M.TECH. INDUSTRIAL BIOTECHNOLOGY

[NANCY SANJAY GUPTA]

2023

"***Identification of potential DNMT1inhibitors in Alzheimer's therapeutics: a drug repurposing and machine learning approach***"

A DISSERTATION

SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

MASTER OF TECHNOLOGY

IN

INDUSTRIAL BIOTECHNOLOGY

Submitted by:

Nancy Sanjay Gupta

2K21/IBT/21

Under the supervision of

Prof. Pravir Kumar

Department of Biotechnology

Delhi Technological University

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

MAY, 2023

Delhi Technological University
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Nancy Sanjay Gupta, 2K21/IBT/21 student of M.Tech Industrial Biotechnology, hereby declare that the project Dissertation titled "**Identification of potential DNMT1 inhibitors in Alzheimer's Disease therapeutics: a drug repurposing and machine learning approach**" which is submitted by me to the department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. The work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

The content present in the thesis has been communicated in Scopus indexed journal and IEEE conference with the below mentioned details:

1. **Title of Paper:** "Perspective of artificial intelligence in healthcare data management: A journey towards precision medicine."

   Journal : Computers in Biology and Medicine (Elsevier)

   Impact Factor: 6.7

   Author's Name: Nancy Sanjay Gupta and Pravir Kumar

   Status of paper: Accepted

   Date of Communication: 11th April, 2023

   Date of Acceptance: 20th May, 2023

2. **Title of Paper:** "DNMT1 inhibitors in Alzheimer's Disease: A drug repurposing approach through FDA-approved drugs."

   Conference: IEEE Bangalore Humanitarian Technology Conference (IEEE B-HTC 2023) organized by JSS Academy of Higher Education and Research, JSS Hospital, Mysuru,

   Author's Name: Nancy Sanjay Gupta and Pravir Kumar

   Status of Paper: Accepted

   Date of Acceptence: 19th February, 2023

   Date of Conference: 24th and 25th March, 2023

3. **Title of Paper:** "TDP-43 Inhibitors in Amyotrophic Lateral Sclerosis: An Application of Drug Repurposing Approach Using FDA Approved Drugs."

   Conference: Computational Intelligence and Sustainable Engineering (CISES-2023)

   Author's Name: Nancy Sanjay Gupta and Pravir Kumar

   Status of Paper: Accepted

   Date of Acceptence: 19th February, 2023

   Date of Conference: 24th and 25th March, 2023

Place: Delhi

Date: MAY, 2023 **NANCY SANJAY GUPTA**

Proof of Publication

A. Journal: Computers in Biology and Medicine

Computers in Biology and Medicine

Available online 30 May 2023, 107051

In Press, Journal Pre-proof What's this?

Perspective of artificial intelligence in healthcare data management: A journey towards precision medicine

Nancy Sanjay Gupta, Pravir Kumar

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.compbiomed.2023.107051 Get rights and content

Abstract

Mounting evidences have highlighted the implementation of big data handling and management in healthcare industry to improve the clinical services. Various private and public companies have generated, store, and analyze different types of big healthcare data, such as omics data, clinical data, electronic health records, personal health records, and sensing data with the aim to move in the direction of precision medicine. Additionally, with the advancement in technologies, researchers are curious to extract the potential involvement of artificial intelligence and machine learning on big healthcare data to enhance the quality of patient's life. However, seeking solutions from big healthcare data requires proper management, storage, and analysis, which imposes hinderance associated with big data handling. Herein, we briefly discuss the implication of big data handling and the role of artificial intelligence in precision medicine. Further, we also highlighted the potential of artificial intelligence in integrating and analyzing the big data that offer personalized treatment. In addition, we briefly discuss the applications of artificial intelligence in personalized treatment, especially in neurological diseases. Lastly, we discuss the challenges and limitations imposed by artificial intelligence in big data management and analysis to hinder precision medicine.

B. IEEE Bangalore Humanitarian Technology Conference (IEEE B-HTC 2023)

JSS ACADEMY OF HIGHER EDUCATION AND RESEARCH
JSS HOSPITAL, MYSURU

Certificate

to certify that

Nancy Sanjay Gupta, Pravir Kumar

has presented paper entitled

DNMT1 inhibitors in Alzheimer's Disease: A drug repurposing approach through FDA approved drugs

in **IEEE Bangalore Humanitarian Technology Conference (IEEE B-HTC 2023)** organized by JSS Academy of Higher Education and Research, JSS Hospital, Mysuru, during 24th and 25th March 2023

Dr. Sudarshan Patil Kulkarni
Chair - IEEE Mysore Subsection
Professor, Dept. of ECE,
SJCE (JSSSTU), Mysuru

Dr. Prashant M Vishwanath
Dean (Research)
JSS AHER, Mysuru -15

C. IEEE- Computational Intelligence and Sustainable Engineering (IEEE CISES-2023)

GL BAJAJ Institute of Technology & Management
FIND YOUR SPARK
Approved by AICTE & Affiliated to AKTU

IEEE Computational Intelligence Society

2nd International Conference
on
Computational Intelligence and Sustainable Engineering Solutions
(CISES-2023)

28-30 April, 2023

CERTIFICATE

Certified that Ms./Mr./Dr. Nancy Sanjay Gupta
from *Delhi Technological University, New Delhi*
has participation / presented paper entitled *TDP-43 Inhibitors in Amyotrophic Lateral Sclerosis: An Application of Drug Repurposing Approach Using FDA Approved Drugs*

in Three Days 2nd International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES-2023) Technically Co-sponsored by IEEE-CIS on 28th April to 30th April, 2023 organized by Department of Master of Computer Applications, G.L. Bajaj Institute of Technology & Management Greater Noida, (U.P.) India.

Convener
Dr. Sanjeev Kumar

Conference Chair
Prof. (Dr.) Madhu Sharma Gaur

General Chair
Prof. (Dr.) Manas Kumar Mishra

DEPARTMENT OF BIOTECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "**Identification of potential DNMT1 inhibitors in Alzheimer's therapeutics: a drug repurposing and machine learning approach**" which is submitted by Nancy Sanjay Gupta, 2K21/IBT/21, Department of Biotechnology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge, this work has not been submitted in part of full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Date: MAY, 2023

Prof. Pravir Kumar
SUPERVISOR

Professor and Head,
Department of Biotechnology
Delhi Technological University
Shahbad Daulatpur Village,
Rohini, Delhi 110042

ACKNOWLEDGEMENT

I want to convey my sincere gratitude and thankfulness to Heavenly Father, for giving me patience and energy to finish my thesis. I owe a debt of gratitude to Prof. Pravir Kumar, my thesis adviser, and mentor, for his steadfast support, wise counsel, and priceless insights during this research project, and I now have a strong enthusiasm for the research career. My intellectual development has been positively affected by Dr. Rashmi Ambasta's trust, guidance and support. I want to express my deep gratitude to my family for their unflagging support, unending love, and everlasting faith in my skills.

I also want to express my gratitude to the Molecular Neuroscience and Functional Genomics laboratory team for their cooperation, friendliness, and desire to impart their knowledge. I would especially want to extend a thankful gesture to Dr. Rohan Gupta and Ms. Smita Modi for their ongoing support and guidance during every stage of my research journey. Their thoughtful comments and helpful criticism considerably improved the caliber of this thesis. My lab partner Mr. Shubham Kumar Shrivastav has made significant contributions, and I would like to thank him for his insightful comments and lively conversations. I also want to thank the faculty

and staff of Delhi Technological University's Department of Biotechnology for giving me a wealth of resources and a supportive academic atmosphere, both of which were essential for finishing my project.

Finally, I want to express my gratitude to all my peers and coworkers who have helped me along the way by offering advice, support, and stimulating conversation. Without the combined efforts and assistance of everyone indicated above, this thesis probably would not have been feasible.

NANCY SANJAY GUPTA

2K21/IBT/21

ABSTRACT

The first part of the thesis examines the use of technology such as big data, artificial intelligence (AI) as well as machine learning (ML) in cognitive healthcare, with a focus on personalized medicine for neurodegenerative diseases and drug discovery and development. Healthcare systems throughout the world face considerable problems from neurological illnesses including Parkinson's disease, Alzheimer's disease, Amyotrophic Lateral Sclerosis, and Huntington's disease. As a result, novel methods of medication development, diagnosis, and treatment are required. The first part of the thesis analyzes the state of neurological healthcare today and the shortcomings of conventional approaches to treating the complexity and variety of neurodegenerative illnesses. It then explores how utilizing enormous amounts of data from sources including genomes, proteomics, imaging, and medical records might change drug development and precision medicine. Intending to identify disease causes and enhance therapeutic approaches, it investigates the incorporation of multi-omics data along with the creation of computer models. The thesis also addresses the difficulties and moral issues related to the application of AI and ML in brain-related treatment in its literature review section. It speaks to the necessity for open and strong validation frameworks as well as data confidentiality, unfairness, and interpretability challenges. This thesis illustrates the promise of AI and ML in improving neurological healthcare through a thorough examination of research, case studies, and computer experiments. It emphasizes how important it is for researchers, physicians, as well as business stakeholders to work together to fully utilize AI and ML for individualized and successful therapies in neurodegenerative illnesses.

In the second part of the thesis, drug discovery and development, which is a segment of the literature review has been considered for the research work for the thesis. Using machine learning algorithms, I predicted drugs to stop the progression of Alzheimer's disease by inhibiting the DNMT1 protein. Out of the drugs mentioned in the predicted list, I have retrieved the best-binding drugs which limit the disease continuation. This work was performed using multiple computational biology tools like virtual screening and molecular docking, which also work on machine learning algorithms. The resulting drugs can be studied in experimental labs to bring the results from the bench to the bedside.

Finally, this research establishes the groundwork for future developments in novel drug discovery and personalized medicine for neurodegenerative illnesses, notably Alzheimer's Disease, and adds to the corpus of understanding in the discipline of AI and ML in brain healthcare.

Keywords: Alzheimer's disease, DNA Methyltransferase 1, artificial intelligence and machine learning, big data, Molecular docking, MD simulation, Binding affinity, personalized medicine.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

AD	Alzheimer's disease
NDD	Neuro Degenerative Disorders
DNMT 1	DNA Methyltransferase 1
ML	machine learning
DL	deep learning
AI	Artificial intelligence
IoT	Internet of things
IBM	International business machine
SNP	Single neucleotide polymorphism
GWAS	Genome wide association study
UTR	Untranscribed region
RFTA	Replication Foci Targeting Sequence
BBB	Blood brain barrier
SVM	Support Vector Machine

CHAPTER 1

INTRODUCTION

Rapid advancement are attained in allied areas of neurobiology along with artificial intelligence (AI) in recent years [1]. Investigations based on AI was closely linked to neurology and psychology at the beginning of the computer era [2], [3]. The interaction has, however, diminished significantly in recent years due to the tremendous complexity growth and academic boundary consolidation of both disciplines [2]. Examining biological intellect carefully has two advantages for the development of AI. First, unlike and in addition to the statistical and logically based techniques and concepts that have predominated conventional approaches to AI, neuroscience offers a rich source of motivation for new kinds of algorithms and architectures [4]. An algorithm's affirmation is a crucial part of a larger general intelligence system so that it can be applied in the study of the brain [3]. When determining how to most effectively allocate resources for a long-term research initiative, these hints can be crucial. For example, in the case, if a tool is not performing at the stage needed or expected but we notice it is vital to brain function, we can conclude that improved engineering efforts aimed at making it operate in artificial systems are likely to be successful [4]. John McCarthy in 1956, coined the terminology of Artificial intelligence and also introduced the world to its principles [5]. AI was entirely based on complex algorithms and the human mind mimicking software for analysis and solving crucial problems like decision-making ability, visual perception, and speech recognition [6]. Tremendous chances and opportunities are created in the healthcare industry because of the involvement of artificial intelligence techniques in computer work and the power to access huge amounts of data with great ease [7]. AI has multiple applications in the field of healthcare including developing protocols, preventing diseases, monitoring patients, keeping a record of medical history, and even helping researchers and clinicians to extract important information from the flood of data [8] **(figure1)**. Innovations in AI have shown legitimate implementations in some critical situations like performing skin cancer analysis better than dermatologists [9], [10] and predicting mortality using prostheses following cardiac surgery [11]. Performance of Ai has been seen from the nano level to the mega level but the usage rate is much lower in the field of healthcare and medicine as compared to other sectors like mechanics,

electronics, electricals, etc. [12], [13]. Intensified loads of neurodegenerative disorders are considered a major reason for speedy aging in the population and at a later date resulting in extravagant expenses in treatment, assistance in the patient's care, hospitalization, and even higher mortality rates [14]. The criteria that differentiate NDDs from neurological disorders is the neuronal dysfunction or neuronal loss in specified regions of the spinal cord or brain. Several physical and mental impairments are witnessed in people suffering from age-related neurological disorders also known as NDDs like speech issues, cognitive decline, forgetfulness, anxiety, continuous changes in mood, and even decline in other organs of the body is also seen [14]. These disorders can be a result of electrical, structural, or biochemical abnormalities in the neurons, spinal cord, and brain. AD, a very common disease among the list of neurodegenerative disorders is predicted to affect approximately 7 million Europeans by 2040. The complete cure for the disease is still unknown, still, the diagnosis, treatment, and care comprise around 130 billion euros per year

Figure 1

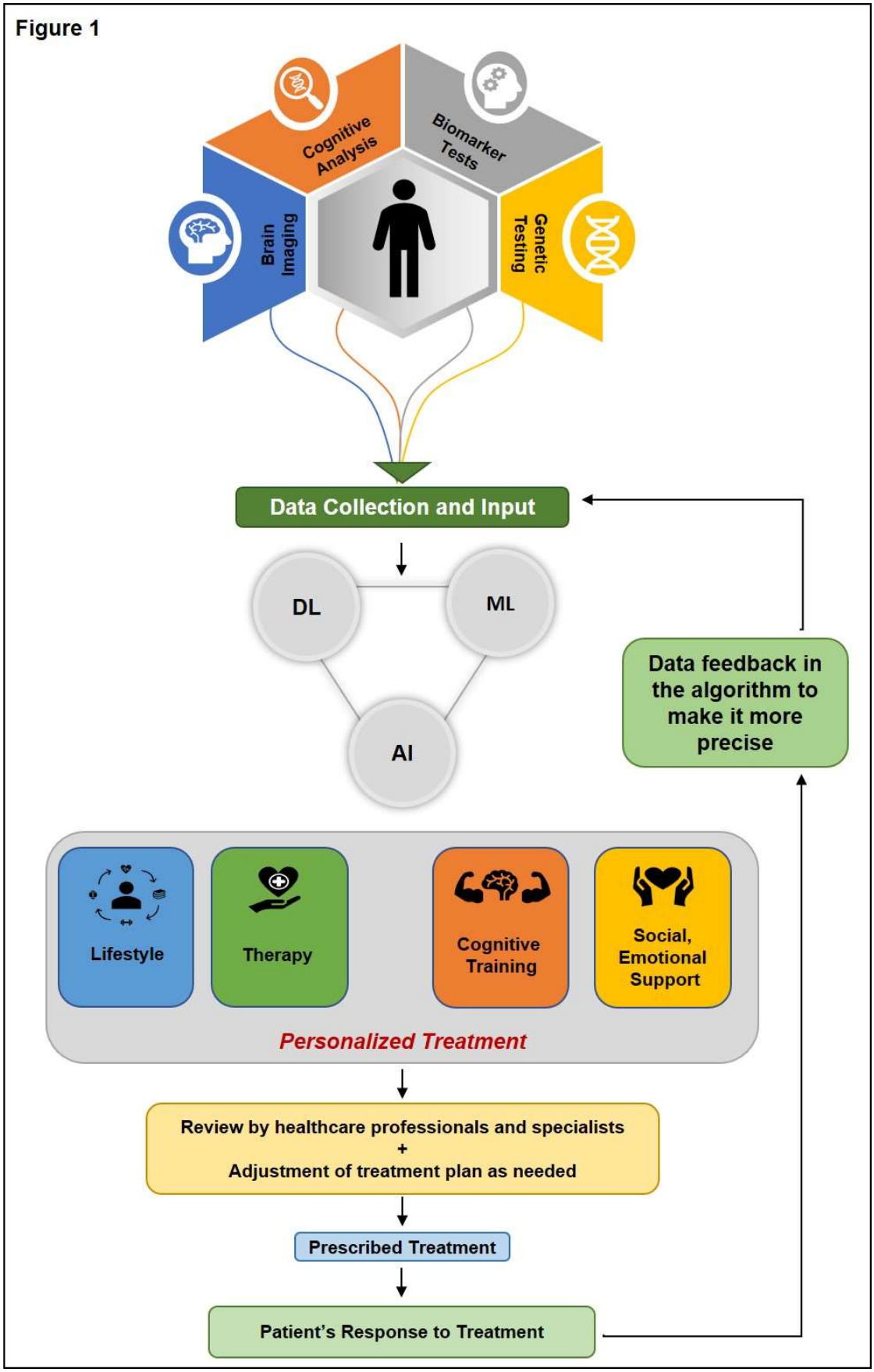
Cognitive Analysis
Biomarker Tests
Brain Imaging
Genetic Testing
Data Collection and Input
DL
ML
AI
Data feedback in the algorithm to make it more precise
Lifestyle
Therapy
Cognitive Training
Social, Emotional Support
Personalized Treatment
Review by healthcare professionals and specialists
+
Adjustment of treatment plan as needed
Prescribed Treatment
Patient's Response to Treatment

Figure 1. The diagram is divided into four components: Patient Data Collection, Artificial Intelligence/Machine Learning, Treatment Plan, and Ongoing Monitoring and adjustment to provide prescribed treatment. Based on this medication, patient's response is recorded and is again fed into the algorithms to enhance its accuracy.

