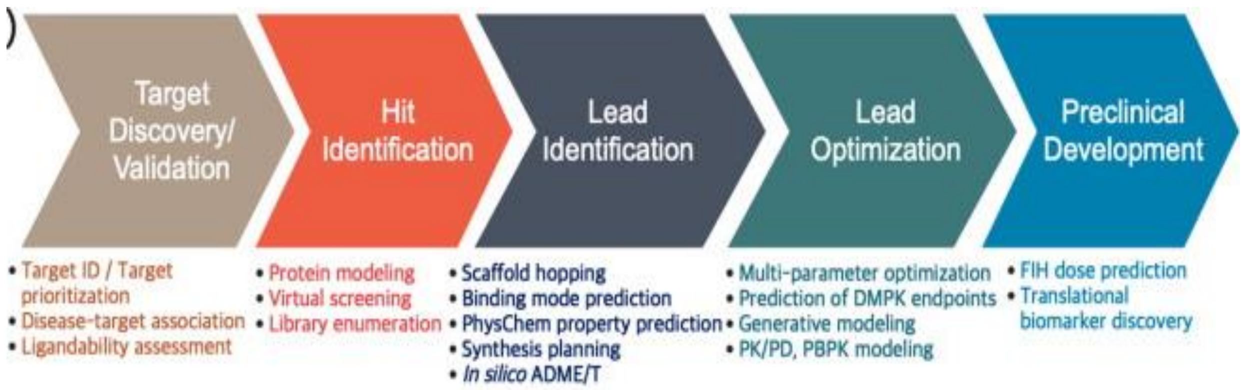


Fig. No. 12: A schematic illustration of the preclinical drug development procedure that shows how AI can be used throughout the drug discovery spectrum.[34]

5.3 Clinical Trials

Clinical trials (CTs) continue to be the cornerstone of secure and efficient medication development. It is critical that businesses and regulators use customized Artificial Intelligence (AI) solutions that facilitate rapid and efficient clinical research, given the changing data-driven and personalized treatment approach in healthcare. The opportunities, difficulties, and possible ramifications of AI in CTs were noted in this research.[26]

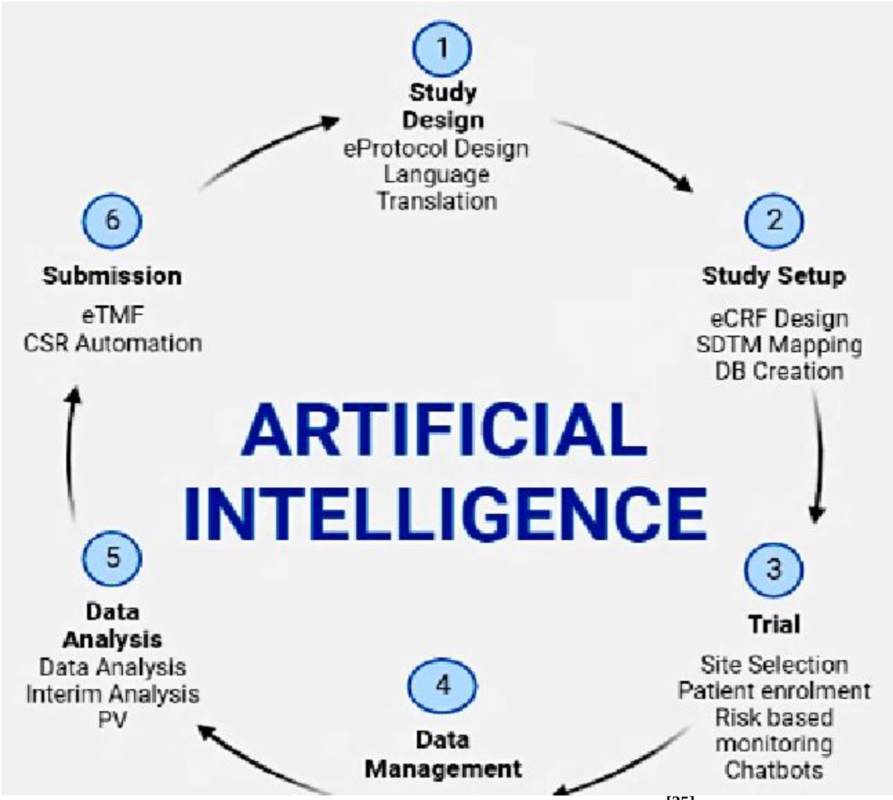


Fig. No. 13: AI in Clinical Study Process.[35]

AI can assist with site selection, for instance. The AI algorithm take could be used to optimize trial operational conduct by assisting in the identification of places with the highest likelihood of successful trial recruitment. Its application can also enable more focused engagement efforts, improve participant recruitment tactics, and improve venue selection. Artificial Intelligence (AI) has already been investigated and applied in a clinical study to predict a participant's clinical outcome based on baseline information. This complements our enrichment plan, for instance, about which we have different guidelines.

