International Journal of Advance Interdisciplinary Research
Vol. 2, Issue 3, Part A, (July-Sept) 2026, pp. 203-222, e-ISSN: 3107-913X
DOI: https://doi.org/10.66095/ijair.2026.v2.i3.a.11

The pursuit and challenges of Machine learning in clinical trials

Manisha Sharma[1], Dr. Sushil Kumar Shukla[2]
[1]Assistant Professor, Department of Clinical Research, Himalayan Institute of Medical Sciences, Swami Ram Himalayan University, Swami Ram Nagar, Dehradun
[2] Professor, Department of Community Medicine, U P University of Medical Sciences, Saifai Etawah, U. P
Corresponding author: E-mail: manisha.october@gmail.com

Abstract-

There has been a significant upward push in artificial intelligence and system mastery. Drug, device, and treatment protocol testing has traditionally relied on trial data. Timelines, trial expenses, and compliance are common trial issues that have historically affected the volume of data gathered. The human workload has decreased, allowing people to live extraordinary lives. Researchers are now able to access, manage, and analyze data more quickly, affordably, and legally than ever before because of machine learning (ML and artificial intelligence (AI).

Clinical development is about to undergo a major transformation due to the rapid integration of massive novel digital data sources, the computing power to locate clinically relevant trends in the data using effective artificial intelligence and machine-learning algorithms, and the regulators' acceptance of this change through new partnerships. In order to speed up the research process and evaluate the risk and cost of clinical trials, machine, learning techniques and tools are utilised at every stage of development. This review offers relevant literature and drug discovery using these methods. The primary source of poor interpretability results that may limit potential uses and drug development is machine learning difficulties. Qualitative and methodological data needs to be collected in order to address a number of issues in clinical trials, such as confirming machine learning methods, improving decision-making, increasing awareness of strategies for machine learning, and detecting risk inadequacies in drug discovery.

This article discusses the difficulties in clinical research as well as the use of AI and system learning to enhance drug development.

Keywords- Machine learning, Clinical trials, Artificial intelligence, AI, ML

Introduction

International Journal of Advance Interdisciplinary Research
Vol. 2, Issue 3, Part A, (July-Sept) 2026, pp. 203-222, e-ISSN: 3107-913X
DOI: https://doi.org/10.66095/ijair.2026.v2.i3.a.11

Over the past 30 years, clinical drug development has mostly not altered. [1] In addition to uncertainty about regulation requirements, risk, versions, and predictability regarding emerging but largely unproven technologies (such as wireless health, monitoring devices, machine, learning, and sensor), there is lack of relevant, practical biomedical data sources, and advanced analytics for developing hypothesis that could encourage the intervention of intriguing diagnostics treatments.

Machine learning is a branch of artificial intelligence that builds classification as well as forecasting models from training data by identifying and comprehending connections in high-dimensional data using computer-based models. During the 1950s, Arthur Samuel, an IBM employee, coined the term "machine learning." Since then, machine learning has come a long way. Other branches of machine learning include incentive learning and both supervised and unsupervised learning. [2, 3]

Finding patterns in unorganized data is the goal of unsupervised machine learning. By doing this, it can automatically find groups of related instances in a dataset. Following their identification, these clusters might be further examined or displayed, for instance, by using common statistics. Questions like "Are there various kinds of a disease?" or “To what extent does a patient with a medical condition differ from normal subjects?" can be especially well-served by unsupervised machine learning techniques. K-nearest-neighbors, principal component analysis, and variation auto encoders—an unsupervised deep learning architecture—are examples of popular unsupervised machine learning methods. [3, 4] Since digital medical records have become widely recognized since automation, it may be possible to better understand illness if scientific facts are converted into clinical knowledge. [5]

Conversely, supervised machine learning techniques use labeled data (such as outcome scores or fit vs. sick) to look for patterns in multi-dimensional data. More specifically, a set of training data with ground truth labels is usually used to construct a model and improve the machine learning models, effectiveness for intended result. [3, 4]