Instruments used in clinical fields are equipped with updated versions of intelligent machine learning abilities which can be used to avoid errors in predictions and decision-making processes. It can also assist clinicians in their treatment process and provide valid and updated information in textbooks and journals [15]. Companies and industries like International Business Machines (IBM), Microsoft, Google, Apple, and DeepMind Technologies Ltd. are heavily investing in healthcare facilities related to artificial intelligence and are using several tools and techniques which integrate AI/ML algorithms for the improvement of the medical sector [15], [16].

CHAPTER 2

REVIEW OF LITERATURE

2.1 Moving towards big data in personalized healthcare

The term 'Big Data' was popularized in the year 1997, in the frame of reference to data visualization [17]. Its approachability is targeted to overcome exceptional challenges in the field of biomedicine and healthcare. Since then, biomedical and healthcare research has witnessed unapprehended changes in clinical as well as molecular knowledge discovery based on big data analysis[18]. For a very long era, big science and business intelligence were major sources of inspiration for big data technologies development [19]. In the year 2009, when people started using google for the prediction of flu-like illnesses just by scrutinizing the queries available, the prediction was highly accurate just like the prognostication of high-priced monitoring networks of the Centers for Disease Control and Prevention. Witnessing this, several analysts affirmed that big data can be a solution to modern healthcare problems [20]. In the last two decades, medical organizations and researchers have initiated to enfold retrieval of data from digitalized clinical records rather than using the traditional methods of data collection. This has resulted in real-time care, cost depreciation, data accuracy with minimal errors, and improved efficiency[21]. Big data analytics is being considered over traditional database management systems because regular systems were unable to administer vast information consisting of details concerning patient care like signs and symptoms, clinical records, the behavior of the patient, imaging data, medications, and insurance claiming data, etc. [22] **(figure 2a)**.

There are majorly four analytical models present that are used for data mining and analysis: Descriptive, Diagnostic, Predictive, and Prescriptive. In the descriptive model, exploitation of already present data is done to provide accurate outcomes. The diagnostic model helps to understand the reason for the occurrence of the disease. The predictive model analyzes the patterns of reasons by which the disease is caused and predicts future risks. The prescriptive model at the end provides some useful recommendations to control the occurrence of the disease [23], [24]. But the analysis of medical data from incalculable patients, determination of clusters and correlations, and the development of new models for disease prediction utilizing the machine learning approach or statistical techniques have been made possible using contemporary big data technologies [20]. The characteristics of big data analytics which have improved the processing of data and the

decision-making ability are the six V's, which represent: variability, variety, veracity, velocity, value, and volume. Volume, which is considered a foremost attribute refers to the amount of data that is being exploited for the analysis and retrieval of desired outcomes [25]. But in rare times, the data present in the databases or the clinical records do not contain enough information depth or the clinical data which is responsible for the expected results is missing.

Even the databases which are used for the evaluation of these data are expensive enough to be obtained by every individual and this sets a limitation in the field of research based on big data analytics [26]. The sources that comprise the surge of data are biosensors, digitizers, scanners, mobile phones, the internet, emails, and social networks [25], [27] **(figure 2b)**. The different types of information which are stored in the form of texts, videos, images, sounds, and geometrics or in the combination of any of these which is used for analysis and processing denotes a variety of data. The data can be both structured as well as unstructured which could have been retrieved from laboratory exams, sensors, simulation results, etc. as presented in some research studies [28], [29]. Velocity can be used to determine the rate of changes occurring, for example, the growth rate in age-related disorders where the number of patients increases with the increase in age. Veracity concerns with the data accuracy and data quality that is to be used in the processing of big data because data collected from clinical research is usually of good quality and high accuracy but at the same time, data from clinical practices may not generate valuable data and it may result in the wastage of time studying them and also distracts from their patient's care[30]. Another V symbolizes value which describes that data can provide valuable outcomes from the already stored information [20]. And finally, the last V means variability which tells how quickly and to what degree the format of your data changes. And how frequently does the nature or format of your data alter? Several algorithms have been proposed in the meantime to overcome the drawbacks like time complexity, and spatial complexity, and for the categorization of imbalanced data into 2 classes [31], [32].

Figure 2(a)

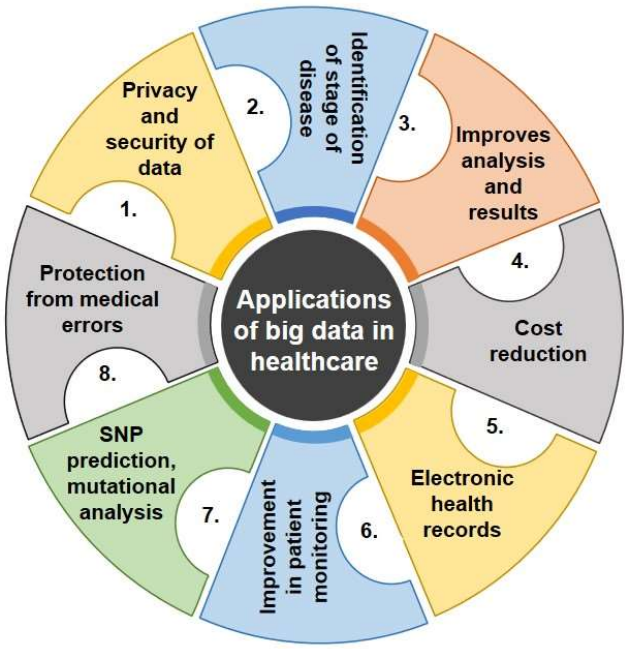

Figure 2(b)

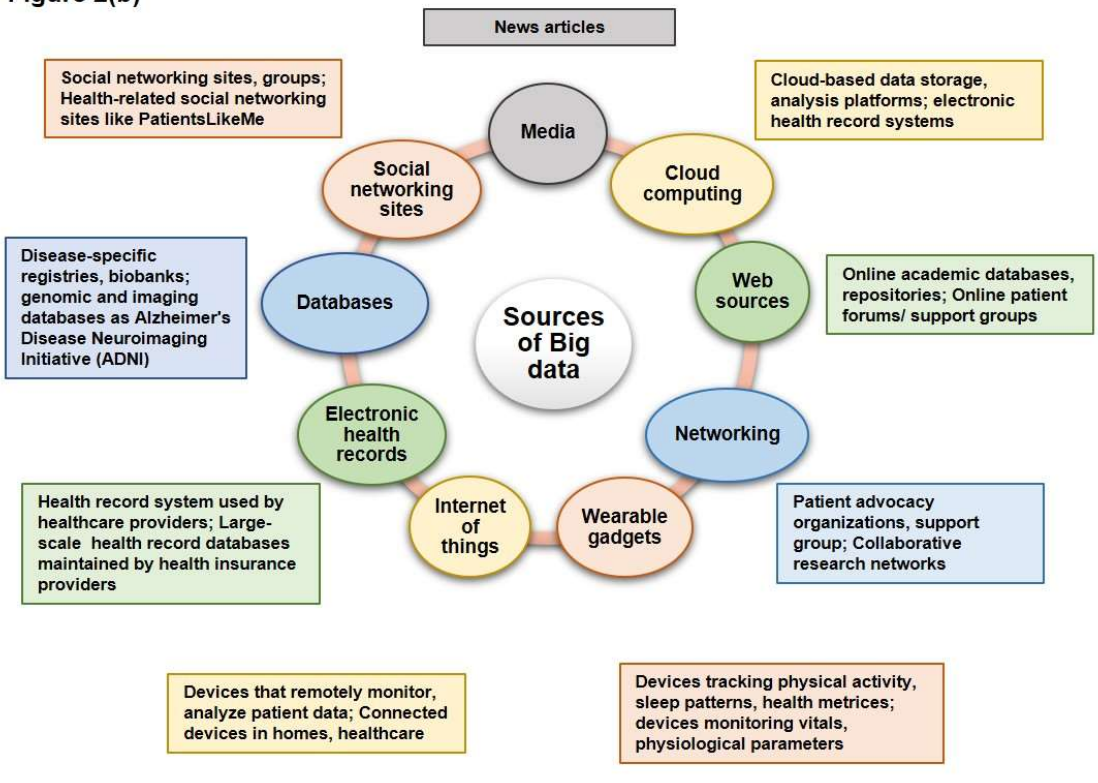

Figure 2: (a). There are multiple sources of big data like databases, electronic health records, cloud computing, wearable gadgets, etc. (b). Applications of big data in

healthcare include error free outcomes, better patient monitoring, cost reduction and enhanced privacy and security of confidential data of patients.

Table 1: Big data-based research in the healthcare industry

S.No.	Work done	Algorithm	Methodology	Sensor type	Merits	Demerits	References
1.	Early detection and preventative steps for limiting chikungunya virus	Fuzzy K-Nearest Neighbor algorithm	Smart health approach	Wearable Internet of Things devices	Easy to carry devices	Security execution is required	[33]
2.	To continuously monitor the patients in ICU	-	Real-time IoT-based device	Wearable sensors, bedside monitor	Permits the staff to invest time in decision-making with precise monitoring	The accuracy of the system is yet to be tested.	[30]
3.	Evaluate the condition of IoT sensors	Korotkoff method	Technology based care	Blood pressure sensors, inertial sensors	Reduction in power gap, meet the power requirement	Skin irritation, uncomfortable for the user	[34]
4.	Healthcare system for detection and limit of chikungunya virus	Fuzzy C-means algorithm	Analyzing social networks	GPS sensors, wearable sensors, climate detector sensor	Quick generation of alerts, high efficiency	High power consumption	[35]
5.	IoT-based system for personal healthcare devices	Fault-tolerant algorithm	Machine-to-machine-based IoT system	Wearable sensors, and personal healthcare devices	Effective supply of healthcare services	Usage of Byzantine fault-tolerant algorithm	[36]
6.	Storage and processing of big data for healthcare purposes	Stochastic gradient descent algorithm	Meta-fog redirection	IoT sensors, wearable sensors	Protection and prevention of big data from intruders	Accessible storage is not provided	[37]
7.	Framework proposal for the mining of healthcare data	-	Tensor-based data mining approach	-	Strong privacy	inconsistent	[38]
8.	Usage of big data analytics for remote monitoring of patients	Quantitative methods	Remote based real-time monitoring system	Body sensors	Early detection, cost reduction	Less accuracy	[39]
9.	ML-based simultaneous aided diagnosis for outpatients	Support vector machine	simultaneous aided diagnosis model	-	Provide better healthcare services	Data collected manually	[40]
10.	Usage of sensors and smartphones for recording signs of users	--	Mobile physiological sensor system	Wireless sensors	Cost-effective, convenient	Issue in consumption of energy	[41]
11.	Smartphone-based study of the ECG pattern of patients	--	Cloud-electrocardiogram system	--	Reducing diagnosis time. Enhancing medical service quality	Security issues	[42]
12.	Mobile application for self-healthcare services	--	Designing research methodology	--	Highly effective technology	Not applicable to remote healthcare services	[43]
13.	Manipulation of sensors and management and	Pattern matching algorithms	Architecture-level modifiability analysis	Biomedical devices	High communication efficiency	Short time monitoring	[44]

	monitoring of cardiac rhythms irregularity						
14.	Investigation of image-based diagnosis	--	Subject task analysis	--	Quick diagnosis	Processing time is unpredictable	[45]

2.2 The implication of artificial intelligence in precision medicine

In earlier days, if a large number of patients used to share similarities in the symptoms they were given generalized treatments irrespective of the cause of their illness. This could result in several allergic reactions. As technology advanced and the healthcare industry went drastic changes, several diagnostic tools were made for diagnosis and treatments [46]. The twenty-first century demands the evolution of the medical sector on the basis of disruptive technologies like bioinformatics, advanced biotechnology, cheap genome sequencing, patient electronic health records, and digital sensors used [47], [48] **(figure 3)**. Precision medicine, as defined by the National Institutes of Health, is "an arising strategy for the therapy and control of a disease that considers variances in genes, environment, and habits for each person." This method enables medical professionals and academics to make more accurate predictions about which disease-specific treatments and preventative measures will be effective in which populations. It needs powerful computers (supercomputers), algorithms that can acquire information on their own quickly (deep learning), and usually, a method that makes use of doctors' thinking skills on a new level [49]. Supercomputers' processing strength has turned into a battlefield as nations contend for supremacy through them. In cancer, dermatology, and cardiology, deep learning systems have demonstrated that they can make decisions at least as well as human doctors [49], [50].

The significance of fusing the expertise of doctors with such systems must be recognized. Competitors developed algorithmic methods for identifying metastatic breast cancer in whole slide pictures of sentinel lymph node biopsies for the International Symposium on Biomedical Imaging's grand challenge. The success percentage of the winning program was 92.5%. The success rate was 96.6% when a pathologist separately examined the same pictures. The pathologist's success rate rose to 99.5% when the deep learning system's forecasts and diagnoses were combined, which represents an 85% decrease in the error rate caused by humans [51]. Stakeholders must now put in place a plan that supports

evidence production, data sharing and integration into healthcare, financial benefits, regulations, payment, and user involvement in order to completely realize the incorporation of precision medicine [52], [53]. Several arguments have taken place discussing whether the involvement of AI/ML-based precision medicine is leading to the loss of the human touch. But this isn't true. AI-based precision medicine requires a lot of tools, codes, patient monitoring, data analysis, and many more which is not possible without the involvement of doctors and researchers.