Clinical trial participant selection may also benefit from this enrichment technique. Throughout the clinical study, adherence can be tracked and enhanced with AI. Tools like electronic medication tracking, smartphone alerts and reminders, and tracking of missed clinical sessions can be used for this. All of this may result in possible non-adherence alarms and, ideally, give us the fundamental knowledge we need to deal with it. Actually, clinical research has already employed AI to enhance medication adherence. Specifically, this is done by using programs to remotely assess adherence, such as digital biomarkers that employ facial and verbal expressions. I can think of one particular instrument that industry and academia are working on to monitor adherence in that way. Participants' access to pertinent trial material and retention may both be enhanced by this technology.[35]

5.4 Drug Repurposing

Artificial Intelligence (AI) presents a promising way forward through accelerated drug repurposing, which enables researchers to analyze large datasets and uncover hidden connections between existing drugs, disease targets, and potential treatments. This approach has several benefits: First, it can expand the search for effective therapies by identifying unexpected connections between drugs and potential new targets; second, it can help mitigate limitations by predicting and minimizing side effects; third, it can optimize drugs for repurposing; and fourth, it can help overcome intellectual property obstacles.[35]

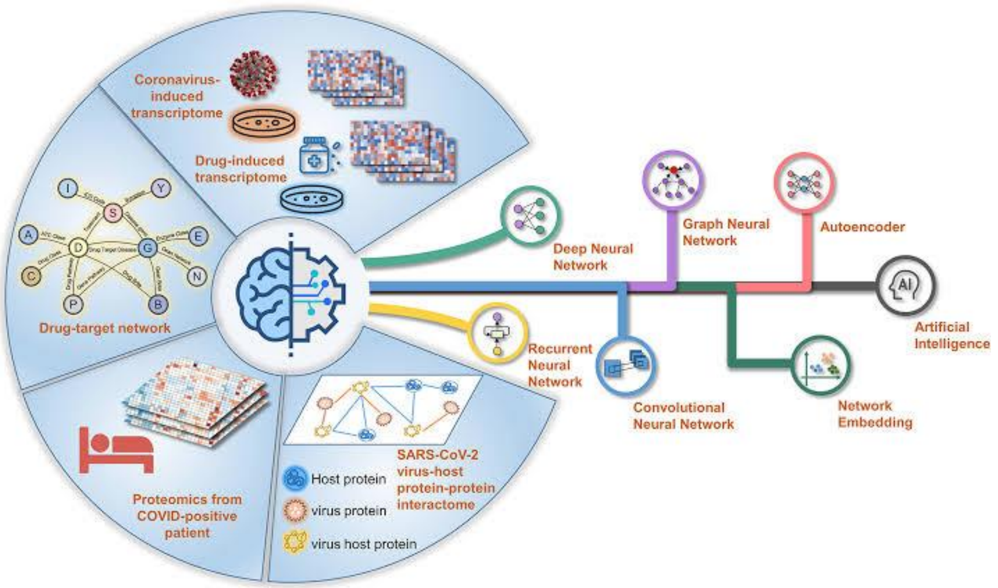


Fig. No: 14.

Deep Learning for Drug Repurposing[35]

TxGNN is the first AI model created especially to find potential medications for uncommon diseases and untreated situations. For almost 17,000 ailments, many of which have no known cures, it found therapeutic candidates among the available medications. This is the most number of diseases that can currently be handled by a single AI model. The model might be used for many more disorders than the 17,000 it was tested on in the early trials, the researchers say. Harvard Medical School researchers led the study, which was published in Nature Medicine on Wednesday.[35]

5.5 Post Market Surveillance

An additional chance to find any adverse drug reactions (ADRs) prior to a new medicine going on sale is through post-market surveillance. It entails sorting through vast amounts of data by definition, which is a laborious process for researchers and physicians to complete by hand and adds years to the R&D timeframes for new treatments. Generative AI excels in analyzing vast amounts of raw data and drawing conclusions about adverse drug reactions (ADRs) and other anomalies in collaboration with human clinicians. And that's only one facet of its post-market potential; it can also assist with updating marketing and packaging materials, generating documentation, and corresponding with doctors and patients.[36]