Despite the fact that machine learning has been studied since mid 1900s, its application in healthcare has grown as a result of improvements in processing, power, data, accessibility, creative methods, and growing body of technological know-how. [6] Models have been investigated for use in area, such as improved cancer detection, ER try, and ICU decision support. [7, 8, 9]

But previous failures to properly apply machine learning systems in medical facilities have brought attention to the technology's shortcomings, leading to a lack of confidence in its ability to change medicine. [10, 11]

Despite the fact that pharmaceutical companies have long recognised the high cost and hazards associated, clinical trial data is generally accepted and is expected to continue to be the goal standard for developing safe and effective medication. [12, 13] However, the application of computational techniques to make robots capable of sensing, thinking, learning, and making decisions is known as artificial intelligence (AI). Technological developments are pushing the development of new types of artificial intelligence, which are expanding quickly in a variety of fields, including financial strategy, face recognition, and self driving cars. [14, 15]

Numerous applications in Clinical trials are being considered and are starting to be investigated in practice as AI is acknowledged as a means of achieving long-term, optimized drug development. Large and complicated amounts of classified and uncategorized clinical, molecular, and imaging data are made available to medical research through the rise and extension of randomized trials. With the goal to effectively accelerate and simplify all various tasks throughout drug research, extensive AI models that have been created and trained with appropriate data set are needed to produce actionable insights from the knowledge that is accessible, even though data access is crucial for trends in information driven and personalized medicine. [13, 16] These days, AI and gadgets have been used in medicine in a variety of ways, such as helping to detect outbreaks of contemporary infectious diseases that could affect public health; integrating genetic, medical, and numerous laboratory results to detect uncommon diseases, commercial activity [17].

Exploring Machine Learning in Drug Discovery

The effectiveness of machine learning approaches in the pharmaceutical industry's development of various medications has been assessed. At this stage, improvements in dataset size and kinds can serve as the foundation for machine learning provided unrestricted storage is offered. [18] As a result, the field of AI has evolved from theoretical understanding to empirical facts. For usage in computers, such as Graphical Processing Unit (GPU), which accelerate processing (i.e. computational processes), data was greatly improved. One machine learning method that was developed lately is the deep learning model.

[19]. It creates models to increase performance in difficulties that occur in broad daylight. [20, 21]

Finding therapeutic targets and producing candidate compounds can be made easier and more successful with the use of machine learning. A machine learning approach may be able to find some sorts of gene illness, relationships from big data sets, even when there are a few pertinent data points, despite the fact that Jia et al. have established drug gene mutation correlations from scientific journals. [22] These methods and other efforts to make extraordinarily vast volumes of biological data. Human understandable assist the researchers identify promising topic for future study by using and avoiding duplicate earlier work. [23] Once pertinent studies subjects have been identified, machine learning (ML), optimizes molecules inside a specific biological system using a gad graph, neural network to produce possible candidate compounds. The use of machine learning algorithms has grown significantly in pharmaceutical companies during the last two years. New drugs were the mainstay of the clinical field when it came to creating a new treatment for chronic illnesses. Recently, a number of medications have been developed to identify active ingredients in conventional therapies like penicillin. Natural materials and tiny chemicals that help in therapeutic treatment identify things like cells or whole organisms are found in chemical laboratories. We refer to this process as "old-style pharmacology."

The human genome has made it possible to apply cleaning techniques and improve protein processing in large quantities, which has frequently led to an increase in high throughput screening, utilizing numerous libraries. The method of altering a disease by employing biological targets to screen for large molecules is known as reverse pharmacology. Sufficiency tests have also been conducted on animals, and screening activities can provide a large number of findings to provide cells. These days, finding hits through screening is a part of drug research, and optimization techniques can improve a drug's efficacy, effectiveness, and metabolic stability.