Figure 3

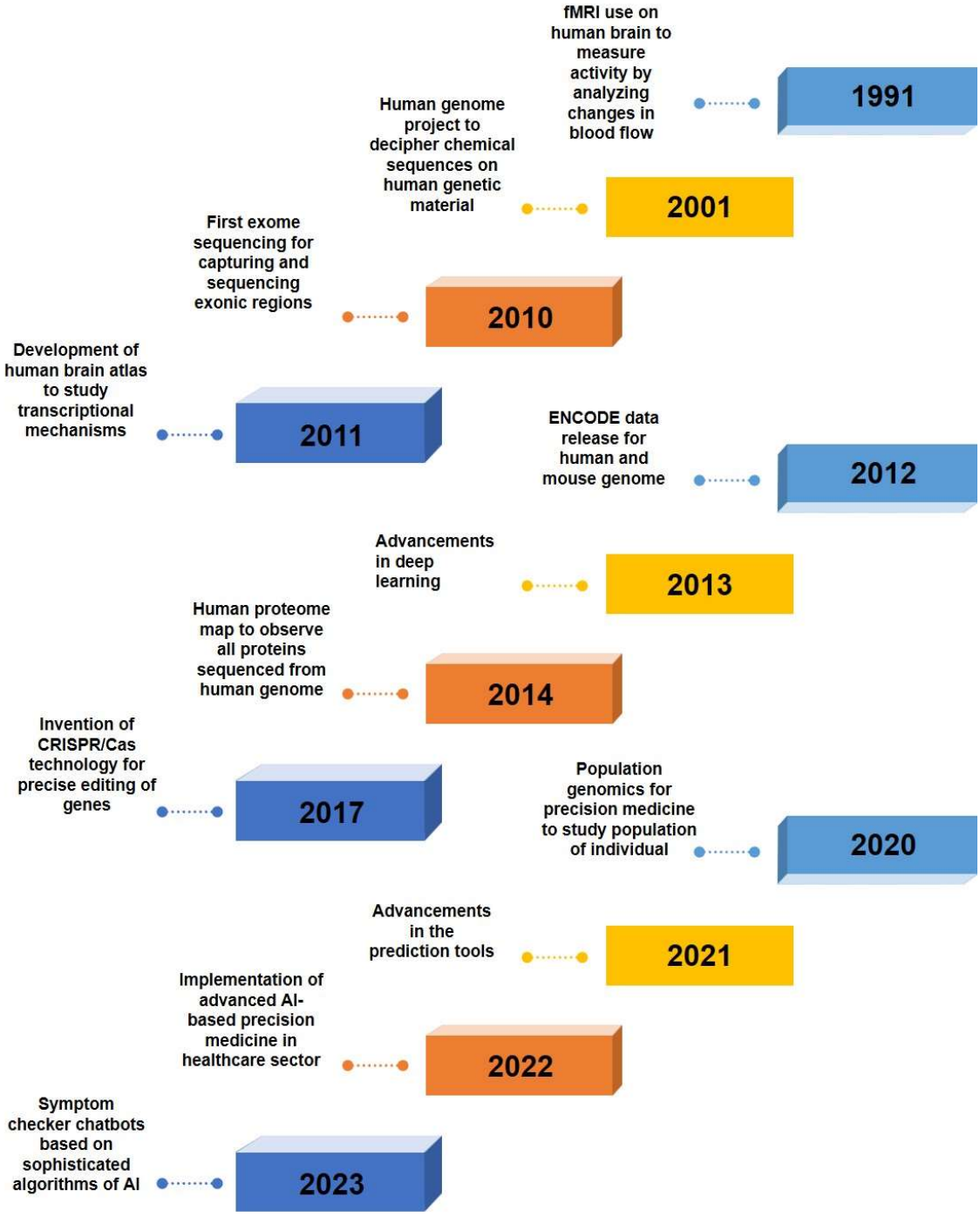

Figure 3. Timeline of the advancements in the technology based on precision medicine and healthcare industry.

2.3 Role of artificial intelligence in integrating clinical and omics data to revalorize personalized medicine

In the 21st century, as the biomedical sector witnesses the big data era, Scientific journals, and funding agencies are encouraged and motivated to produce large volumes and a variety of data so that the accessibility to data gets much easier and quicker. To make data publically available, multiple databases and repositories were entrenched [54]–[57]. Standardized frameworks for data management approaches, cross-platform interoperability, and common processes for data sharing and data analysis, on the other hand, trailed behind an exponentially growing faster data output, hurting model implementation and insight generation. Researchers' abilities to acquire, combine, and analyze frequently noisy, complex, and high-dimensional data continue to face significant challenges as a result of the isolation of data from EHRs and the multi-omics[58], [59]. But several steps are being taken to avoid false or negative data like an increase in the sample size, reduction in the heterogenicity of the studied population, reduction in the heterogenicity of the measurement methods, reduction in feature space to lower data's background noise, performing 2-3 studies to validate the results, and performing experiments both in-vitro and in-vivo to validate the outcomes [58]. UK Biobanks and Million Veterans Programs are standard examples of biobanks containing data that are used to avoid the issue of sample size. The data present here is in the form of biochemical, demographic, or anthropometric data [60], [61]. Despite the fact that omics research is restricted by large analytical heterogeneity and design constraints based on experiments, leading to a low signal-to-noise rate, data merging of multi-omics information remains a key provocation in precision medication [62]. Genetic and genomic study along with the environmental conditions of the person proves to have a notable part in the precision medicine revolution launching, and this also shows that the majority of the precision medicine data is taken from genetic and genomic studies [63], [64]. This approach of precision medicine has saved a lot of money and time which was earlier used in genetic studies and wet laboratory experiments [65]. Numerous molecular attributes, including genes, proteins, and the relationships between those elements, are responsible for biological processes. By grouping these molecules based on their structural or functional similarity, omics refers to their thorough classification and measurement [66]. Integrating multiple levels of omics data allows researchers to model biochemical systems and better comprehend how distinct biological systems communicate at the molecular level [66], [67]. For instance, an increase in a gene's messenger RNA expression may not result in

an increase in its protein expression, and an increase in its protein expression may not result in an increase in its activity. Additionally, a single assessment of a single omics region may not disclose time- or dynamically-dependent processes. In order to adequately describe biochemical signatures indicative of the phenotype at the exact time the sample is collected, these difficulties call for deliberate mixtures of omics data [68].

Omics integration is an effective strategy that can connect even tiny data sets across orthogonal biochemical regions, amplifying physiologically important signals in the process [69]. Common methods for doing this include empirical connections (correlation), functional settings like pathways, data drawn from a single study design, and meta-analyses that combine the findings of several studies [70]. However, only DL provides the ability to integrate omics data both unstructured and controlled. For example, to create a multi-omics library and predictive model, researchers investigating the growth kinetics of Escherichia coli used a recurrent neural network (RNN) as part of an ML-based data integration approach [71], [72]. Integration of omics data is important for neurodegenerative diseases (NDDs) like AD as well as PD because the given diseases typically have a multifactorial origin, diverse clinical manifestations, and mixed pathologies [73]. The two types of multi-omics data are multi-feature data and multi-relational data, respectively, depending on the number of features and sample sets that are engaged in the same event or system. However, some data design variants are feasible, such as multi-class data that measure different sample sets employing a similar set of characteristics and at last, tensor data that measure the same sample collection of objects using the identical set of features in various scenarios [74]. A multi-omics-based statistical analysis of data in Neurodegenerative Disorders may be done to identify sub-phenotypes using clustering methods or to look for potential indicators and druggable targets. Understanding the relationships between an array of features may also be important in comprehending the pathogenic mechanisms underneath various illness phenotypes, every one of which is identified by a unique combination of biomarkers to create a phenotypic subtype with its own set of appropriate personalized treatments [75].

2.4 Applications of artificial intelligence in precision medicine

Currently, a variety of uses of precision medicine support health treatment at various stages of life. To determine the likelihood of transmitting genetic diseases to future generations, genetic screening may be used before conception [76]. Non-invasive prenatal testing can be used to check for the developmental defects caused by trisomy 13, 18, and 21 in the fetus between eight and twelve weeks of gestation [77]. Whole-genome

sequencing of the fetus has even been done. Sequencing has been used at delivery to quickly identify a variety of serious diseases for which there may be a remedy, lowering morbidity and mortality [78]. Later in life, these methods can be used to identify a number of illnesses, most notably cancer by detecting DNA from moving tumor cells [79]. Here we are going to discuss some applications of precision medicine which involve the integration of artificial intelligence and machine learning [80] **(figure 4)**.

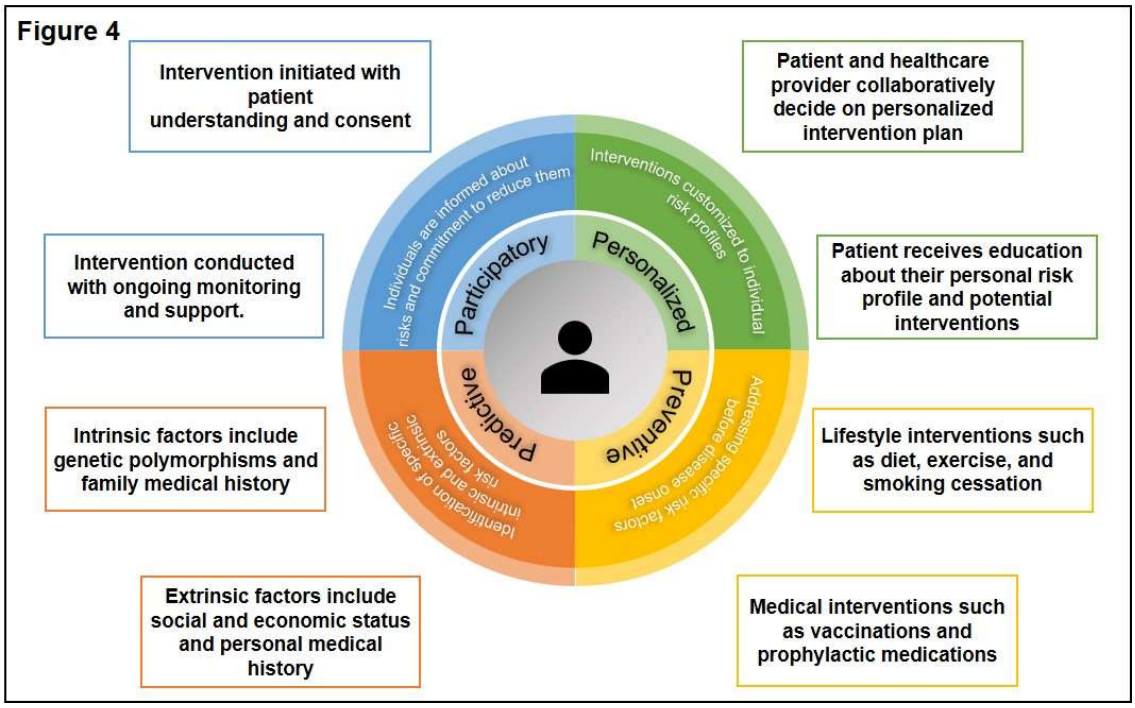

Figure 4: P4 Medicine consists of four quadrants: predictive, preventive, personalized, and participatory, with precision medicine as the overarching concept. The predictive component uses data and analytics to anticipate risk, the preventive component focuses on measures to prevent or delay disease onset, the personalized component customizes healthcare based on an individual's unique characteristics, and the participatory component promotes patient involvement in their own care. This holistic and personalized approach aims to achieve optimal outcomes.

2.4.1 Data hampering and privacy

Big data comprises a huge amount of unstructured or unstructured data which is collected from several sources like smart devices, gadgets, web server logs, phone records, and many more. The gathered information is in the form of videos, images, audio, written records, etc. that is to be linked with several sources for analysis, cleaning, and transformation [81]. The entire data is extremely vulnerable and need strong security and privacy in order to avoid trust issues [82], [83]. If encryption is not provided, the patient's crucial data may be subjected to the attacks of cybercrime and can be severely misused [84]. Differential privacy, a model for data protection received Godel Prize in the year 2017 and was introduced to overcome the shortcoming and limitations of already existing privacy models [85]. This model helps in the identification of individuals from datasets with the help of auxiliary information. The differential privacy model promises to provide valid information when studying a dataset, even if the individual is involved in the study or not [86]. Efforts are invested to fill the gap between the differential privacy model and the k-anonymity model to provide accurate results and maximize privacy [87]. Earlier, features of the differential privacy model were a major limitation in precision medicine because the data dimensionality was unable to match up with the model's privacy mechanism [88]. But now the implementation of differential privacy models for GWAS is helping users to analyze the databases. This initiative has resulted in providing accurate outcomes with good quality [89]–[91].

2.4.2 Medical devices assistant

Regardless of the extensive research that has helped to identify a wide range of illnesses, it has been difficult to translate that research into clinical practice. Machine learning or ML can fill the gap between precise medical evaluation and the translation of pertinent clinical data [19]. Researchers are using AI-enabled smart devices, such as smartwatches, smartphones, and tablets, for the identification and categorization of arrhythmias [92]. They are also using these devices for the detection of aspiration pneumonia in which the patients suffer from swallowing difficulties which were caused by stroke or dementia [93] and to increase medication compliance in anticoagulation therapy receiving patients [94]. Specialists are using iPads with AI capabilities to develop differential diagnostic and treatment plans for uncommon epilepsy disorders [95], as well as smart devices comprising of wrist-based sensors, for the identification of seizure like activity [96]. The tracking of cardiovascular illnesses, pulmonary diseases, anemia, and sleep apnea [97] is attainable by the addition of photoplethysmography devices in gadgets. Additionally,

Parkinson's disease signs like tremors and altered gait, posture, and speaking patterns could be detected and measured by wearable devices [98]. Personal monitoring devices offer the chance to direct behavioral adjustments [99], but the accuracy of the data they gather can vary [100]. . In addition, one-third of all smart device owners in the US gave them up within six months of purchase, indicating the potential of the technology to influence long-term attitudinal change [101]. To determine the best strategies for maximizing wearables' efficacy in the maintenance and promotion of health, more study is required [102].

Table 2: Artificial intelligence-based technologies and their applications in concern to neurological disorders

S.no.	AI/ML-based medical devices	Description	Applications	References
1.	smartphones	• An iPhone-based application • Medical history based 8-minutes based on medical history of patients and their families	Helpful in the classification and stratification of tremors in patients.	[103]
2.	Smartwatches	• Smart device system-based application used to track movements. • 2 watches are worn during the neurological exams. • The data collected from smartwatches are assessed based on amplitude and frequency to distinguish different types of movement disorders.	Capturing and detecting high-resolution tremors, and movement disorders and the information is transferred to the user's mobile phone.	[103]
3.	Smart patches	• A wearable smart device • embedded electronics attached to the skin of the patient • monitoring physiological symptoms like a pulse	Deliver bio-electrical signals for Sensation restoration and skin nerve regeneration.	[104]
4.	Tablet-based devices	• iPad-device-based assessment • test of Archeamedian spiral drawing to test neurological patients • data collected is used for feature detection, pattern deviation, direction inversion.	The sum of the data collected from the test is used for the prediction and diagnosis of movement neurological disorders.	[103]
5.	Epifinder – AI-integrated application	• iPad-based pattern recognition applications • better medication and therapeutic approaches to neurological disorders like epilepsy. • Neurologists use the data history saved in application for treatment of patient.	Comparision of knowledge representation and symptoms	[105]
6.	Photoplethysmography	• technique for detection of heartbeat • analyzing change in skin color and absorption of light. • Change in the light intensity is detected by PPG via the reflection from the tissue.	Cell-phone-based PPG records the heartbeat and variability which is considered for paramedics.	[106]