Ways in which AI benefits in Post Market Surveillance

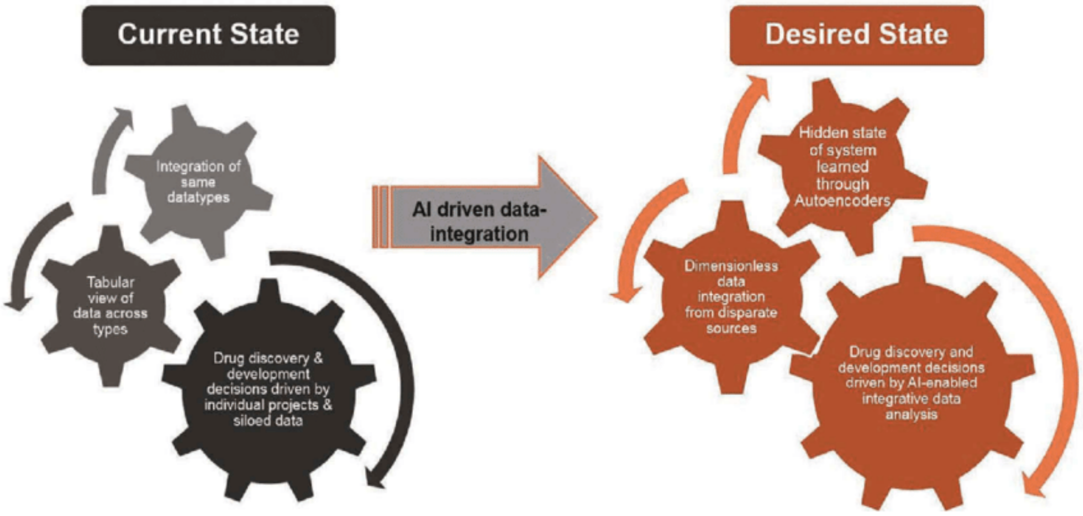
- Chatbots: You may be aware of the 24/7 services offered by a number of businesses. Customers can complain to a chatbot from anywhere in the world, and it can address their issues right away. As a result, it becomes easier and faster to file complaints and reply to customers. Consumers can contact the chatbot directly to file a complaint. The bot can offer the customer pre-programmed automated responses if it recognizes the report as a question or comment, personalizing the exchanges. If there is ever a serious incident during the encounter, there may also be a choice to escalate the matter to a real person.
- 2. Complaint Analysis: Using exact pre-established procedures and criteria, customer messages can be quickly examined as legitimate or invalid complaints. If there is a complaint, the complaint management staff may step in. AI can assess the complaint right away, and swift action will be made. Additionally, the AI will notify users if a follow-up

on the complaint is necessary or if any crucial information pertaining to the complaint is lacking.

- AI is capable of actively monitoring and identifying events that may be subject to reporting. For instance, the AI can determine if a complaint should be reported if the incident involved the medical device and resulted in death or a significant decline in health, or if the incident was such that it could cause death or a major decline in health if it happened again. Data points and trends from earlier data will serve as the foundation for this. The complaints investigation team will be notified of such incidents right away for additional assessment and action. The AI will also assess the incident's seriousness and choose the appropriate reporting schedule.
- Report Generation: Using pre-defined data sets and information, AI can automate the creation of reports. This shortens the time needed to create regulatory reports such as Periodic Safety Update Reports (PSUR) and Post Market Surveillance Reports (PMSR). A standard template can be filled with all of the information that has been examined. Data like sales volume and the number of major and non-serious occurrences should be recorded based on the year and location of the incidents.
- Signal detection and trend reporting A trend is a visual depiction of the recorded complaints (the tendency of a variable) over time, showing whether they have increased, decreased, or remained constant. It can be used to ascertain whether there has been a notable rise in consumer complaints in comparison to their predictable frequency. A control chart is typically used to illustrate the patterns. It should be mentioned that trend evaluation is done using reliable statistical techniques.[36]

4. Synergy between Data integration and Artificial Intelligence (AI)

The pharmacological data science method offers effective medication candidate selection for tasks including development and repurposing. The injection of a medication to alter a gene product known as the drug target forms the conceptual basis of contemporary pharmacological treatment of a disease. Through the application of pharmacological data science, we have created a novel concept known as "process pharmacology," which centres medication therapy around the disease as defined by the biological processes involved in its pathophysiology. Merely serving as a conduit between the medicine and the altered biological processes, the molecular drug targets directly support the therapeutic context of illness treatment. Globally accessible databases contain the genetic determinants of the pharmacological targets, respectively.[36]



7.1 Data Integration

- Improved Efficiency for Medical Professionals: One major benefit is that medical professionals save a lot of time on each patient. They have instant access to thorough and useful insights, including integrated patient histories, real-time health data, and predictive analytics, thanks to unified data platforms. This effectiveness enables faster, more precise decision-making in patient treatment in addition to cutting down on the amount of time spent on administrative duties.
- Streamlined Workflows: Healthcare practitioners can ensure that patients receive timely and effective care by streamlining their workflows with integrated data at their fingertips.
- Improved Decision-Making: Clinicians are equipped with thorough insights thanks to a consolidated view of patient data, including inputs from wearable devices and medical records. Treatment strategies and diagnosis are improved by this data-driven method.
- Better Patient Outcomes: By ensuring that each intervention is founded on the full patient medical history, integrated data systems improve patient outcomes.
- Cost-Effective Care Delivery: Healthcare firms can attain operational efficiency and save a substantial

amount of money by getting rid of data silos and redundancies.

- Real-time Data Access: While on-site equipment and prompt medical response are essential in critical care, real-time access to patient data, which offers a thorough patient history and current health condition, supports these efforts. This can play a pivotal role in directing suitable interventions and continuous observation. Furthermore, real-time data integration's real power is found in its capacity to assist long-term care management by detecting patterns and anomalies, which helps to spot any health problems before they become serious.

7.2 Artificial Intelligence

Artificial Intelligence (AI) has revolutionized the way we approach scientific research and has led to groundbreaking discoveries and innovations across various fields. From drug discovery to climate modeling, AI is transforming the scientific landscape in unprecedented ways. In this article, we will explore some of the recent advances in science that have been made possible by AI. In the ever-evolving landscape of science and medicine, Artificial Intelligence (AI) has emerged as a transformative force, propelling advancements at an unprecedented pace. From drug discovery to personalized medicine, AI is reshaping the way we approach research, diagnosis, and treatment. This article delves into the remarkable contributions of AI to recent advances in science and medicine, illuminating the promising future it unlocks.