Clinical studies will be conducted to generate a specific medication if the chemical meets all standards. Target identification and validation, lead optimization, clinical trials, and hit discovery are all essential phases in the drug development and discovery process. [24]

Computational intelligence offers a variety of methods for learning and analysis in drug development. It also describes how AI was used to identify a pharmaceutical company and locate several medications in an organized and cohesive manner. Because of this, a lot of pharmaceutical corporations are more willing to give to technology and resources that can assist recover trustworthy drug discovery results. [25]

International Journal of Advance Interdisciplinary Research
Vol. 2, Issue 3, Part A, (July-Sept) 2026, pp. 203-222, e-ISSN: 3107-913X
DOI: https://doi.org/10.66095/ijair.2026.v2.i3.a.11

- *The function of machine learning in managing clinical trial participants*

Target patient population choosing, patient recruitment, and participant attrition are all included in clinical trial participant management. However, patient dropout and failing to adhere frequently result in trials going over budget or time limits or failing to provide useful data, even when substantial resources—such as time, preparation, and trial coordinator effort—are typically allocated to participant management. Given that between 33.6% and 52.4% of phase 1 to 3 clinical trial supporting development do not proceed to the next trial phase, the actual chance that a medication evaluated in phase 1 will be authorised is 13.8%. [26] By carefully choosing specific patient populations for research, it would be able to lower the sample size required to identify a significant impact. Stated differently, fewer people may be exposed to treatments that are doubtful to be helpful to them if the patient group is better chosen. [27]

In addition to making patient population selection easier by facilitating the rapid analysis of sizeable databases of prior research, unsupervised machine learning applied to patient populations can identify trends in particular traits that can be used to select a patient phenotypes that are most likely to benefit from the recommended medication or intervention (as previously mentioned). [28] Phenotyping and the creation of representative cohorts need the consideration of a structured data; nevertheless, the identification of robust and meaningful cohort requires the concentration of additional patient data [29].

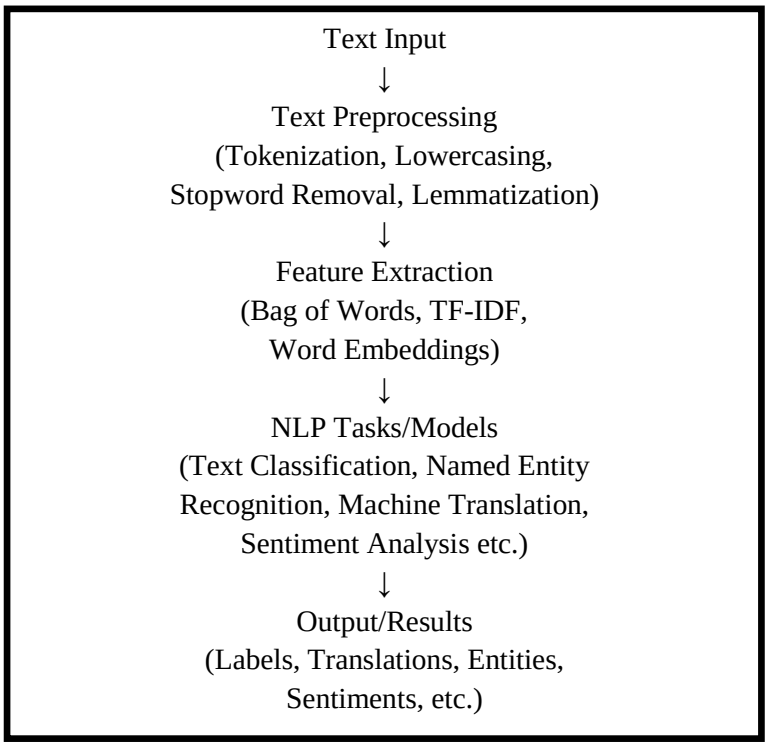
For example, three distinct sub groups of type II diabetes with various phenotypic presentations were identified by unsupervised learning of genetic data from 11,210 individuals and data from electronic health records. These populations could respond differently to a recommended treatment and have different requirements. [30].

In attempt to leverage the benefits of targeted patient, population selection, the start-up Bullfrog AI has analyzed clinical trial datasets "to anticipate which individuals will receive advantages from particular medications in development, thus enhancing inclusion/exclusion criteria and ensuring the main results are achieved" [31].