7.	Electrocardiography monitors	• A phone-based electrocardiogram • records heartbeats in form of electrical signals • check condition of the heart.	Used to detect atrial flutter, ventricular premature beats, atrial fibrillation, etc.	[106]
8.	Neuroimaging	• Quantitative and qualitative techniques • to study the activity and structure of the brain.	Machine learning algorithms are used in diagnosis and treatment decisions.	[107]
9.	Self-organizing map	• ML-based approach • studying phenotypic difference between concussion patients based on balance, vestibular diagnostics results	Supports the identification of different phenotypic aspects of concussion.	[108]
10.	Imaging-based time since stroke onset	• ML-based approach to extract imaging sequences and hidden representations. • Demonstration of improved classification by integrating deep features is also done.	Analysis of magnetic resonance imaging is advanced for stroke treatment by getting closer to operational decision support tools.	[107]
11.	Augmented EEG-learning	• Several ML-based models • SVM and linear discriminant analysis are used for the prediction of seizures.	Neurophysiological and pathological factors are used for the prediction of the results of epilepsy surgery.	[109]
12.	Imaging services	• CT scan and MRI • Imaging services model working on clinical clues and severity of the patient.	Easier and quicker imaging evaluation is performed.	[110]
13.	Electrodermal activity sensor	• wrist-annotated wearable sensor uses a video-based EEG monitor. • Each epoch is classified into 3 classifiers based on probability estimations • classifier 1 has 19 features, classifier 2 has 46 features and classifier 3 has 25 features.	Detection of precision-based seizure activity.	[111]–[113]
14.	EnsoSleep	• AI-based algorithm for classification of sleep disorders.	It is a sleep-scoring solution used in sleep centers and hospitals for testing sleep volume and management of patient care.	[114]
15.	HIPAA-compliant applications	• application is designed for improved medication, and therapeutics identification of patients, dosing, and treatment.	Patients suffering from Alzheimer's, stroke, and epilepsy uses this application to elevate medication adherence.	[115]

2.4.3 SNPs prediction and mutational analysis

The majority of the human genome, or about 1/1,000th of the typical human genome, is made up of single nucleotide polymorphisms (SNP) [116], [117]. SNPs are traditionally considered to be biallelic since two of four familiar nucleotides are present in the given location and the most uncommon nucleotide is present in over one percent of the entire population [118], [119]. The current studies are heavily focused on the distribution and purpose of SNPs. Reviews for comprehension that how SNPs influence the structure of the protein, using SNPs in genetic research, and locating functional variations in contender genes are all accessible [120]. Bioinformatics, and artificial intelligence in the

form of web resources and software tools has played a major role in SNP characterization. Accessibility to visualization and functional annotations are provided in the NCBI databases like Ensembl, dbSNP, and OMIM. SNPper and Goldenpath are some major resources for the analysis and visualization of SNP [121]. Finding a link between complex illnesses and genetic risk factors is a current scientific concern. The method that links variations with traits is known as a genome-wide association study (GWAS), and it has been used the most frequently so far. Analyzing single nucleotide polymorphism is required in this. Understanding the illness development linked to SNPs is one of the main significances of SNP research. The main challenge for the experts is to create a method that can efficiently mine the millions of functional SNPs present in a database [122]. SNPs are more commonly found in the 5′UTR, 3′UTR, and introns of the genome, which are non-coding sections [123], [124]. While SNPs in the 3′UTR typically have an impact on gene translation, those in the 5′UTR are engaged in transcriptional activity. The handling of mRNA will be impacted by the polymorphism found in intronic areas [125]. Because non-synonymous SNPs in the coding sections of the genome are bringing mutations at the amino acid level, which can be detrimental to the structure or function of the protein, researchers are paying close attention to coding SNPs in particular. Proteins are harmed by a variety of complicated illnesses because of their collusion in protein folding, protein-protein interactions, and protein stability alterations [126]. Latest advancements which include summing up temporal and spatial prospects of gene or protein expression data help in the understanding of regulatory activities [127]–[130]. Recently, the SNPs that are associated with phenotypes and genetic risk factors are included in risk modeling. This approach has benefitted in the accurate prediction of disease [131]. In order to assign genetic risk for the individual outcome, polygenic risk scoring plays a major role [132]. Studies have shown that bigger datasets have elevated the chance to predict smaller effect sizes of SNPs and such SNPs if get indulged in polygenic risk models lead to accurate predictions of the illness [133].

CHAPTER 3

3.1 The emergence of artificial intelligence-based drug discovery and development in neurological diseases

Intensified loads of neurodegenerative disorders are considered a major reason for speedy aging in the population and at a later date resulting in extravagant expenses in treatment, assistance in the patient's care, hospitalization, and even higher mortality rates [14]. The criteria that differentiate NDDs from neurological disorders is the neuronal dysfunction or neuronal loss in specified regions of the spinal cord or brain. Several physical and mental impairments are witnessed in people suffering from age-related neurological disorders also known as NDDs like speech issues, cognitive decline, forgetfulness, anxiety, continuous changes in mood, and even decline in other organs of the body is also seen [14]. These disorders can be a result of electrical, structural, or biochemical abnormalities in the neurons, spinal cord, and brain. AD, being the most common disease among the list of neurodegenerative disorders is predicted to affect approximately 7 million Europeans by 2040. The complete cure for the disease is still unknown, still, the diagnosis, treatment, and care comprise around 130 billion euros per year [14]. Facing such challenges, improvement in the healthcare sector is a need of society. Applications of artificial intelligence have started to contribute to the upgradation of neurological healthcare as the algorithms help to study the risk factors like genetics, epigenetics, and environmental factors which are responsible for the diseased body, also at the same time reducing manual labor [46]. To make research easy, several databases have been made which provide most of the knowledge regarding the diseases specifically the most studied NDDs like Alzheimer's Disease and Parkinson's disease [14], [134]. In the current era, computational work has significantly contributed to the field of research and healthcare, and a large portion of scientific research can rely on the analysis, predictions, and results produced by high-technology-based computational tools and algorithms [135].

Precision medicine, an artificial intelligence-based medication can pave the path to studying neurological disorders on a deeper level and provide a much more comprehensive description as compared to traditional medication. The ultimate goal of the researchers should be to take practical applications from bench to bedside, analysis, and discriminate patients based on risk factors, find the stage or level of disease, provide tailored medications, and set up strategies to provide the right treatment to the right person

at the right time [134]. AI has proven to be richly beneficial by giving a promising future in upgrading neurological care procedures, analysis of brain tumors, monitoring of symptoms and progression, and outcome predictions. In the past few years, several instruments, databases, and gadgets are invented which in turn increased efficiency in diagnostics, prognosis, functionality, and treatment of patients [136]. AI-based precision medicine is not only a ground breaking miracle in the field of science but it has also become a battleground through which several countries are demonstrating their powers [46]. Improvements in the techniques which help in understanding and parsing out the intermediate pathways that expedite neuronal loss and dysfunction is a major step of integration of precision medicine in the treatment of neurological disorders.

Table 3: Machine learning algorithms with their mechanism of action, applications in neurology

S.NO.	ALGORITHMS	MECHANISM OF ACTION	APPLICATIONS	REFERENCES
1.	Random forest	Classification of new objects from attributes	Prediction, classification	[137]
2.	Regression	Finding the correlations between dependent and independent variables	interpolation	[138]
3.	Decision tree	Dividing the population into more than 2 homogenous sets	classification	[139]
4.	K means	Classification of datasets in a specific number of clusters	clustering	[140]
5.	k-Nearest neighbors algorithm	Classification of new cases based on majority voting and storing all cases for prediction	Interpolation, clustering	[141]
6.	Logistic regression	Using logic function finding the event's occurrence probability	classification	[142]
7.	Naïve Bayes	Prediction is done by assuming a feature that is different from other features	classification	[143]
8.	Support vector machine	Plotting data items as a point in n-dimensional space	classification	[144]
9.	gradient boosting algorithm	Predictions did by assembling the learning algorithms	prediction	[145]
10.	Double Ratchet Algorithm	Conversion of high dimensional data into low dimensional data	Interpolation, classification	[146]

In recent years there has been a blood of scientific articles based on artificial intelligence in healthcare, big data in healthcare, AI-based precision medicine, and implementation of precision medicine in chronic diseases like cancer, diabetes, and neurological disorders. Here, we have made graphs of the number of papers got published on various neurological

disorders like Alzheimer's disease (A.D.), Parkinson's disease (P.D.), Huntington's disease (H.D.), and Neuropsychiatric disorders. These graphs were made using the keywords precision medicine, AI/ML, and neurological disorders (A.D., P.D., H.D., and neuropsychiatric disorders, etc) on literature databases like PubMed and Google Scholar **(figure 5)**.

Figure 5(a)

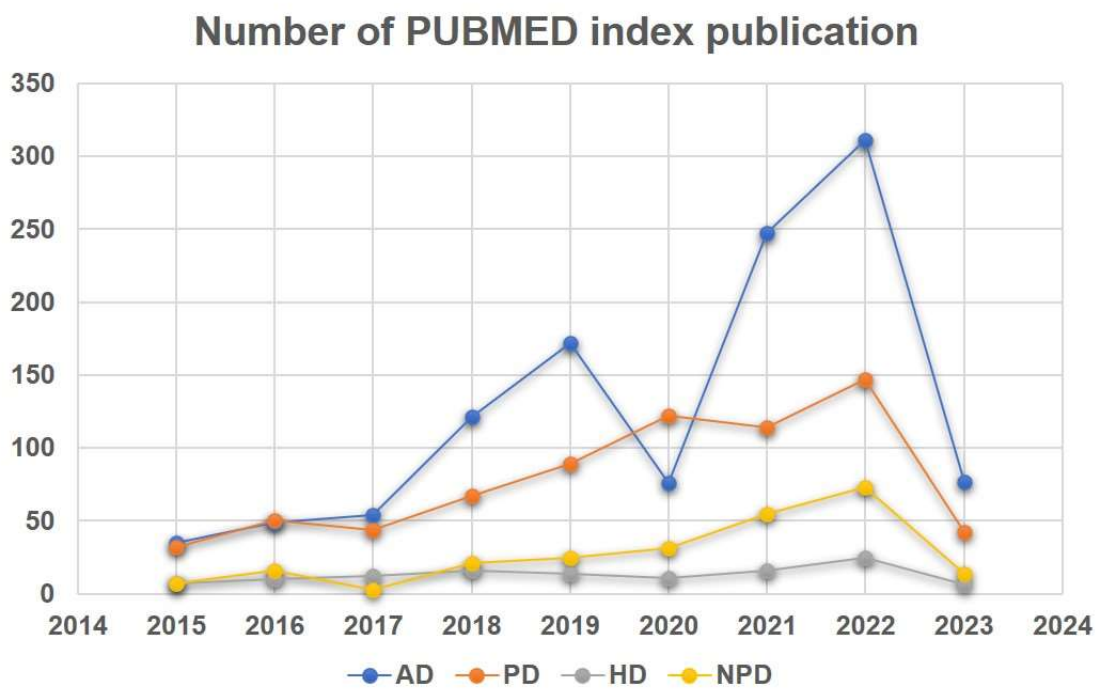

Figure 5(b)

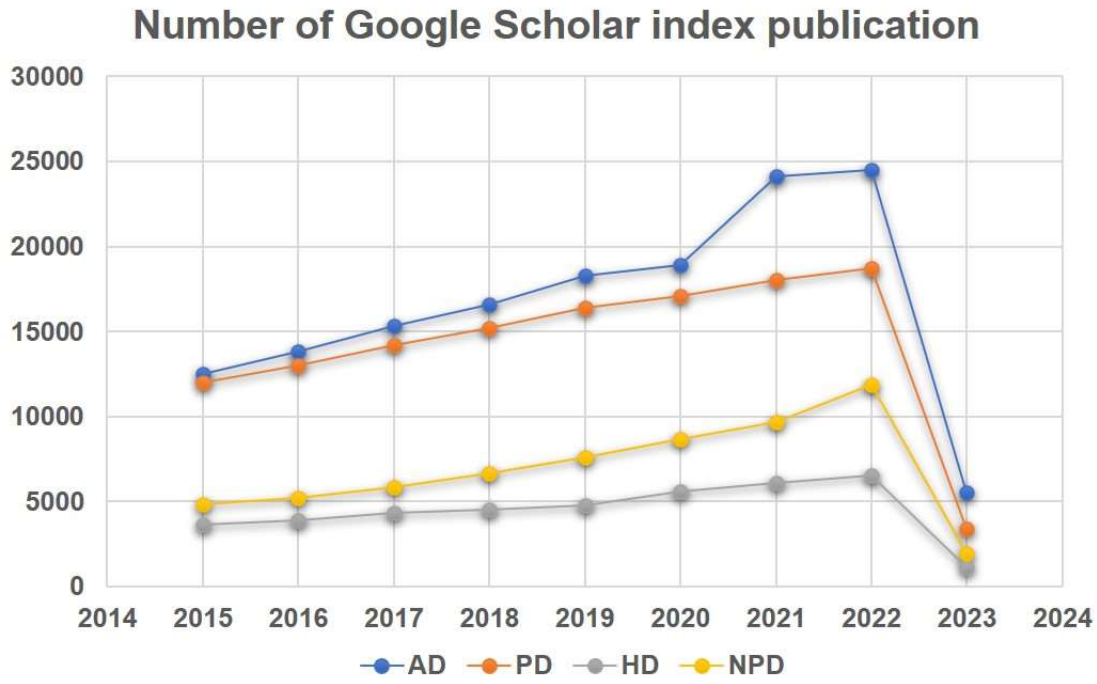

Figure 5: (a) This graph displays the number of PUBMED indexed publications on the topics of Precision medicine and AD, PD, HD, and NPD. Each topic is represented by a different coloured line, and the y-axis shows the number

of publications. (b) shows the graph of the number of publications present in Google Scholar database, related to neurological disorders likeAD, PD, HD, NPD. The data present in both the graphs is from the year 2014 to March 2023.

In the thesis, latest literature is collected and drug discovery and development, an important application of AI/ML has been taken into consideration for the research work to defend the thesis work. Machine learning algorithms have been employed to retrieve, sort, and indentify multipurpose drugs for the treatment of neurological disease, AD. AD has been a deadly disease for which accurate medications are not yet present. As the urbanization is increasing, canges in lifestyles, environment and even in the genetic makeup are occurring, which is resulting in the mutation in genes and proteins and later leading to neurological conditions in a person. Eventually, machine learning tools were used to validate the results. The entire methodology, results and discussion have been done in the upcoming parts of the thesis.

3.2 ALZHEIMER'S DISEASE

Alzheimer's disease (AD) is an untreatable neurological condition that gradually erodes cognitive abilities until dementia emerges. AD is distinguished from other forms of dementia by the aggregation of intracellular neurofibrillary tangles and extracellular amyloid β plaques. Medications such as Donegal, galantamine, rivastigmine, and memantine are presently available to treat AD. However, clinicians work against the clock to identify drugs that can slow or stop disease progression. Amyloid-beta-targeting medicines, DNMT 1 inhibiting medicines have been developed for many decades, but promising medications have failed to exhibit therapeutic efficacy in the clinical trial phase III research. Even the positive Aducanumab findings from Biogen are still unclear, and further study is needed to evaluate the drug's long-term effectiveness. Consequently, scientists concentrate their efforts on tau-targeting therapies, as tau seems to be more directly linked to cognitive decline than amyloid. Along with this protein, other genes and proteins like DNMT 1 and BACE 1 and many more are being studied. Some of the apoptotic signaling pathways that affect AD are involved in the interaction between trophic factors with signaling pathways, including PI3K/AKT, JNK, MAPK, and mTOR. Immunotherapies make up most drugs in clinical trials, even though they are still relatively young. Objective cognitive impairments are the basis for the clinical diagnosis of AD (which are, typically, prominent memory impairments) [147]. Atypical AD presentations with deficits in non-amnesic areas have been seen in some instances. But AD has many clinical traits in common with other neurodegenerative dementias, such as Lewy body dementia, frontotemporal disorders, and vascular dementia, rendering early and alternative identification proves to be challenging in the initial stages of the disease [148], [149]. Precision medicine in AD and other neurological diseases, aims to offer a therapeutic or preventive strategy that is personalized to the discovered molecular pattern for vulnerability and disease progression. This is done by identifying risk factors, and its underlying pathophysiologic mechanism, and administering a therapeutic or preventive intervention. Drug discovery and development has been an advanced application of machine learning which is being utilized for predicting the appropriate drugs for the treatment of AD, without using any animal model for experimentation purpose. This also helps cost and time reduction along with saving the animals from getting exploited. Determining the risk profile, molecular mechanisms, and analysis of disease pathology profile are some of the other applications of precision medicine in A.D [150], [151].