There are wide range of major applications of AI that are becoming very common and is providing a great benefit in our day-to-day life. Some of the applications and uses of artificial intelligence are as follows

- **Drug Discovery and Development**

The field of drug development and discovery has benefited greatly from artificial intelligence. Artificial intelligence (AI) algorithms are being used by pharmaceutical companies to more effectively find possible medication candidates by analysing large amounts of chemical and biological data. AI-driven drug discovery has sped up the process of identifying new chemicals, which has resulted in the creation of breakthrough therapies for illnesses like cancer, Alzheimer's disease, and uncommon genetic abnormalities.

- **Genomics and Precision Medicine**

AI is essential to the advancement of personalised medicine and genomics research. AI systems are able to recognise genetic variants linked to illnesses, forecast patient outcomes, and suggest customised therapy regimens through the analysis of massive genomic datasets. This cleared the path for precision medicine, which improves patient outcomes and care by tailoring medicines based on a patient's genetic composition.

- **Material Science**

Artificial Intelligence is being utilised in material science and nanotechnology to expedite the identification and creation of novel materials possessing distinctive characteristics. AI systems are helping scientists create new materials for use in electronics, energy storage, and medical applications by modelling atomic structures and forecasting material behaviour. This might completely transform industries and spur the development of more sustainable and effective technology.

- **Biomedical Imaging and Diagnostics**

The precision and efficacy of medical imaging procedures have been greatly enhanced by AI-powered image analysis and diagnostic technologies. Medical pictures such as MRI scans, X-rays, and histopathology slides can be analysed using machine learning algorithms to help with early disease identification, tumour categorization, and treatment planning. As a result, patient outcomes have improved and healthcare personnel' competencies have been strengthened.

- **Personalized Medicine: Tailoring Treatment to Individuals**

Personalised healthcare is one of the areas where AI in medicine is having the biggest effects. Artificial Intelligence (AI) makes it easier to identify genetic markers linked to diseases by analysing genomic data using machine learning techniques. This allows medical professionals to customise treatment plans for each patient according to their genetic composition, maximising effectiveness and reducing side effects. This allows medical professionals to customise treatment plans for each patient according to their genetic composition, maximising effectiveness and reducing side effects.

- **Empowering Clinical Decision-Making**

In the era of big data, healthcare providers are inundated with vast amounts of patient information. AI offers a solution by providing clinical decision support systems that sift through complex datasets to extract valuable insights. By analyzing electronic health records, medical literature, and real-time patient data, AI algorithms assist clinicians in diagnosing diseases, predicting treatment outcomes, and devising personalized care plans, ultimately enhancing the quality and efficiency of healthcare delivery.

- **Advancing Disease Prevention and Public Health**

Beyond individual patient care, AI contributes to broader public health initiatives by leveraging predictive analytics and epidemiological modeling. By analyzing population-level data, AI can forecast disease outbreaks, identify high-risk communities, and inform preventive interventions. From tracking the spread of infectious diseases to optimizing vaccination strategies, AI-powered tools play a vital role in safeguarding public health and mitigating health disparities.

- **Facilitating Research and Discovery**

In the realm of scientific research, AI serves as a catalyst for innovation and discovery. Natural Language Processing (NLP) algorithms analyze vast repositories of scientific literature, accelerating the pace of knowledge dissemination and facilitating interdisciplinary collaborations. By uncovering hidden patterns, identifying novel hypotheses, and streamlining literature reviews, AI empowers researchers to push the boundaries of scientific understanding and tackle complex challenges more effectively.

5. Challenges and Limitations

8.1 Faced by Data Integration

- Heterogeneity of Data Formats: From many sources may be captured in different forms due to heterogeneity, which can make it challenging to integrate them into a cohesive whole at times. It's similar to attempting to fit together jigsaw pieces that don't match. Databases, various IT systems, and even paper records can provide data for pharmaceutical companies. Therefore, in order to be used successfully, they must be processed correctly.
- Ensuring Data Quality and Reliability: Providing Reliability and Quality Data For judgments to be made correctly, all data must be complete and accurate. What if there are mistakes or the information is out of date? Erroneous analysis and conclusions may result from this. This is why using strong validation and verification processes in the data integration process is crucial.
- Data Security and Privacy Issues: Serious repercussions, such patient health issues or financial losses for the business, are frequently caused by data integrity breaches. To guarantee data security, every business needs thus put in place the proper policies and resources, which could include:
  - System validations,
  - Backups, and
  - Access path audits.

Artificial Intelligence in Drug Discovery: Case Studies and Real-World Applications

1. Identification of Novel Drug Candidates

Case Study: BenevolentAI

- **Application**: BenevolentAI, a UK-based biotech company, leveraged AI to identify existing drugs that could treat new diseases.
- **Outcome**: During the COVID-19 pandemic, BenevolentAI's platform identified Baricitinib, originally for rheumatoid arthritis, as a treatment for severe COVID-19 symptoms. This discovery took weeks, compared to the years traditional methods would have required.
- **Impact**: The drug was approved for emergency use, showcasing AI's ability to repurpose existing drugs for new indications quickly (Dias & Torkamani, 2019).