Natural language processing (NLP), which is often a labor-intensive procedure, is showing promise in identifying patients who fit the target phenotype once the particular cohort has been chosen. For instance, by concurrently merging enrolment criteria (text) and patient records (tabular data) into a distinct, hidden space, cross model, inference, learning, model

International Journal of Advance Interdisciplinary Research
Vol. 2, Issue 3, Part A, (July-Sept) 2026, pp. 203-222, e-ISSN: 3107-913X
DOI: https://doi.org/10.66095/ijair.2026.v2.i3.a.11

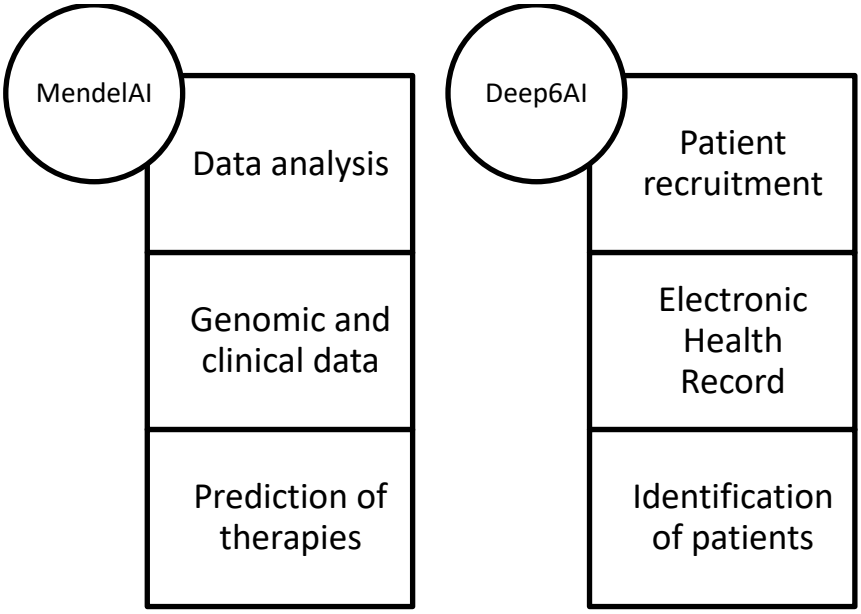
algorithm, and trials utilizing HR data are significantly more effective than earlier machine learning approaches. [32].



A small descriptive flowchart for Natural Language Processing (NLP)

The usefulness of MendelAI and Deep6AI, two business enterprises offering similar services, is questioned because to the dearth of peer-reviewed data regarding their growth and achievement of metrics [33, 34]. There are two main ways to use ML-assisted techniques to increase protocol adherence and participant engagement. The first is to determine and deal with individuals who are at high risk of non-compliance with the study by using machine learning (ML) to gather and evaluate vast volumes of data. The second strategy is to employ machine learning (ML) to lessen the study load of participants and enhance their experiences.

International Journal of Advance Interdisciplinary Research
Vol. 2, Issue 3, Part A, (July-Sept) 2026, pp. 203-222, e-ISSN: 3107-913X
DOI: https://doi.org/10.66095/ijair.2026.v2.i3.a.11



The difference between MendelAI and Deep6AI explain via a flow chart

AiCure is a for-profit business that ensures patients take their medications as prescribed by using facial recognition technology. In both an anticoagulation study, and a Schizophrenia research, it was demonstrated that AICure was more effective than a modified instantaneously observed treatment technique at detecting and improving compliance with medications in those who had previously experienced a stroke. [35, 36]

Wearable device patient data has demonstrated the ability to recognize patient falls differentiate between neuropsychiatric symptomatology patterns, and link participant engagement with the International Parkinson and Movement Disorders Society Unified Parkinson's Disease Rating Scale. [37, 38, 39].