How is technology getting advanced in other medical applications? For diagnostic, prognosis, or progression modeling, recent works on the application of AI in AD study use language and speech data gathered in various methods and computational speech processing. This field of technology includes techniques for listening to, identifying, and comprehending spoken language. It suggests that at least some of the AD detection procedures might be automatic. The core of this study program has been machine learning techniques. The development of predictive models that are "learned" directly from data is the focus of the AI discipline known as machine learning, where the student enhances its performance through "experience." Positive findings from research on automated speech and language processing using AI and machine learning techniques have sparked a growing interest. Signal processing, computational linguistics, and human-robot interactions are some approaches that have been studied with respect to A.D. diagnosis and treatment [152]–[154]. The work here is based on the drug discovery and development application for the inhibition of DNMT 1 to stop the progression of AD, and make more and more drugs easily feasible to the patients which cause less side-effects to their health.

3.3 DNA METHYLTRANSFERASE 1

A big protein having a C-terminal catalytic domain and an N-terminal regulatory region, DNMT1 has about 1600 amino acids. CXXC zinc finger domain, Proliferating cell nuclear (PCNA) binding domain (PBD), two bromo-adjacent homologies domains replication foci targeting sequence (RFTS) domain are only a few of the functionally significant domains found in the regulatory N-terminal region [155]. 7 lysyl-glycyl dipeptide repeats serve as the link between the C-terminal as well as N-terminal domains. The multiple domains that collectively comprise the N-terminal region each have a unique role in how DNMT1 operates as a whole [156]. While DNMT1 interacts with replication machinery via the PBD domain, the RFTS domain improves DNMT1 recruitment to replication foci, whereas the CXXR zinc finger domain aids in DNMT1 binding to unmethylated DNA [157]. It is not unexpected that DNMT1 experiences conformational modifications with the goal to communicate with domain-specific proteins given the distinct functional functions that these domains perform. Furthermore, it is not unusual for proteins to undergo specific conformational changes during structural interactions that lead to functional activity. Recruiting DNMT1 to heterochromatin replication foci and DNMT1 autoinhibition are the two primary functions of the DNMT1 RFTS domain [158]. To stop the enzyme, the domain binds to the binding pocket of the catalytic domain. It has been discovered that the degree of inhibition depends on the intensity of the connection between RFTS and the catalytic domain: the more intense the interaction, the decreased DNA methylation activity [157]. The RFTS domain must be moved in order for the DNMT to function. It's also crucial to remember that a helical connector between the CXXR domain, as well as the BAH1 domain of the DNMT1, regulates the autoinhibition caused by RFTS. Mutations in the RFTS domain impact DNA methylation since it is an autoinhibitory domain, which also results in aberrant functioning [155]. DNMT1 if mutated is find responsible for the cause of AD. It is seen to perform the regulatory function in central nervous system development but if mutated leads to the diposition of abnormal proteins in the brain which in turn cause memory issues and can be non-treatable and fatal.

3.4 DRUG DISCOVERY AND DEVELOPMENT

Healthcare systems always require new medications to address unmet medical requirements in a range of medical specialties, and pharmaceutical companies are largely focused on developing new treatments through the difficult procedures of research and development for drugs. The discovery process involves a number of processes, including as target selection and validation, hit detection, lead generation and optimization, and target identification for possible future development. On the other hand, the process of development includes refining the production of chemicals and formulations, performing animal toxicological research, running clinical trials, and finally obtaining regulatory approval. [159]. Both of these procedures take a lot of time and money, and the business is now under pressure due to incredibly strict regulatory demands, environmental concerns, and decreased profits as a result of patent expirations.

The goal of a preclinical drug discovery program is to generate a small number of pharmaceutical candidate molecules with sufficient evidence of biological activity at a disease-relevant target, sufficient security especially sufficient drug-like properties to be evaluated in people. A great deal of discovery programs tries to produce a variety of prospective molecules since numerous compounds do not proceed through the entire process owing to problems with dependability, motion, efficacy, privacy, or other characteristics.[160][161]. The enormous cost of failure implies that researchers proposing novel targets or compounds for research should carefully consider the traits that accompany effective research programs, even as funding in AD treatments continues[159]. It is more preferable for failures to occur sooner rather than later in development in a situation where disappointment is the norm. Additionally, it is ideal if every investigation in the development program produces data that offers a strong case for ending, continuing, or specifically changing the compound development program. Data that are challenging to interpret scientifically can result in additional studies and delayed decisions without the possibility of better studies in the future [161]. For this method multiple machine learning algorithms and computational tools are required which helps in time saving and are always cost effective.

CHAPTER 4

METHODOLOGY

Initialy macine learning algorithms like support vector machine, and random forest were trained to retrive the entire inhibtors that were established against DNMT1. These were devided into active, inactive and intermediate classes based on the multiple parameters like molecular weight, structure, atoms present , etc. Then the shortlisted drugs were validated on the basis of binding affinity and RMSD value using molecular docking technique. The methodology for entire data collection till docking and results retrieval is given below.

A. Collection of data

To repurpose the drugs against DNA methyltransferase 1, a list of anti-cancerous, anti-diabetic, and anti-hypertensive Drugs was retrieved from ChEMBL. ChEMBL is a chemical database managed by the European Bioinformatics Institute and is responsible for the information regarding the bioactive molecules specifying their drug-producing properties.

The SDF structure of all the listed drugs was retrieved from PubChem using https://pubchem.ncbi.nlm.nih.gov/. PubChem is a database handled by the National Center For Biotechnology Information and is responsible for the activities of chemical molecules against biological assays.

The blood-brain barrier (BBB) permeability of these drugs was checked by screening all the drugs through a BBB predictor using https://cbligand.org/BBB/predictor.php. BBB permeability protects the brain from inhibiting harmful molecules to pass by because they can be severely toxic to neurons.

Protein Data Bank or PDB, a file format used to specify the 3-D structure of molecules was used to download the structures of wild-type DNMT1 and its mutant using the link https://www.rcsb.org/. The basic information of literature was derived from literature databases like PubMed and google scholar which consists of all the studies and research work in form of published papers. It was done using https://pubmed.ncbi.nlm.nih.gov/, and https://scholar.google.com/.

B. Preparation of receptor and ligand

AutoDock vina was used to prepare the DNMT1 molecule structure. Water molecules were removed from the structure and polar hydrogen atoms, and Kollman charges were introduced in the structure. The prepared structure of the receptor was saved in the PDBQT format.

Similarly, ligands in the form of drugs that were downloaded in the SDF format were initially converted to PDB format and eventually to the PDBQT format. This entire process was carried out in the docking tool. The grid map for DNMT1 and its mutant was prepared whose dimensions were 80x80x80 and whose center was - 9.167, -5.056, 12.000

C. Molecular docking

Molecular docking was carried out using the AutoDock Vina tool and the top ten drugs were shortlisted based on their free binding energy and Root Mean Square Deviation (RMSD) value. <1 angstrom was the value for RMSD and -9kcal/mol was standard free binding energy.

D. Enlisting top drugs and studying the best drug

A list of best drugs was made and the best binding ligand was taken for docking again and the details were retrieved. A table was made and the images were prepared at 300dpi.

TOOL USED : AUTODOCK VINA 4

Molecular docking is a computational method used for the prediction of bound confirmations and the affinity required for binding. This procedure tries to predict the binding of receptors and ligands. The prediction of binding is necessary for the virtual screening of libraries of drug-like molecules. Hence this entire procedure is beneficial in the drug discovery and development process [162].

Virtual screening and molecular docking can be done by AutoDock Vina, which is new software and highly efficient than other docking software i.e. AutoDock 4. It not only provides accurate results but also calculates the grid maps and sums up results in a much more presentable and understandable manner. AutoDock Vina uses the Fletcher-Goldfarb-Shanno method for local optimization which uses gradient along with the scoring function concerning its argument [163].

Some other advancements in docking are like inverse molecular docking in which the cellular mechanism and the clinical application of known and unknown drugs can be studied. Cross docking is used to study the binding of antigen-antibody interactions. Movement of the side chains of the binding site is permitted with the use of induced fit docking. The movement takes place during the docking process [164].

CHAPTER 5

RESULTS AND DISCUSSION

When the list of best drugs was analyzed, some of them did not pass the blood-brain barrier permeability. So, a total of 15 drugs were listed out of which 10 drugs that passed the standard parameters were taken. After analysis of all the parameters, those drugs were finalized for the list of 10 drugs.

Table 4: LIST OF TOP 10 DRUGS WITH ChEMBL IDs, BINDING AFFINITY, AND BLOOD-BRAIN BARRIER PERMEABILITY

ChEMBL ID	Compound Name	Binding Affinity (kcal/mol)	IC50 value (nM)	Blood-Brain Barrier
CHEMBL225071	Raltitrexed	-13.8	14.44	+
CHEMBL409	Bicalutamide	-13.2	160	+
CHEMBL1095097	Eplerenone	-13.2	81	+
CHEMBL477772	Pazopanib	-12.6	1010	+
CHEMBL2028663	Dabrafenib	-12.5	0.7	+
CHEMBL139835	Cyproterone acetate	-12.2	7.1	+
CHEMBL1023	Bexarotene	-11.9	83	+
CHEMBL1201139	Megestrol acetate	-11.9	$3x10^4$	+
CHEMBL435	Hydrochlorothiazide	-11.6	7300	+
CHEMBL1481	Glimepiride	-11.4	7.3	+

Raltitrexed (ChEMBL ID: ChEMBL225071) proved to be the best drug. It is an anti-cancerous drug that can be repurposed for Alzheimer's disease. The binding affinity of Raltitrexed came to -13.8 kcal/mol. It is sold under the brand name Tomudex® and the drug is manufactured by AstraZeneca which is a pharmaceutical industry. The drug bank accession number is DB00293. The half-life of Raltitrexed is 198 hours.

Consumption of the drug can result in some side effects like unusual bleeding, chest pain, fever, pale skin, ulcers, etc. Considering the traditional medications, this drug is used for

the treatment of cancer and lung issues but now it can be researched for the treatment of neurodegenerative diseases like A.D. because it results in the inhibition of DNMT1which is one of the causes of Alzheimer's disease.

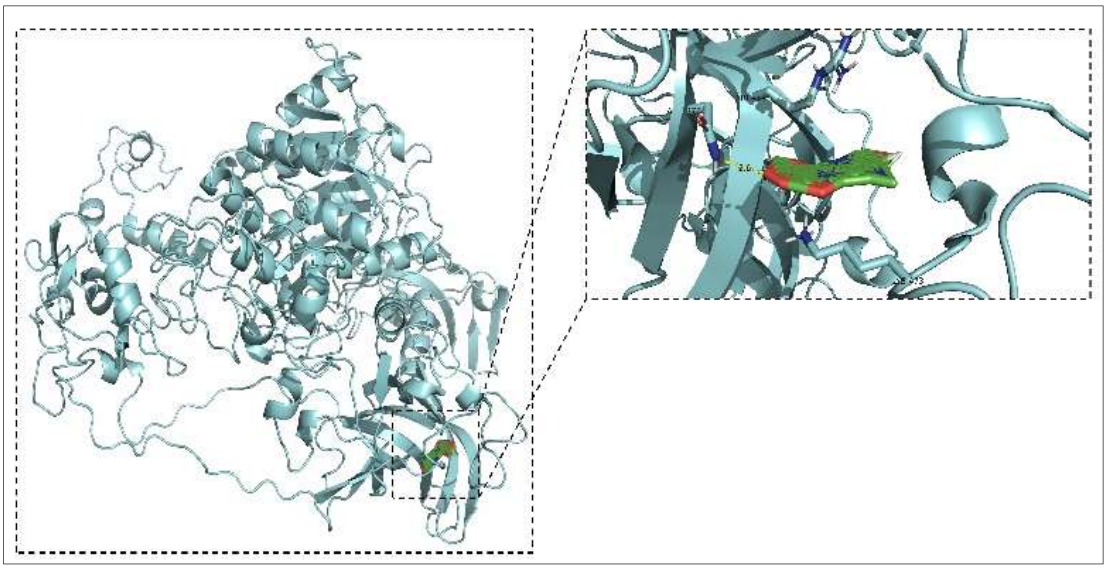

Figure 6: BINDING OF DNMT1 WITH RALTITREXED

It is well known that there have not been many advancements in the history of neurodegenerative treatments. But taking a step towards the solution to the problem can give us some positive results. As technology is advancing, various methods have been designed to work on the NDD treatment. In-silico studies are one of them where the structure and affinity of the compound can be studied. Virtual screening and molecular docking can be done to repurpose any drug for any disease. For example, we took anti-cancer, anti-diabetic, and anti-hypertension drugs to select a few drugs which can be effective in the treatment of AD and the inhibition of DNMT1 with some specific mutation.

Utilizing the advancements in the field of medicine, we repurposed Raltitrexed which is an anti-cancer drug, for the treatment of AD. It can be a useful drug for treatment because of its high binding energy of -13.8kcal/mol and Ic50 value of 14.44nM. IC50 value denotes the quantity of the drug or medicine required for the inhibition of half biological process. The use of Raltitrexed will deactivate all the pathways of DNMT1 mutants which results in the progression of the disease.

CHAPTER 5

CONCLUSION

In order to comprehend and cure human disease, drug discovery and development is a cutting-edge method of clinical practice and study. The categorization and forecasting of outcomes for both people and groups are made possible by the analysis and processing of multi-omics data using machine learning methods, including the recently developed deep learning models. Thus, the aim of AI/ML is to emulate human neurological functions. It is ushering in a growing phase in healthcare, fuelled by the increasing accessibility of medical data along with the fast advents of analytics tools. In the current era, drug discovery and development has also evolved greatly thanks to genetics and genomics, which are also well suited to machine learning technology. Drug discovery and development and AI offer great promise for the future of humankind as they have already made significant contributions to our knowledge of human health and illness. A lot of research is to be done in the field of neurological sciences with respect to AI/ML but still, great progress has been made that helps in the early detection of diseases with the help of biomarkers and omics studies. AI/ML is going to be a wonderful tool in the field of medicine and healthcare.

REFERENCES

[1] R. R. Shamir, T. Dolber, A. M. Noecker, B. L. Walter, and C. C. McIntyre, "Machine Learning Approach to Optimizing Combined Stimulation and Medication Therapies for Parkinson's Disease," *Brain Stimul*, vol. 8, no. 6, pp. 1025–1032, Nov. 2015, doi: 10.1016/j.brs.2015.06.003.

[2] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick, "Neuroscience-Inspired Artificial Intelligence," *Neuron*, vol. 95, no. 2, pp. 245–258, Jul. 2017, doi: 10.1016/j.neuron.2017.06.011.

[3] P. S. Churchland and T. J. Sejnowski, "Perspectives on Cognitive Neuroscience," *Science (1979)*, vol. 242, no. 4879, pp. 741–745, Nov. 1988, doi: 10.1126/science.3055294.

[4] "Is the brain a good model for machine intelligence?," *Nature*, vol. 482, no. 7386, pp. 462–463, Feb. 2012, doi: 10.1038/482462a.

[5] U. K. Patel *et al.*, "Artificial intelligence as an emerging technology in the current care of neurological disorders," *J Neurol*, vol. 268, no. 5, pp. 1623–1642, May 2021, doi: 10.1007/s00415-019-09518-3.

[6] K. Amit, *Artificial Intelligence and Soft Computing*. CRC Press, 2018. doi: 10.1201/9781315219738.

[7] P. Apell and H. Eriksson, "Artificial intelligence (AI) healthcare technology innovations: the current state and challenges from a life science industry perspective," *Technol Anal Strateg Manag*, vol. 35, no. 2, pp. 179–193, Feb. 2023, doi: 10.1080/09537325.2021.1971188.