2. Deep Learning in Drug Design

Case Study: Insilico Medicine

- **Application**: Insilico Medicine used generative adversarial networks (GANs) to design novel molecules with specific therapeutic properties.
- **Outcome**: In a landmark case, the company identified a promising drug candidate for idiopathic pulmonary fibrosis (IPF) in less than 18 months.
- **Impact**: The AI-generated molecule entered preclinical trials, demonstrating how AI can drastically reduce drug development time and improve precision in designing new compounds (Popova et al., 2018).

3. Predicting Drug-Drug Interactions (DDIs)

Case Study: IBM Watson for Drug Safety

- **Application**: IBM Watson's AI model analyzed vast amounts of biomedical literature to predict adverse drug interactions.
- **Outcome**: It identified potential DDIs that were previously unrecognized by human researchers.
- **Impact**: This proactive approach helps mitigate risks during drug development and ensures patient safety (Percha et al., 2012).

4. AI-Driven Target Identification

Case Study: Atomwise

- **Application**: Atomwise used AI to identify potential inhibitors for Ebola virus. The technology analyzed billions of molecules to predict their binding affinity with target proteins.
- **Outcome**: The AI models shortlisted molecules within days, which would traditionally take years through experimental screening.
- **Impact**: This expedited the path toward developing treatments for Ebola and highlighted AI's role in tackling global health crises (Russell & Norvig, 2016).

5. Integration of Multi-Omics Data

Case Study: Multi-Omics Drug Development by AstraZeneca

- **Application**: AstraZeneca used AI to integrate genomic, transcriptomic, and proteomic data to identify biomarkers and therapeutic targets.
- **Outcome**: The insights allowed the company to design targeted therapies for cancer, including combinations for drug-resistant tumors.
- **Impact**: AI enhanced precision medicine, leading to more effective treatments tailored to individual patients (Ivanisevic & Sewduth, 2023).

6. Virtual Clinical Trials

Case Study: PathAI in Oncology

- **Application**: PathAI developed AI algorithms to analyze pathology data and predict the efficacy of oncology drugs.

- **Outcome**: The insights provided by AI streamlined trial designs, reducing the number of patients needed and improving trial success rates.
- **Impact**: Virtual trials save costs and time, making drug development more efficient and patient-centric (Askin et al., 2023).

7. Predicting Toxicity and Safety Profiles
Case Study: Novartis and Microsoft Collaboration

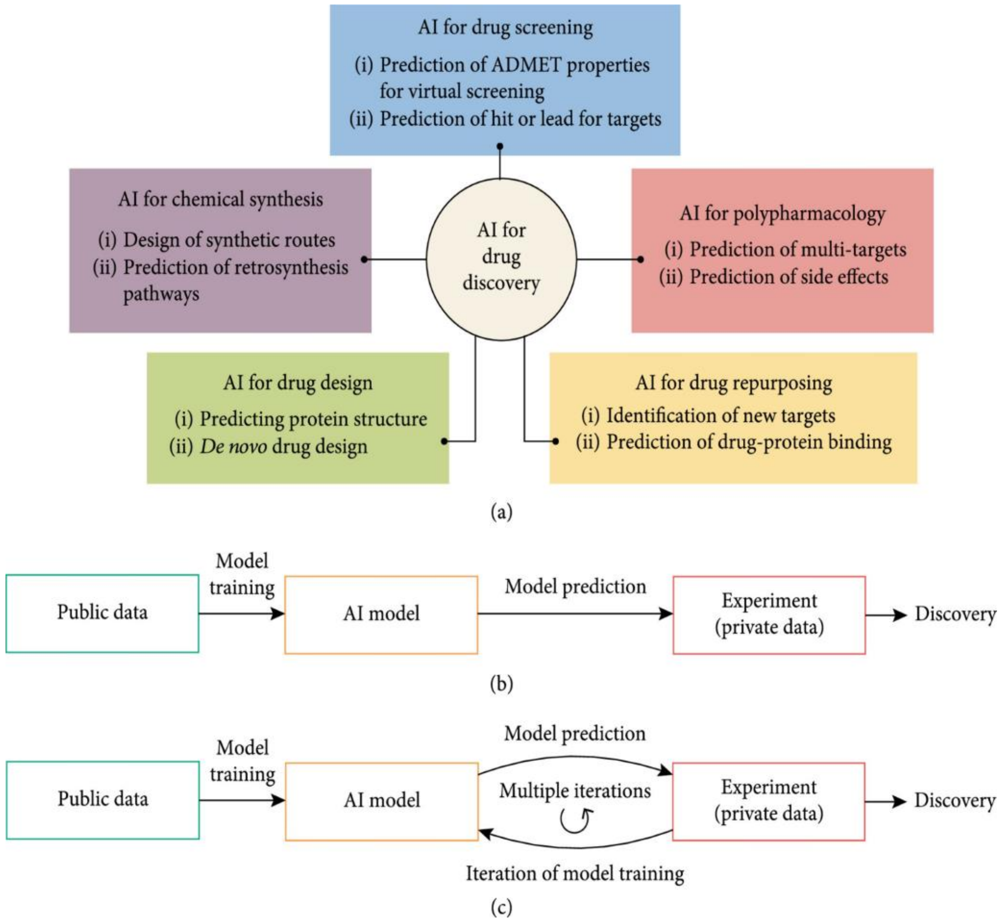
- **Application**: AI was employed to predict the safety and toxicity profiles of drug candidates during early-stage development.
- **Outcome**: By identifying potential safety issues early, Novartis avoided costly late-stage failures.
- **Impact**: This partnership demonstrated the potential of AI to de-risk drug development pipelines (Kirtania et al., 2024).