- *Using machine learning techniques in precision medicine clinical trials*

Patients are known to react significantly to the same therapy. [40] The need to choose the best course of action for individual patient has led to the emergence of precision medicine, sometimes referred to as customized medicine, as a fast growing discipline. [41]

A key component of precision medicines, resurgence is genomics, the study of an individual's genome and how it interact with their environment, and genetics, the study of

genes and how they affect heredity. As consequence, genetics and genomics accounted for a large portion of precision medicine's the original priority. In fact, the domains of genomics and genetics provide a great deal of the data that is now accessible for precision medicine. [42]

There are a growing variety of genomic tests accessible in practice, including genomic diagnostics for pediatric and uncommon illnesses, non-invasive prenatal screening, and tumor genome sequencing for targeted treatment. [43] For instance, genome-wide association studies (GWAS) which analyze genomic data to find variations linked to a trait or phenotype, have made substantial use of machine learning. Wei et al. used GWAS data to train and evaluate a support vector machine model to detect people with Type 1 Diabetes. Based on their research, they concluded that diseases where a considerable risk proportion is gathered by single nucleotide polymorphism (SNP) array would be suitable for a genotype-based assesement of disease risk. [44]

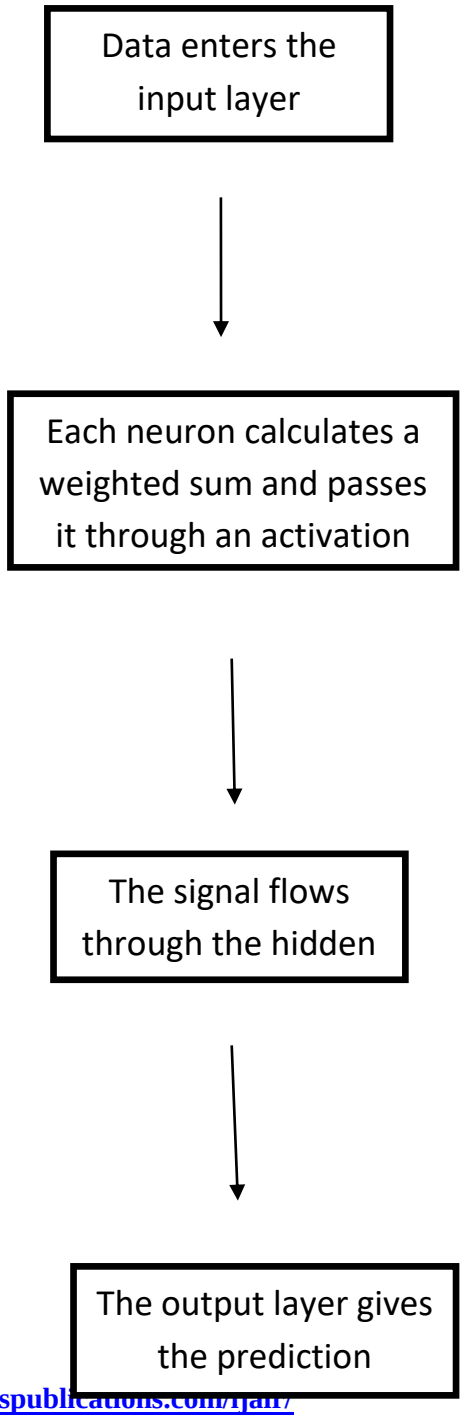
- *Apart from in-person trials: Machine Learning and Simulation*

Due to the cost and time involved in conducting clinical trials, many researchers are now focusing on alternate study designs to validate novel treatments and diagnostics. New approaches are currently being developed to find patient-effective therapies without setting up prospective trials, thanks to the growing availability of extensive health databases. Additionally, regulatory agencies are realising more and more how important it is to support the scientific evidence of a drug's efficacy with real world data in addition to data from clinical studies. [45]. However, the same fundamental problems with data collecting that affect traditional statistical studies also affect ML models, including uneven data quality, selection bias, and confounding [46, 47, 48]