[8] A. Pantelopoulos and N. G. Bourbakis, "A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 1, pp. 1–12, Jan. 2010, doi: 10.1109/TSMCC.2009.2032660.

[9] H. A. Haenssle *et al.*, "Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," *Annals of Oncology*, vol. 29, no. 8, pp. 1836–1842, Aug. 2018, doi: 10.1093/annonc/mdy166.

[10] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017, doi: 10.1038/nature21056.

[11] G. N. Brock, P. Mukhopadhyay, V. Pihur, C. Webb, R. M. Greene, and M. M. Pisano, "MmPalateMiRNA, an R package compendium illustrating analysis of miRNA microarray data," *Source Code Biol Med*, vol. 8, no. 1, p. 1, Dec. 2013, doi: 10.1186/1751-0473-8-1.

[12] E. Laurenza, M. Quintano, F. Schiavone, and D. Vrontis, "The effect of digital technologies adoption in healthcare industry: a case based analysis," *Business Process Management Journal*, vol. 24, no. 5, pp. 1124–1144, Sep. 2018, doi: 10.1108/BPMJ-04-2017-0084.

[13] M. R. Lennon *et al.*, “Readiness for Delivering Digital Health at Scale: Lessons From a Longitudinal Qualitative Evaluation of a National Digital Health Innovation Program in the United Kingdom,” *J Med Internet Res*, vol. 19, no. 2, p. e42, Feb. 2017, doi: 10.2196/jmir.6900.

[14] C. Strafella *et al.*, “Application of precision medicine in neurodegenerative diseases,” *Front Neurol*, vol. 9, no. AUG, p. 701, Aug. 2018, doi: 10.3389/FNEUR.2018.00701/BIBTEX.

[15] K. Ganapathy, S. Abdul, and A. Nursetyo, “Artificial intelligence in neurosciences: A clinician’s perspective,” *Neurol India*, vol. 66, no. 4, p. 934, 2018, doi: 10.4103/0028-3886.236971.

[16] J. Powles and H. Hodson, “Google DeepMind and healthcare in an age of algorithms,” *Health Technol (Berl)*, vol. 7, no. 4, pp. 351–367, Dec. 2017, doi: 10.1007/s12553-017-0179-1.

[17] M. Cox and D. Ellsworth, “Application-Controlled Demand Paging for Out-of-Core Visualization,” 1997.

[18] R. Kitchin, “Big Data, new epistemologies and paradigm shifts,” *Big Data Soc*, vol. 1, no. 1, Jul. 2014, doi: 10.1177/2053951714528481.

[19] J.-H. Huh, “Big Data Analysis for Personalized Health Activities: Machine Learning Processing for Automatic Keyword Extraction Approach,” *Symmetry (Basel)*, vol. 10, no. 4, p. 93, Apr. 2018, doi: 10.3390/sym10040093.

[20] M. Viceconti, P. Hunter, and R. Hose, “Big Data, Big Knowledge: Big Data for Personalized Healthcare,” *IEEE J Biomed Health Inform*, vol. 19, no. 4, pp. 1209–1215, Jul. 2015, doi: 10.1109/JBHI.2015.2406883.

[21] D. Cirillo and A. Valencia, “Big data analytics for personalized medicine,” *Current Opinion in Biotechnology*, vol. 58. Elsevier Ltd, pp. 161–167, Aug. 01, 2019. doi: 10.1016/j.copbio.2019.03.004.

[22] J. Andreu-Perez, C. C. Y. Poon, R. D. Merrifield, S. T. C. Wong, and G. Z. Yang, “Big Data for Health,” *IEEE J Biomed Health Inform*, vol. 19, no. 4, pp. 1193–1208, Jul. 2015, doi: 10.1109/JBHI.2015.2450362.

[23] M. A. Weigelt *et al.*, “Advanced Wound Diagnostics: Toward Transforming Wound Care into Precision Medicine,” *Adv Wound Care (New Rochelle)*, vol. 11, no. 6, pp. 330–359, Jun. 2022, doi: 10.1089/wound.2020.1319.

[24] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, “Big data in healthcare: management, analysis and future prospects,” *J Big Data*, vol. 6, no. 1, p. 54, Dec. 2019, doi: 10.1186/s40537-019-0217-0.

[25] Y. Wang, L. Kung, and T. A. Byrd, “Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations,” *Technol Forecast Soc Change*, vol. 126, pp. 3–13, Jan. 2018, doi: 10.1016/j.techfore.2015.12.019.

[26] T. Hulsen, “Challenges and solutions for big data in personalized healthcare,” in *Big Data in Psychiatry #x0026; Neurology*, Elsevier, 2021, pp. 69–94. doi: 10.1016/B978-0-12-822884-5.00016-7.

[27] Y. Wang, L. Kung, W. Y. C. Wang, and C. G. Cegielski, “An integrated big data analytics-enabled transformation model: Application to health care,” *Information & Management*, vol. 55, no. 1, pp. 64–79, Jan. 2018, doi: 10.1016/j.im.2017.04.001.

[28] D. Schmitter *et al.*, “An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer’s disease,” *Neuroimage Clin*, vol. 7, pp. 7–17, 2015, doi: 10.1016/j.nicl.2014.11.001.

[29] A. Karlsson *et al.*, “Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI,” *Journal of Magnetic Resonance Imaging*, vol. 41, no. 6, pp. 1558–1569, Jun. 2015, doi: 10.1002/jmri.24726.

[30] V. Jagadeeswari, V. Subramaniyaswamy, R. Logesh, and V. Vijayakumar, “A study on medical Internet of Things and Big Data in personalized healthcare system,” *Health Inf Sci Syst*, vol. 6, no. 1, p. 14, Dec. 2018, doi: 10.1007/s13755-018-0049-x.

[31] J. Zhai, S. Zhang, and C. Wang, “The classification of imbalanced large data sets based on MapReduce and ensemble of ELM classifiers,” *International Journal of Machine Learning and Cybernetics*, vol. 8, no. 3, pp. 1009–1017, Jun. 2017, doi: 10.1007/s13042-015-0478-7.

[32] X. Wu, W. Fan, J. Peng, K. Zhang, and Y. Yu, “Iterative sampling based frequent itemset mining for big data,” *International Journal of Machine Learning and Cybernetics*, vol. 6, no. 6, pp. 875–882, Dec. 2015, doi: 10.1007/s13042-015-0345-6.

[33] S. Rani, S. H. Ahmed, and S. C. Shah, “Smart Health: A Novel Paradigm to Control the Chickungunya Virus,” *IEEE Internet Things J*, vol. 6, no. 2, pp. 1306–1311, Apr. 2019, doi: 10.1109/JIOT.2018.2802898.

[34] H. Baali, H. Djelouat, A. Amira, and F. Bensaali, “Empowering Technology Enabled Care Using <italic>IoT</italic> and Smart Devices: A Review,” *IEEE Sens J*, vol. 18, no. 5, pp. 1790–1809, Mar. 2018, doi: 10.1109/JSEN.2017.2786301.

[35] S. K. Sood and I. Mahajan, “Wearable IoT sensor based healthcare system for identifying and controlling chikungunya virus,” *Comput Ind*, vol. 91, pp. 33–44, Oct. 2017, doi: 10.1016/j.compind.2017.05.006.

[36] M. W. Woo, J. Lee, and K. Park, “A reliable IoT system for Personal Healthcare Devices,” *Future Generation Computer Systems*, vol. 78, pp. 626–640, Jan. 2018, doi: 10.1016/j.future.2017.04.004.

[37] G. Manogaran, R. Varatharajan, D. Lopez, P. M. Kumar, R. Sundarasekar, and C. Thota, “A new architecture of Internet of Things and big data ecosystem for secured smart healthcare monitoring and alerting system,” *Future Generation Computer Systems*, vol. 82, pp. 375–387, May 2018, doi: 10.1016/j.future.2017.10.045.

[38] R. Sandhu, N. Kaur, S. K. Sood, and R. Buyya, “TDRM: tensor-based data representation and mining for healthcare data in cloud computing environments,” *J Supercomput*, vol. 74, no. 2, pp. 592–614, Feb. 2018, doi: 10.1007/s11227-017-2163-y.

[39] N. Kalid, A. A. Zaidan, B. B. Zaidan, O. H. Salman, M. Hashim, and H. Muzammil, “Based Real Time Remote Health Monitoring Systems: A Review on Patients Prioritization and

Related ‘Big Data’ Using Body Sensors information and Communication Technology,” *J Med Syst*, vol. 42, no. 2, p. 30, Feb. 2018, doi: 10.1007/s10916-017-0883-4.

[40] Y. Hu, K. Duan, Y. Zhang, M. S. Hossain, S. M. Mizanur Rahman, and A. Alelaiwi, “Simultaneously aided diagnosis model for outpatient departments via healthcare big data analytics,” *Multimed Tools Appl*, vol. 77, no. 3, pp. 3729–3743, Feb. 2018, doi: 10.1007/s11042-016-3719-1.

[41] F. Leu, C. Ko, I. You, K.-K. R. Choo, and C.-L. Ho, “A smartphone-based wearable sensors for monitoring real-time physiological data,” *Computers & Electrical Engineering*, vol. 65, pp. 376–392, Jan. 2018, doi: 10.1016/j.compeleceng.2017.06.031.

[42] M. Esposito, A. Minutolo, R. Megna, M. Forastiere, M. Magliulo, and G. De Pietro, “A smart mobile, self-configuring, context-aware architecture for personal health monitoring,” *Eng Appl Artif Intell*, vol. 67, pp. 136–156, Jan. 2018, doi: 10.1016/j.engappai.2017.09.019.

[43] X.-S. Zhang, F.-Y. Leu, C.-W. Yang, and L.-S. Lai, “Healthcare-based on Cloud Electrocardiogram System: A Medical Center Experience in Middle Taiwan,” *J Med Syst*, vol. 42, no. 3, p. 39, Mar. 2018, doi: 10.1007/s10916-018-0892-y.

[44] J. E. Venson, F. Bevilacqua, J. Berni, F. Onuki, and A. Maciel, “Diagnostic concordance between mobile interfaces and conventional workstations for emergency imaging assessment,” *Int J Med Inform*, vol. 113, pp. 1–8, May 2018, doi: 10.1016/j.ijmedinf.2018.01.019.

[45] H.-Y. Kao, C.-W. Wei, M.-C. Yu, T.-Y. Liang, W.-H. Wu, and Y. J. Wu, “Integrating a mobile health applications for self-management to enhance Telecare system,” *Telematics and Informatics*, vol. 35, no. 4, pp. 815–825, Jul. 2018, doi: 10.1016/j.tele.2017.12.011.

[46] B. Mesko, “The role of artificial intelligence in precision medicine,” *http://dx.doi.org/10.1080/23808993.2017.1380516*, vol. 2, no. 5, pp. 239–241, 2017, doi: 10.1080/23808993.2017.1380516.

[47] B. Meskó, Z. Drobni, É. Bényei, B. Gergely, and Z. Győrffy, “Digital health is a cultural transformation of traditional healthcare,” *Mhealth*, vol. 3, pp. 38–38, Sep. 2017, doi: 10.21037/mhealth.2017.08.07.

[48] E. Elenko, L. Underwood, and D. Zohar, “Defining digital medicine,” *Nat Biotechnol*, vol. 33, no. 5, pp. 456–461, May 2015, doi: 10.1038/nbt.3222.

[49] N. Nagarajan, E. K. Y. Yapp, N. Q. K. Le, B. Kamaraj, A. M. Al-Subaie, and H.-Y. Yeh, “Application of Computational Biology and Artificial Intelligence Technologies in Cancer Precision Drug Discovery,” *Biomed Res Int*, vol. 2019, pp. 1–15, Nov. 2019, doi: 10.1155/2019/8427042.

[50] A. Esteva *et al.*, “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017, doi: 10.1038/nature21056.

[51] G. S. Ginsburg and K. A. Phillips, “Precision Medicine: From Science To Value,” *Health Aff*, vol. 37, no. 5, pp. 694–701, May 2018, doi: 10.1377/hlthaff.2017.1624.

[52] K. A. Phillips, J. R. Trosman, R. K. Kelley, M. J. Pletcher, M. P. Douglas, and C. B. Weldon, "Genomic Sequencing: Assessing The Health Care System, Policy, And Big-Data Implications," *Health Aff*, vol. 33, no. 7, pp. 1246–1253, Jul. 2014, doi: 10.1377/hlthaff.2014.0020.

[53] D. E. Pritchard, F. Moeckel, M. S. Villa, L. T. Housman, C. A. McCarty, and H. L. McLeod, "Strategies for integrating personalized medicine into healthcare practice," *Per Med*, vol. 14, no. 2, pp. 141–152, Mar. 2017, doi: 10.2217/pme-2016-0064.

[54] S. Maudsley, V. Devanarayan, B. Martin, and H. Geerts, "Intelligent and effective informatic deconvolution of 'Big Data' and its future impact on the quantitative nature of neurodegenerative disease therapy," *Alzheimer's & Dementia*, vol. 14, no. 7, pp. 961–975, Jul. 2018, doi: 10.1016/j.jalz.2018.01.014.

[55] D. Grapov, J. Fahrmann, K. Wanichthanarak, and S. Khoomrung, "Rise of Deep Learning for Genomic, Proteomic, and Metabolomic Data Integration in Precision Medicine," *OMICS*, vol. 22, no. 10, pp. 630–636, Oct. 2018, doi: 10.1089/omi.2018.0097.

[56] C. Manzoni, P. A. Lewis, and R. Ferrari, "Network Analysis for Complex Neurodegenerative Diseases," *Curr Genet Med Rep*, vol. 8, no. 1, pp. 17–25, Mar. 2020, doi: 10.1007/s40142-020-00181-z.

[57] Y. Perez-Riverol *et al.*, "Quantifying the impact of public omics data," *Nat Commun*, vol. 10, no. 1, p. 3512, Aug. 2019, doi: 10.1038/s41467-019-11461-w.

[58] N. Perakakis, A. Yazdani, G. E. Karniadakis, and C. Mantzoros, "Omics, big data and machine learning as tools to propel understanding of biological mechanisms and to discover novel diagnostics and therapeutics," *Metabolism*, vol. 87, pp. A1–A9, Oct. 2018, doi: 10.1016/j.metabol.2018.08.002.

[59] A. L. Young *et al.*, "Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference," *Nat Commun*, vol. 9, no. 1, p. 4273, Oct. 2018, doi: 10.1038/s41467-018-05892-0.

[60] J. M. Gaziano *et al.*, "Million Veteran Program: A mega-biobank to study genetic influences on health and disease," *J Clin Epidemiol*, vol. 70, pp. 214–223, Feb. 2016, doi: 10.1016/j.jclinepi.2015.09.016.

[61] C. Sudlow *et al.*, "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age," *PLoS Med*, vol. 12, no. 3, p. e1001779, Mar. 2015, doi: 10.1371/journal.pmed.1001779.

[62] D. Grapov, J. Fahrmann, K. Wanichthanarak, and S. Khoomrung, "Rise of Deep Learning for Genomic, Proteomic, and Metabolomic Data Integration in Precision Medicine," *OMICS*, vol. 22, no. 10, pp. 630–636, Oct. 2018, doi: 10.1089/omi.2018.0097.

[63] Q. Wang, W.-X. Peng, L. Wang, and L. Ye, "Toward multiomics-based next-generation diagnostics for precision medicine," *Per Med*, vol. 16, no. 2, pp. 157–170, Mar. 2019, doi: 10.2217/pme-2018-0085.

[64] R. W. Peck, "Precision Medicine Is Not Just Genomics: The Right Dose for Every Patient," *Annu Rev Pharmacol Toxicol*, vol. 58, no. 1, pp. 105–122, Jan. 2018, doi: 10.1146/annurev-pharmtox-010617-052446.

[65] R. Hodson, “Precision medicine,” *Nature*, vol. 537, no. 7619, pp. S49–S49, Sep. 2016, doi: 10.1038/537S49a.