Future Prospective
Emerging AI Technologies

Blockchain technology and artificial intelligence (AI) have the potential to completely transform drug research by establishing cooperative ecosystems and offering novel approaches. An extensive examination of these developments is provided below:

- Generative AI Models for Molecular Design
  - The creation of novel molecular structures with desired properties has been demonstrated to be possible with generative AI models, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs). These models learn from existing chemical data to generate new compounds, potentially speeding up the drug discovery process, but there are still issues with ensuring the chemical validity and synthesizability of AI-generated molecules. As stated in a blog post on Practical Cheminformatics, generative models offer useful ideas, but substantial manual curation is still necessary to weed out chemically implausible suggestions.[37]
- Use of Blockchain for Secure, Decentralized Data Integration Pharmaceutical research data security and integrity are improved by blockchain technology, which provides a decentralized and unchangeable ledger system. Blockchain can simplify drug supply chain supervision, enhance clinical trial management, and enable safe sharing of private health information by utilizing smart contracts and encryption. Blockchain has the potential to improve privacy, openness, and accessibility in healthcare settings, according to a thorough analysis published in the National Center for Biotechnology Information (NCBI).[38]

Through the use of data and computer analysis, AI helps researchers find therapeutic targets, comprehend disease causes, and expedite the drug-discovery process.[38]



- ***Broader Collaboration***

Partnership between Pharmaceutical Companies, Academies and Technological Firms:

Collaboration across multiple industries is required for the incorporation of blockchain and artificial intelligence into medicine development. To take advantage of specialized knowledge and resources, pharmaceutical corporations are increasingly collaborating with academic institutions and tech companies. As demonstrated by the creation of AlphaFold, the partnership between Google DeepMind and pharmaceutical firms has resulted in notable progress in the prediction of protein structures. These collaborations are essential to converting AI research into useful drug discovery applications.[38]

- Increased Public – Private Investments in AI driven Research:
  - Public-private investments to promote innovation have increased as a result of the realization of AI's revolutionary potential in healthcare. Initiatives that promote AI-driven medication discovery and development are being funded by both public and private organizations. The Times article highlights the substantial economic benefit of investing in AI technology by stating that McKinsey predicts that generative AI may provide $60–110 billion in value for pharmaceutical businesses yearly.[38]
- Through the use of data and computer analysis, AI helps researchers find therapeutic targets, comprehend disease causes, and expedite the drug-discovery process.[39]

CONCLUSION

Data integration has developed into a tool that helps healthcare organisations, carers, and many others improve their normal business procedures in recent years. Medical data integration has several advantages for hospitals who have already integrated our cutting-edge technology into their operations, including enhanced patient experiences, efficient disease prevention, and workflow automation. In conclusion, the integration of AI into scientific research has opened up new frontiers and accelerated progress across diverse disciplines. As AI continues to evolve, its potential to drive scientific discovery and innovation remains immense. The collaboration between AI technology and scientific expertise holds promise for addressing some of the most pressing challenges facing humanity and shaping the future of scientific exploration. The ongoing advancements in AI-driven science are poised to transform our understanding of the world and improve the quality of life for generations to come. As we stand on the brink of a new era in science and medicine, the transformative power of AI is undeniable. From revolutionizing drug discovery to enabling personalized healthcare, AI holds the key to addressing some of the most pressing challenges in healthcare and biomedical research. By harnessing the capabilities of AI-driven technologies, we can unlock new frontiers of knowledge, enhance patient care, and ultimately, transform the future of health and well-being. As we continue to harness the potential of AI, the possibilities for innovation and progress are limitless.