[66] N. Perakakis, A. Yazdani, G. E. Karniadakis, and C. Mantzoros, “Omics, big data and machine learning as tools to propel understanding of biological mechanisms and to discover novel diagnostics and therapeutics,” *Metabolism*, vol. 87, pp. A1–A9, Oct. 2018, doi: 10.1016/j.metabol.2018.08.002.

[67] A. R. Sonawane, S. T. Weiss, K. Glass, and A. Sharma, “Network Medicine in the Age of Biomedical Big Data,” *Front Genet*, vol. 10, Apr. 2019, doi: 10.3389/fgene.2019.00294.

[68] T. A. Olafsdottir *et al.*, “Comparative Systems Analyses Reveal Molecular Signatures of Clinically tested Vaccine Adjuvants,” *Sci Rep*, vol. 6, no. 1, p. 39097, Dec. 2016, doi: 10.1038/srep39097.

[69] P. J. Fabres, C. Collins, T. R. Cavagnaro, and C. M. Rodríguez López, “A Concise Review on Multi-Omics Data Integration for Terroir Analysis in Vitis vinifera,” *Front Plant Sci*, vol. 8, Jun. 2017, doi: 10.3389/fpls.2017.01065.

[70] K. Wanichthanarak, J. F. Fahrmann, and D. Grapov, “Genomic, Proteomic, and Metabolomic Data Integration Strategies,” *Biomark Insights*, vol. 10s4, p. BMI.S29511, Jan. 2015, doi: 10.4137/BMI.S29511.

[71] M. Kim, N. Rai, V. Zorraquino, and I. Tagkopoulos, “Multi-omics integration accurately predicts cellular state in unexplored conditions for Escherichia coli,” *Nat Commun*, vol. 7, no. 1, p. 13090, Oct. 2016, doi: 10.1038/ncomms13090.

[72] D. Lin, J. Zhang, J. Li, H. He, H.-W. Deng, and Y.-P. Wang, “Integrative analysis of multiple diverse omics datasets by sparse group multitask regression,” *Front Cell Dev Biol*, vol. 2, Oct. 2014, doi: 10.3389/fcell.2014.00062.

[73] P. L. De Jager, H.-S. Yang, and D. A. Bennett, “Deconstructing and targeting the genomic architecture of human neurodegeneration,” *Nat Neurosci*, vol. 21, no. 10, pp. 1310–1317, Oct. 2018, doi: 10.1038/s41593-018-0240-z.

[74] Y. Li, F.-X. Wu, and A. Ngom, “A review on machine learning principles for multi-view biological data integration,” *Brief Bioinform*, p. bbw113, Dec. 2016, doi: 10.1093/bib/bbw113.

[75] A. J. Espay *et al.*, “Biomarker-driven phenotyping in Parkinson’s disease: A translational missing link in disease-modifying clinical trials,” *Movement Disorders*, vol. 32, no. 3, pp. 319–324, Mar. 2017, doi: 10.1002/mds.26913.

[76] C. J. Bell *et al.*, “Carrier Testing for Severe Childhood Recessive Diseases by Next-Generation Sequencing,” *Sci Transl Med*, vol. 3, no. 65, Jan. 2011, doi: 10.1126/scitranslmed.3001756.

[77] J. O. Kitzman *et al.*, “Noninvasive Whole-Genome Sequencing of a Human Fetus,” *Sci Transl Med*, vol. 4, no. 137, Jun. 2012, doi: 10.1126/scitranslmed.3004323.

[78] C. J. Saunders *et al.*, “Rapid Whole-Genome Sequencing for Genetic Disease Diagnosis in Neonatal Intensive Care Units,” *Sci Transl Med*, vol. 4, no. 154, Oct. 2012, doi: 10.1126/scitranslmed.3004041.

[79] E. J. Topol, “Individualized Medicine from Prewomb to Tomb,” *Cell*, vol. 157, no. 1, pp. 241–253, Mar. 2014, doi: 10.1016/j.cell.2014.02.012.

[80] J. J. McCarthy, H. L. McLeod, and G. S. Ginsburg, “Genomic Medicine: A Decade of Successes, Challenges, and Opportunities,” *Sci Transl Med*, vol. 5, no. 189, Jun. 2013, doi: 10.1126/scitranslmed.3005785.

[81] B. Khattak, A. Khan, K. Khan, W. Khan, M. Kamran, and M. Fahad, “Empirical Analysis of Recent Advances, Characteristics and Challenges of Big Data,” *ICST Transactions on Scalable Information Systems*, vol. 0, no. 0, p. 159621, Jul. 2018, doi: 10.4108/eai.13-7-2018.159621.

[82] C. S. Kruse, R. Goswamy, Y. Raval, and S. Marawi, “Challenges and Opportunities of Big Data in Health Care: A Systematic Review,” *JMIR Med Inform*, vol. 4, no. 4, p. e38, Nov. 2016, doi: 10.2196/medinform.5359.

[83] B. Blobel, D. M. Lopez, and C. Gonzalez, “Patient privacy and security concerns on big data for personalized medicine,” *Health Technol (Berl)*, vol. 6, no. 1, pp. 75–81, Jun. 2016, doi: 10.1007/s12553-016-0127-5.

[84] E.-H. W. Kluge, “Health Information, the Fair Information Principles and Ethics,” *Methods Inf Med*, vol. 33, no. 04, pp. 336–345, Feb. 1994, doi: 10.1055/s-0038-1635041.

[85] C. Dwork, “Differential Privacy,” 2006, pp. 1–12. doi: 10.1007/11787006_1.

[86] K. Nissim *et al.*, “Differential Privacy: A Primer for a Non-technical Audience * (Preliminary version),” 2017. [Online]. Available: http://privacytools.seas.harvard.edu

[87] J. Soria-Comas, J. Domingo-Ferrer, D. Sánchez, and S. Martínez, “Enhancing data utility in differential privacy via microaggregation-based $$k$$ k -anonymity,” *The VLDB Journal*, vol. 23, no. 5, pp. 771–794, Oct. 2014, doi: 10.1007/s00778-014-0351-4.

[88] USENIX Association., *Proceedings of the seventeenth Large Installation Systems Administration Conference (LISA XVII) : October 26-31, 2003 San Diego, CA, USA*. USENIX Association, 2003.

[89] A. Johnson and V. Shmatikov, “Privacy-preserving data exploration in genome-wide association studies,” in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA: ACM, Aug. 2013, pp. 1079–1087. doi: 10.1145/2487575.2487687.

[90] F. Yu and Z. Ji, “Scalable privacy-preserving data sharing methodology for genome-wide association studies: an application to iDASH healthcare privacy protection challenge,” *BMC Med Inform Decis Mak*, vol. 14, no. S1, p. S3, Dec. 2014, doi: 10.1186/1472-6947-14-S1-S3.

[91] S. Simmons and B. Berger, “Realizing privacy preserving genome-wide association studies,” *Bioinformatics*, vol. 32, no. 9, pp. 1293–1300, May 2016, doi: 10.1093/bioinformatics/btw009.

[92] K. H. C. Li *et al.*, “The Current State of Mobile Phone Apps for Monitoring Heart Rate, Heart Rate Variability, and Atrial Fibrillation: Narrative Review,” *JMIR Mhealth Uhealth*, vol. 7, no. 2, p. e11606, Feb. 2019, doi: 10.2196/11606.

[93] K. C. Ho, W. Speier, H. Zhang, F. Scalzo, S. El-Saden, and C. W. Arnold, “A Machine Learning Approach for Classifying Ischemic Stroke Onset Time From Imaging,” *IEEE Trans Med Imaging*, vol. 38, no. 7, pp. 1666–1676, Jul. 2019, doi: 10.1109/TMI.2019.2901445.

[94] D. L. Labovitz, L. Shafner, M. Reyes Gil, D. Virmani, and A. Hanina, “Using Artificial Intelligence to Reduce the Risk of Nonadherence in Patients on Anticoagulation Therapy,” *Stroke*, vol. 48, no. 5, pp. 1416–1419, May 2017, doi: 10.1161/STROKEAHA.116.016281.

[95] E. M. Okazaki *et al.*, “Usage of EpiFinder clinical decision support in the assessment of epilepsy,” *Epilepsy & Behavior*, vol. 82, pp. 140–143, May 2018, doi: 10.1016/j.yebeh.2018.03.018.

[96] Ming-Zher Poh, T. Loddenkemper, N. C. Swenson, S. Goyal, J. R. Madsen, and R. W. Picard, “Continuous monitoring of electrodermal activity during epileptic seizures using a wearable sensor,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, IEEE, Aug. 2010, pp. 4415–4418. doi: 10.1109/IEMBS.2010.5625988.

[97] S. Majumder, T. Mondal, and M. Deen, “Wearable Sensors for Remote Health Monitoring,” *Sensors*, vol. 17, no. 12, p. 130, Jan. 2017, doi: 10.3390/s17010130.

[98] M. Pastorino, M. T. Arredondo, J. Cancela, and S. Guillen, “Wearable sensor network for health monitoring: the case of Parkinson disease,” *J Phys Conf Ser*, vol. 450, p. 012055, Jun. 2013, doi: 10.1088/1742-6596/450/1/012055.

[99] K. Mercer, M. Li, L. Giangregorio, C. Burns, and K. Grindrod, “Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis,” *JMIR Mhealth Uhealth*, vol. 4, no. 2, p. e40, Apr. 2016, doi: 10.2196/mhealth.4461.

[100] J. Takacs, C. L. Pollock, J. R. Guenther, M. Bahar, C. Napier, and M. A. Hunt, “Validation of the Fitbit One activity monitor device during treadmill walking,” *J Sci Med Sport*, vol. 17, no. 5, pp. 496–500, Sep. 2014, doi: 10.1016/j.jsams.2013.10.241.

[101] R. Yang, E. Shin, M. W. Newman, and M. S. Ackerman, “When fitness trackers don’t ‘fit,’” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, NY, USA: ACM, Sep. 2015, pp. 623–634. doi: 10.1145/2750858.2804269.

[102] J. Clawson, J. A. Pater, A. D. Miller, E. D. Mynatt, and L. Mamykina, “No longer wearing,” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, NY, USA: ACM, Sep. 2015, pp. 647–658. doi: 10.1145/2750858.2807554.

[103] J. Varghese *et al.*, “A Smart Device System to Identify New Phenotypical Characteristics in Movement Disorders,” *Front Neurol*, vol. 10, Jan. 2019, doi: 10.3389/fneur.2019.00048.

[104] M. Tan *et al.*, “Self-powered smart patch promotes skin nerve regeneration and sensation restoration by delivering biological-electrical signals in program,” *Biomaterials*, vol. 283, p. 121413, Apr. 2022, doi: 10.1016/j.biomaterials.2022.121413.

[105] E. M. Okazaki *et al.*, "Usage of EpiFinder clinical decision support in the assessment of epilepsy," *Epilepsy & Behavior*, vol. 82, pp. 140–143, May 2018, doi: 10.1016/j.yebeh.2018.03.018.

[106] K. H. C. Li *et al.*, "The Current State of Mobile Phone Apps for Monitoring Heart Rate, Heart Rate Variability, and Atrial Fibrillation: Narrative Review," *JMIR Mhealth Uhealth*, vol. 7, no. 2, p. e11606, Feb. 2019, doi: 10.2196/11606.

[107] K. C. Ho, W. Speier, H. Zhang, F. Scalzo, S. El-Saden, and C. W. Arnold, "A Machine Learning Approach for Classifying Ischemic Stroke Onset Time From Imaging," *IEEE Trans Med Imaging*, vol. 38, no. 7, pp. 1666–1676, Jul. 2019, doi: 10.1109/TMI.2019.2901445.

[108] R. M. S. Visscher, N. Feddermann-Demont, F. Romano, D. Straumann, and G. Bertolini, "Artificial intelligence for understanding concussion: Retrospective cluster analysis on the balance and vestibular diagnostic data of concussion patients," *PLoS One*, vol. 14, no. 4, p. e0214525, Apr. 2019, doi: 10.1371/journal.pone.0214525.

[109] K. R. V *et al.*, "Machine learning detects EEG microstate alterations in patients living with temporal lobe epilepsy," *Seizure*, vol. 61, pp. 8–13, Oct. 2018, doi: 10.1016/j.seizure.2018.07.007.

[110] K. L. Chan *et al.*, "Early Identification of High-Risk TIA or Minor Stroke Using Artificial Neural Network," *Front Neurol*, vol. 10, Mar. 2019, doi: 10.3389/fneur.2019.00171.

[111] F. Onorati *et al.*, "Multicenter clinical assessment of improved wearable multimodal convulsive seizure detectors," *Epilepsia*, vol. 58, no. 11, pp. 1870–1879, Nov. 2017, doi: 10.1111/epi.13899.

[112] Ming-Zher Poh, T. Loddenkemper, N. C. Swenson, S. Goyal, J. R. Madsen, and R. W. Picard, "Continuous monitoring of electrodermal activity during epileptic seizures using a wearable sensor," in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, IEEE, Aug. 2010, pp. 4415–4418. doi: 10.1109/IEMBS.2010.5625988.

[113] R. Vasta *et al.*, "The application of artificial intelligence to understand the pathophysiological basis of psychogenic nonepileptic seizures," *Epilepsy & Behavior*, vol. 87, pp. 167–172, Oct. 2018, doi: 10.1016/j.yebeh.2018.09.008.

[114] E. Urtnasan, E. Y. Joo, and K. H. Lee, "AI-Enabled Algorithm for Automatic Classification of Sleep Disorders Based on Single-Lead Electrocardiogram," *Diagnostics*, vol. 11, no. 11, p. 2054, Nov. 2021, doi: 10.3390/diagnostics11112054.

[115] D. L. Labovitz, L. Shafner, M. Reyes Gil, D. Virmani, and A. Hanina, "Using Artificial Intelligence to Reduce the Risk of Nonadherence in Patients on Anticoagulation Therapy," *Stroke*, vol. 48, no. 5, pp. 1416–1419, May 2017, doi: 10.1161/STROKEAHA.116.016281.

[116] P. Taillon-Miller, Z. Gu, Q. Li, L. Hillier, and P.-Y. Kwok, "Overlapping Genomic Sequences: A Treasure Trove of Single-Nucleotide Polymorphisms," *Genome Res*, vol. 8, no. 7, pp. 748–754, Jul. 1998, doi: 10.1101/gr.8.7.748.

[117] A. Auton *et al.*, “A global reference for human genetic variation,” *Nature*, vol. 526, no. 7571, pp. 68–74, Oct. 2015, doi: 10.1038/nature15393.

[118] S. Sunyaev, W. Lathe, and P. Bork, “Integration of genome data and protein structures: prediction of protein folds, protein interactions and ‘molecular phenotypes’ of single nucleotide polymorphisms,” *Curr Opin Struct Biol*, vol. 11, no. 1, pp. 125–130, Feb. 2001, doi: 10.1016/S0959-440X(00)00175-5.

[119] R. E. Steward, M. W. MacArthur, R. A. Laskowski, and J. M. Thornton, “Molecular basis of inherited diseases: a structural perspective,” *Trends in Genetics*, vol. 19, no. 9, pp. 505–513, Sep. 2003, doi: 10.1016/S0168-9525(03)00195-1.

[120] J. C. Fay, G. J. Wyckoff, and C.-I. Wu, “Positive and Negative Selection on the Human Genome,” *Genetics*, vol. 158, no. 3, pp. 1227–1234, Jul. 2001, doi: 10.1093/genetics/158.3.1227.

[121] S. Mooney, “Bioinformatics approaches and resources for single nucleotide polymorphism functional analysis,” *Brief Bioinform*, vol. 6, no. 1, pp. 44–56, Jan. 2005, doi: 10.1093/bib/6.1.44.

[122] J. Shen, P. L. Deininger, and H. Zhao, “Applications of computational algorithm tools to identify functional SNPs in cytokine genes,” *Cytokine*, vol. 35, no. 1–2, pp. 62–66, Jul. 2006, doi: 10.1016/j.cyto.2006.07.008.