REFERENCES

1. Tobin D. The data integration landscape in healthcare: Improving patient care. Integrate.io. 2023 [cited 2024 Mar 10]. Available from: https://www.integrate.io/blog/healthcare-data-integration/
2. Why Data Integration is Important? [Internet]. WADIC. 2019 [cited 2024 Mar 10]. Available from: https://wadic.net/data-integration-importance/
3. Richman J. 11+ most common data integration challenges & solutions [Internet]. Estuary.dev. Estuary; 2023 [cited 2024 Mar 10]. Available from: https://estuary.dev/data-integration-challenges
4. BasuMallick C. Here’s why data integration is important to organizations [Internet]. Spiceworks. [cited 2024 Mar 10]. Available from: https://www.spiceworks.com/tech/devops/articles/data-integration/amp/
5. King T. The history of data integration: A 90-second primer [Internet]. Solutionsreview.com. Solutions Review; 2018 [cited 2024 Mar 10]. Available from: https://solutionsreview.com/data-integration/the-history-of-data-integration-a-90-second-primer/
6. EdD PMP, EdD F. Demystifying Ai in healthcare: Historical perspectives and current considerations, 2021; 8(1): 59–66. Available from: https://www.proquest.com/openview/1d554e977b9e48119fb07abddd9fcf2d/1?pq-origsite=gscholar&cbl=2037550
7. Jubilantbiosys.com. [cited 2024 Mar 10]. Available from: https://www.jubilantbiosys.com/integrated-drug-discovery-services/
8. Ivanisevic T, Sewduth RN. Multi-omics integration for the design of novel therapies and the identification of novel biomarkers. Proteomes [Internet], 2023; [cited 2024 Mar 10]; 11(4): 34. Available from: https://www.mdpi.com/2227-7382/11/4/34
9. Percha B, Garten Y, Altman RB. Discovery and explanation of drug-drug interactions via text mining. Pac Symp Biocomput, 2012; 410–21.
10. Drug-drug interaction [Internet]. Hiv.gov. [cited 2024 Mar 10]. Available from: https://clinicalinfo.hiv.gov/en/glossary/drug-drug-interaction
11. What are the most common data integration standards and protocols and how do you apply them? [Internet]. Linkedin.com. www.linkedin.com; 2023 [cited 2024 Mar 10]. Available from: https://www.linkedin.com/advice/3/what-most-common-data-integration-standards-protocols
12. Bizdata. Bizdata Inc.; 2020 [cited 2024 Mar 10]. Available from: https://www.bizdata360.com/how-to-connect-and-automate-with-ipaas/

13. Hopkins AL. Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol [Internet], 2008; 4(11): 682–90. Available from: http://dx.doi.org/10.1038/nchembio.118
14. Berger SI, Iyengar R. Network analyses in systems pharmacology. Bioinformatics [Internet], 2009; 25(19): 2466–72. Available from: http://dx.doi.org/10.1093/bioinformatics/btp465
15. Learn more. Pharmacovigilance [Internet]. Who.int. [cited 2024 Mar 10]. Available from: https://www.who.int/teams/regulation-prequalification/regulation-and-safety/pharmacovigilance
16. CTG - datasharing [Internet]. Albany.edu. [cited 2024 Mar 10]. Available from: https://www.ctg.albany.edu/publications/datasharing/
17. Lötsch J, Ultsch A. Process pharmacology: A pharmacological data science approach to drug development and therapy. CPT Pharmacometrics Syst Pharmacol [Internet]. 2016 [cited 2024 Mar 10];5(4):192–200. Available from: http://dx.doi.org/10.1002/psp4.12072
18. Kamat S. Data integration in healthcare & pharma industries [Internet]. algoscale. [cited 2024 Mar 10]. Available from: https://algoscale.com/blog/transformative-power-of-data-integration-in-the-healthcare-and-pharma-industries/
19. Russell, S.J.; Norvig, P. Artificial Intelligence: A Modern Approach; Pearson Education Limited: London, UK, 2016.
20. Enholm, I.M.; Papagiannidis, E.; Mikalef, P.; Krogstie, J. Artificial intelligence and business value: A literature review. Inf. Syst. Front, 2022; 24: 1709–1734.
21. Wang, Z.; Li, M.; Lu, J.; Cheng, X. Business Innovation based on artificial intelligence and Blockchain technology. Inf. Process. Manag, 2022; 59: 102759.
22. Chen, X.; Xie, H.; Zou, D.; Hwang, G. Application and theory gaps during the rise of Artificial Intelligence in Education. Comput. Educ. Artif. Intell, 2020; 1: 100002.
23. Hossen, M.S.; Karmoker, D. Predicting the Probability of COVID-19 Recovered in South Asian Countries Based on Healthy Diet Pattern Using a Machine Learning Approach. In Proceedings of the 2nd International Conference on Sustainable Technologies for Industry 4.0, Dhaka, Bangladesh, December 2020; 19–20: 1–6.
24. Liu, J.; Ma, J.; Li, J.; Huang, M.; Sadiq, N.; Ai, Y. Robust watermarking algorithm for medical volume data in internet of medical things. IEEE Access, 2020; 8: 93939–93966.
25. Zhou, L. A rapid, accurate and machine-agnostic segmentation and quantification method for CT-Based COVID-19 diagnosis. IEEE Trans. Med. Imaging, 2020; 39: 2638–2652.
26. Bhattamisra, Subrat Kumar, Priyanka Banerjee, Pratibha Gupta, Jayashree Mayuren, Susmita Patra, and Mayuren Candasamy. 2023. "Artificial Intelligence in Pharmaceutical and Healthcare Research." Big Data and Cognitive Computing 7: https://doi.org/10.3390/bdcc7010010.
27. Blanco-González, Alexandre, Alfonso Cabezón, Alejandro Seco-González, Daniel Conde-Torres, Paula Antelo-Riveiro, Ángel Piñeiro, and Rebeca Garcia-Fandino. n.d. "The Role of AI in Drug Discovery: Challenges, Opportunities, and Strategies." Available from: https://arxiv.org/ftp/arxiv/papers/2212/2212.08104.pdf.
28. Santoro, Helen. 2022. "Bringing Principles of Ethics to AI and Drug Design." Scope, April 2022.
29. Zhou S-F, Zhong W-Z. Drug design and discovery: Principles and applications. Molecules [Internet]. 2017 [cited 2024 Mar 10];22(2):279. Available from: http://dx.doi.org/10.3390/molecules22020279
30. New drugs definition [Internet]. Law Insider. [cited 2024 Mar 10]. Available from: https://www.lawinsider.com/dictionary/new-drugs
31. Paul Ehrlich's foundational concept of "magic bullets" (Ehrlich, 1906) laid the groundwork for target-based drug discovery. Application of cell-based assays in drug discovery is extensively reviewed by Fink et al., 2016.
32. The design and conduct of Phase II trials are detailed by Chow et al., 2018.
33. Nih.gov. [cited 2024 Mar 10]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6342261/
34. The FDA provides a detailed guide for NDAs: https://www.fda.gov/.
35. The benefits of big data in drug development [Internet]. Contract Pharma. [cited 2024 Mar 10]. Available from: https://www.contractpharma.com/issues/2024-01-02/view_features/the-benefits-of-big-data-in-drug-development/
36. Lamberti MJ, Wilkinson M, Donzanti BA, Wohlhieter GE, Parikh S, Wilkins RG, et al. A study on the application and use of artificial intelligence to support drug development. Clin Ther [Internet], 2019; 41(8): 1414–26. Available from: http://dx.doi.org/10.1016/j.clinthera.2019.05.018
37. Srivastava S. The potential of AI in drug discovery and its impact on healthcare [Internet]. Appinventiv. 2023 [cited 2024 Mar 10]. Available from: https://appinventiv.com/blog/ai-in-drug-discovery/
38. SAP. What is enterprise integration? Data integration overview [Internet]. [cited 2024 Mar 10]. Available from: https://www.sap.com/africa/products/technology-platform/what-is-enterprise-integration/data-integration.html
39. Tyagi MR. The role of data integration in the pharmaceutical industry. Adeptia. 2024 [cited 2025 Jan 21]. Available from:

https://www.adeptia.com/blog/data-integration-in-pharmaceutical-industry.
40. Natural language processing in healthcare medical records. ForeSee Medical. [cited 2025 Jan 21]. Available from: https://www.foreseemed.com/natural-language-processing-in-healthcare.
41. Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. Genome Med., 2019; 11(1): 70. Available from: http://dx.doi.org/10.1186/s13073-019-0689-8
42. Kostetska L. How real-world data impacts drug development. Binariks.com. Binariks, 2024; [cited 2025 Jan 21]. Available from: https://binariks.com/blog/real-world-data-in-drug-development/
43. Kirtania MD, Sinha D, Biswas S, Sultana S, Kirtania R. Artificial Intelligence in modernization of pharmaceutical and healthcare industry: A review. Current Artificial Intelligence, 2024; [cited 2025 Jan 21]; 02(1): 34–43. Available from: https://www.eurekaselect.com/article/141341
44. Popova M, Isayev O, Tropsha A. Deep reinforcement learning for de novo drug design. Sci Adv., 2018; 4(7): eaap7885. Available from: http://dx.doi.org/10.1126/sciadv.aap7885
45. Askin S, Burkhalter D, Calado G, El Dakrouni S. Artificial Intelligence Applied to clinical trials: opportunities and challenges. Health Technol (Berl) [Internet], 2023; 13(2): 203–13. Available from: http://dx.doi.org/10.1007/s12553-023-00738-2
46. Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. Genome Med., 2019; 11(1): 70. Available from: http://dx.doi.org/10.1186/s13073-019-0689-8
47. Popova M, Isayev O, Tropsha A. Deep reinforcement learning for de novo drug design. Sci Adv., 2018; 4(7): eaap7885. Available from: http://dx.doi.org/10.1126/sciadv.aap7885
48. Percha B, Garten Y, Altman RB. Discovery and explanation of drug-drug interactions via text mining. Pac Symp Biocomput, 2012; 410–21.
49. Russell, S.J.; Norvig, P. Artificial Intelligence: A Modern Approach; Pearson Education Limited: London, UK, 2016.
50. Ivanisevic T, Sewduth RN. Multi-omics integration for the design of novel therapies and the identification of novel biomarkers. Proteomes [Internet]. 2023 [cited 2024 Mar 10]; 11(4): 34. Available from: https://www.mdpi.com/2227-7382/11/4/34
51. Askin S, Burkhalter D, Calado G, El Dakrouni S. Artificial Intelligence Applied to clinical trials: opportunities and challenges. Health Technol (Berl) [Internet], 2023; 13(2): 203–13. Available from: http://dx.doi.org/10.1007/s12553-023-00738-2
52. Kirtania MD, Sinha D, Biswas S, Sultana S, Kirtania R. Artificial Intelligence in modernization of pharmaceutical and healthcare industry: A review. Current Artificial Intelligence. 2024 [cited 2025 Jan 21]; 02(1): 34–43. Available from: https://www.eurekaselect.com/article/141341