[123] W. Schierding *et al.*, “GWAS on prolonged gestation (post-term birth): analysis of successive Finnish birth cohorts,” *J Med Genet*, vol. 55, no. 1, pp. 55–63, Jan. 2018, doi: 10.1136/jmedgenet-2017-104880.

[124] P. M. Visscher *et al.*, “10 Years of GWAS Discovery: Biology, Function, and Translation,” *The American Journal of Human Genetics*, vol. 101, no. 1, pp. 5–22, Jul. 2017, doi: 10.1016/j.ajhg.2017.06.005.

[125] Z. Zhang, M. A. Miteva, L. Wang, and E. Alexov, “Analyzing Effects of Naturally Occurring Missense Mutations,” *Comput Math Methods Med*, vol. 2012, pp. 1–15, 2012, doi: 10.1155/2012/805827.

[126] N. J. Schork, D. Fallin, and J. S. Lanchbury, “Single nucleotide polymorphisms and the future of genetic epidemiology,” *Clin Genet*, vol. 58, no. 4, pp. 250–264, Oct. 2000, doi: 10.1034/j.1399-0004.2000.580402.x.

[127] T. Fadason, W. Schierding, T. Lumley, and J. M. O’Sullivan, “Chromatin interactions and expression quantitative trait loci reveal genetic drivers of multimorbidities,” *Nat Commun*, vol. 9, no. 1, p. 5198, Dec. 2018, doi: 10.1038/s41467-018-07692-y.

[128] T. Fadason, C. Ekblad, J. R. Ingram, W. S. Schierding, and J. M. O’Sullivan, “Physical Interactions and Expression Quantitative Traits Loci Identify Regulatory Connections for Obesity and Type 2 Diabetes Associated SNPs,” *Front Genet*, vol. 8, Oct. 2017, doi: 10.3389/fgene.2017.00150.

[129] W. Schierding, J. Antony, W. S. Cutfield, J. A. Horsfield, and J. M. O’Sullivan, “Intergenic GWAS SNPs are key components of the spatial and regulatory network for human growth,” *Hum Mol Genet*, vol. 25, no. 15, pp. 3372–3382, Aug. 2016, doi: 10.1093/hmg/ddw165.

[130] W. Schierding and J. M. O'Sullivan, "Connecting SNPs in Diabetes: A Spatial Analysis of Meta-GWAS Loci," *Front Endocrinol (Lausanne)*, vol. 6, Jul. 2015, doi: 10.3389/fendo.2015.00102.

[131] X. Wang, G. Strizich, Y. Hu, T. Wang, R. C. Kaplan, and Q. Qi, "Genetic markers of type 2 diabetes: Progress in genome-wide association studies and clinical application for risk prediction," *J Diabetes*, vol. 8, no. 1, pp. 24–35, Jan. 2016, doi: 10.1111/1753-0407.12323.

[132] G. Abraham and M. Inouye, "Genomic risk prediction of complex human disease and its clinical application," *Curr Opin Genet Dev*, vol. 33, pp. 10–16, Aug. 2015, doi: 10.1016/j.gde.2015.06.005.

[133] G. Abraham and M. Inouye, "Genomic risk prediction of complex human disease and its clinical application," *Curr Opin Genet Dev*, vol. 33, pp. 10–16, Aug. 2015, doi: 10.1016/j.gde.2015.06.005.

[134] A. Delamarre, W. M.-L. P. Médicale, and undefined 2017, "Epidemiology, environmental risk factors and genetics of Parkinson's disease," *Elsevier*, Accessed: Feb. 27, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0755498217300015

[135] "Artificial Intelligence | Syneos Health Communications." https://syneoshealthcommunications.com/perspectives/artificial-intelligence (accessed Feb. 27, 2023).

[136] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick, "Neuroscience-Inspired Artificial Intelligence," *Neuron*, vol. 95, no. 2. Cell Press, pp. 245–258, Jul. 19, 2017. doi: 10.1016/j.neuron.2017.06.011.

[137] M. Monteiro *et al.*, "Using Machine Learning to Improve the Prediction of Functional Outcome in Ischemic Stroke Patients," *IEEE/ACM Trans Comput Biol Bioinform*, vol. 15, no. 6, pp. 1953–1959, Nov. 2018, doi: 10.1109/TCBB.2018.2811471.

[138] "Pattern Recognition and Machine Learning," *J Electron Imaging*, vol. 16, no. 4, p. 049901, Jan. 2007, doi: 10.1117/1.2819119.

[139] T. Shaikhina, D. Lowe, S. Daga, D. Briggs, R. Higgins, and N. Khovanova, "Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation," *Biomed Signal Process Control*, vol. 52, pp. 456–462, Jul. 2019, doi: 10.1016/j.bspc.2017.01.012.

[140] E. H. Park *et al.*, "Basic Facts of Breast Cancer in Korea in 2014: The 10-Year Overall Survival Progress," *J Breast Cancer*, vol. 20, no. 1, p. 1, 2017, doi: 10.4048/jbc.2017.20.1.1.

[141] Y. Zhang *et al.*, "Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: decision tree, *k* -nearest neighbors, and support vector machine," *Simulation*, vol. 92, no. 9, pp. 861–871, Sep. 2016, doi: 10.1177/0037549716666962.

[142] R. Vashistha, D. Yadav, D. Chhabra, and P. Shukla, "Artificial intelligence integration for neurodegenerative disorders," in *Leveraging Biomedical and Healthcare Data:*

Semantics, Analytics and Knowledge, Elsevier, 2018, pp. 77–89. doi: 10.1016/B978-0-12-809556-0.00005-8.

[143] S. I. R. Bhagyashree, K. Nagaraj, M. Prince, C. H. D. Fall, and M. Krishna, "Diagnosis of Dementia by Machine learning methods in Epidemiological studies: a pilot exploratory study from south India," *Soc Psychiatry Psychiatr Epidemiol*, vol. 53, no. 1, pp. 77–86, Jan. 2018, doi: 10.1007/s00127-017-1410-0.

[144] G. Orrù, W. Pettersson-Yeo, A. F. Marquand, G. Sartori, and A. Mechelli, "Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review," *Neurosci Biobehav Rev*, vol. 36, no. 4, pp. 1140–1152, Apr. 2012, doi: 10.1016/j.neubiorev.2012.01.004.

[145] A. Eloyan *et al.*, "Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging," *Front Syst Neurosci*, vol. 6, 2012, doi: 10.3389/fnsys.2012.00061.

[146] A. Subasi and M. Ismail Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Syst Appl*, vol. 37, no. 12, pp. 8659–8666, Dec. 2010, doi: 10.1016/j.eswa.2010.06.065.

[147] D. S. Knopman *et al.*, "Alzheimer disease," *Nat Rev Dis Primers*, vol. 7, no. 1, p. 33, May 2021, doi: 10.1038/s41572-021-00269-y.

[148] I. G. McKeith *et al.*, "Diagnosis and management of dementia with Lewy bodies," *Neurology*, vol. 89, no. 1, pp. 88–100, Jul. 2017, doi: 10.1212/WNL.0000000000004058.

[149] G. M. Sancesario and S. Bernardini, "Diagnosis of neurodegenerative dementia: where do we stand, now?," *Ann Transl Med*, vol. 6, no. 17, pp. 340–340, Sep. 2018, doi: 10.21037/atm.2018.08.04.

[150] S. T. Pendlebury and P. M. Rothwell, "Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis," *Lancet Neurol*, vol. 8, no. 11, pp. 1006–1018, Nov. 2009, doi: 10.1016/S1474-4422(09)70236-4.

[151] C. Reitz, "Toward precision medicine in Alzheimer's disease," *Ann Transl Med*, vol. 4, no. 6, pp. 107–107, Mar. 2016, doi: 10.21037/atm.2016.03.05.

[152] B. Mirheidari *et al.*, "Computational Cognitive Assessment: Investigating the Use of an Intelligent Virtual Agent for the Detection of Early Signs of Dementia," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, May 2019, pp. 2732–2736. doi: 10.1109/ICASSP.2019.8682423.

[153] M. R. Kumar *et al.*, "Dementia Detection from Speech Using Machine Learning and Deep Learning Architectures," *Sensors*, vol. 22, no. 23, p. 9311, Nov. 2022, doi: 10.3390/s22239311.

[154] K. C. Fraser, J. A. Meltzer, and F. Rudzicz, "Linguistic Features Identify Alzheimer's Disease in Narrative Speech," *Journal of Alzheimer's Disease*, vol. 49, no. 2, pp. 407–422, Nov. 2015, doi: 10.3233/JAD-150520.

[155] D. Cui and X. Xu, “DNA Methyltransferases, DNA Methylation, and Age-Associated Cognitive Function,” *Int J Mol Sci*, vol. 19, no. 5, p. 1315, Apr. 2018, doi: 10.3390/ijms19051315.

[156] T. H. Bestor, “The DNA methyltransferases of mammals,” *Hum Mol Genet*, vol. 9, no. 16, pp. 2395–2402, Oct. 2000, doi: 10.1093/hmg/9.16.2395.

[157] L. Shi and J. Wu, “Epigenetic regulation in mammalian preimplantation embryo development,” *Reproductive Biology and Endocrinology*, vol. 7, no. 1, p. 59, Dec. 2009, doi: 10.1186/1477-7827-7-59.

[158] M. Okano, D. W. Bell, D. A. Haber, and E. Li, “DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development,” *Cell*, vol. 99, no. 3, pp. 247–257, Oct. 1999, doi: 10.1016/S0092-8674(00)81656-6.

[159] S. Sinha and D. Vohora, “Drug Discovery and Development,” in *Pharmaceutical Medicine and Translational Clinical Research*, Elsevier, 2018, pp. 19–32. doi: 10.1016/B978-0-12-802103-3.00002-X.

[160] D. Cook *et al.*, “Lessons learned from the fate of AstraZeneca’s drug pipeline: a five-dimensional framework,” *Nat Rev Drug Discov*, vol. 13, no. 6, pp. 419–431, Jun. 2014, doi: 10.1038/nrd4309.

[161] J. L. Cummings, T. Morstorf, and K. Zhong, “Alzheimer’s disease drug-development pipeline: few candidates, frequent failures,” *Alzheimers Res Ther*, vol. 6, no. 4, p. 37, 2014, doi: 10.1186/alzrt269.

[162] J. Eberhardt, D. Santos-Martins, A. F. Tillack, and S. Forli, “AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings,” *J Chem Inf Model*, vol. 61, no. 8, pp. 3891–3898, Aug. 2021, doi: 10.1021/acs.jcim.1c00203.

[163] J. Eberhardt, D. Santos-Martins, A. F. Tillack, and S. Forli, “AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings,” *J Chem Inf Model*, vol. 61, no. 8, pp. 3891–3898, Aug. 2021, doi: 10.1021/acs.jcim.1c00203.

[164] O. Trott and A. J. Olson, “AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading,” *J Comput Chem*, p. NA-NA, 2009, doi: 10.1002/jcc.21334.

turnitin

Similarity Report ID: oid:27535:36559792

10% Overall Similarity

Top sources found in the following databases:

- 8% Internet database
- Crossref database
- 6% Submitted Works database
- 4% Publications database
- Crossref Posted Content database

31/05/2023

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	dspace.dtu.ac.in:8080 Internet	3%
2	Richard C. Mohs, Nigel H. Greig. "Drug discovery and development: Rol... Crossref	<1%
3	liebertpub.com Internet	<1%
4	tandfonline.com Internet	<1%
5	Demis Hassabis, Dharshan Kumaran, Christopher Summerfield, Matthe... Crossref	<1%
6	mjsat.com.my Internet	<1%
7	Aberystwyth University on 2023-04-28 Submitted works	<1%
8	CSU, San Marcos on 2022-07-02 Submitted works	<1%

9	hindawi.com Internet	<1%
10	University of Northumbria at Newcastle on 2022-04-25 Submitted works	<1%
11	ebin.pub Internet	<1%
12	mdpi.com Internet	<1%
13	era.ed.ac.uk Internet	<1%
14	patents.google.com Internet	<1%
15	The University of Manchester on 2019-04-12 Submitted works	<1%
16	libcatalog.embl.de Internet	<1%
17	frontiersin.org Internet	<1%
18	City University of Hong Kong on 2022-01-22 Submitted works	<1%
19	dokumen.pub Internet	<1%
20	escholarship.org Internet	<1%

21	openaccess.library.uitm.edu.my Internet	<1%
22	pubmed.ncbi.nlm.nih.gov Internet	<1%
23	researchgate.net Internet	<1%
24	Bournemouth University on 2022-12-15 Submitted works	<1%
25	Liyan Sun, Guixia Liu, Rongquan Wang. "SHEIB-AGM: A Novel Stochasti... Crossref	<1%
26	Piyush Agrawal, Navami Jain, Vishaka Gopalan, Annan Timon, Arashde... Crossref posted content	<1%
27	res.mdpi.com Internet	<1%
28	publicsafety.gc.ca Internet	<1%
29	slideshare.net Internet	<1%
30	Institute of Graduate Studies, UiTM on 2014-03-25 Submitted works	<1%
31	Napier University on 2018-11-01 Submitted works	<1%
32	University of Sheffield on 2017-12-10 Submitted works	<1%

33	University of Stirling on 2022-02-25 Submitted works	<1%
34	University of York on 2022-08-22 Submitted works	<1%
35	atm.amegroups.com Internet	<1%
36	ediss.uni-goettingen.de Internet	<1%
37	ipfs.io Internet	<1%
38	mts.intechopen.com Internet	<1%
39	pastebin.com Internet	<1%
40	pure.rug.nl Internet	<1%
41	jddtonline.info Internet	<1%
42	"2015 Alzheimer's disease facts and figures", Alzheimer s & Dementia, ... Crossref	<1%
43	"Computational Systems Biology", Springer Nature, 2018 Crossref	<1%
44	"DNA Methyltransferases - Role and Function", Springer Science and B... Crossref	<1%

45	Almir Badnjević, Lejla Gurbeta Pokvić, Mehrija Hasičić, Lejla Bandić et ... Crossref	<1%
46	Delhi Technological University on 2018-12-02 Submitted works	<1%
47	Glyndwr University on 2023-01-03 Submitted works	<1%
48	Griffith College Dublin on 2023-05-13 Submitted works	<1%
49	Hiroki Tanaka, Hiroyoshi Adachi, Norimichi Ukita, Manabu Ikeda, Hiroa... Crossref	<1%
50	Leiden University on 2021-02-10 Submitted works	<1%
51	Mostafa Haghi Kashani, Mona Madanipour, Mohammad Nikravan, Parv... Crossref	<1%
52	Rajat Vashistha, Dinesh Yadav, Deepak Chhabra, Pratyoosh Shukla. "Ar... Crossref	<1%
53	Raneem Alharthi, Rajwa Alharthi, Benjamin Guthier, Abdulmotaleb El S... Crossref	<1%
54	Sofia de la Fuente Garcia, Craig Ritchie, Saturnino Luz. "Artificial Intelli... Crossref	<1%
55	University of Lincoln on 2021-05-19 Submitted works	<1%
56	University of Piraeus on 2022-09-29 Submitted works	<1%

57	University of Sydney on 2022-06-24 Submitted works	<1%
58	drops.dagstuhl.de Internet	<1%
59	ejournal.uin-malang.ac.id Internet	<1%
60	ira.lib.polyu.edu.hk Internet	<1%
61	journals.lww.com Internet	<1%
62	link.springer.com Internet	<1%
63	mitofit.org Internet	<1%
64	nrl.northumbria.ac.uk Internet	<1%
65	repository.tudelft.nl Internet	<1%
66	ru.scribd.com Internet	<1%
67	biorxiv.org Internet	<1%
68	igi-global.com Internet	<1%

69	nature.com Internet	<1%
70	wjgnet.com Internet	<1%