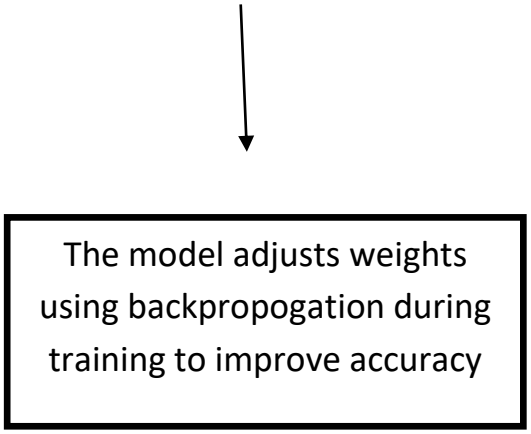
Natural Language Processing [NLP] and other machine learning technologies might need advancement. Cohort and retrospective data sets that met the eligibility requirements for patient receiving treatment in target trials for identified as part of these efforts. [49] in clinical studies where control group cannot be realistically recruited because of a sex trial for a small number of participants, such as in uncommon illnesses, synthetic cohorts, provide good support for some conclusions, even though more work is probably needed to ensure comparison throughout the life and simulated study groups. [50, 51].

International Journal of Advance Interdisciplinary Research
Vol. 2, Issue 3, Part A, (July-Sept) 2026, pp. 203-222, e-ISSN: 3107-913X
DOI: https://doi.org/10.66095/ijair.2026.v2.i3.a.11

Last but not least, several research teams have already demonstrated that neural networks may discover common patterns of traits (i.e., representations) among individuals undergoing various types of therapy [52]. By simulating clinical results when a certain therapy is added or removed, these systems may be utilized to quantify the benefits of various treatments after improving the detection of similarities across patients in distinct treatment arms [53, 54]. In essence, these methods produce virtual avatars, or "digital twins," of specific patients that can afterwards be safely exposed to experimental treatments [55, 56, 57]



International Journal of Advance Interdisciplinary Research
Vol. 2, Issue 3, Part A, (July-Sept) 2026, pp. 203-222, e-ISSN: 3107-913X
DOI: https://doi.org/10.66095/ijair.2026.v2.i3.a.11



A flowchart describing the functioning of neural networks

- *Using Digital Biomarkers, Therapeutic Software, and Remote Monitoring to Revolutionize Trial Design*

Machine Learning may potentially be used to improve clinical trial efficacy by mitigating several of the disadvantages of traditional, centralized research approach. Machine learning facilitates trial administration in a number of ways. MyStudies is a smartphone app created by the US Food and Drug Administration (USFDA) to assist in informing people during a coronavirus outbreak [58].

It is possible to increase system reliability by teaching picture classification in every effort to verify the authenticity of patient signs.

Comparable methods have been used to verify medication compliance in patient groups, including those dealing with mental illness or drug abuse. For instance, facial recognition technology has been used by AiCure, an analytics startup that specializes in remote clinical trial assistance, to verify that patients suffering from opioid addiction are following their prescribed treatment schedules [59].

Ingestable sensors have also been used by Tokyo-based Otsuka Pharmaceuticals to track antipsychotic medication intake in schizophrenic patients [60]. By automatically identifying risky changes in a participant's condition, remote monitoring of variables like blood

chemistry and vital signs may also help discover adverse events in clinical trials early [61]. Given that machine learning system can detect anomalies in continuous data, software programs that use variable or implanted sensors to learn patient's unique physiological patterns may open the door for customize warning system during experimental drug trials. [62].

In order to track treatment outcomes, machine learning algorithms can also discover completely new patterns in routinely obtained data, creating a new class of digital biomarkers [63, 65]. In order to optimize the quantity of information utilized for decision-making, automated systems can be taught to recognize these indicators from a single data source (such as an ECG) or from interactions of many modalities (such as blood glucose, skin conductance, and pulse oximetry). Furthermore, physiological cues or digital indicators of actual function, such the measurement of mobility using wearable sensors, may eventually operate as clinical efficacy goals in and of themselves [65, 66].

Last but not least, there is growing evidence that new digital technologies could serve as remedies in and of themselves rather than only aiding in the creation of conventional medications and equipment. Prescription video games and smartphone apps are examples of such "digital therapeutics" [67] that are now being developed to treat a variety of illnesses, including multiple sclerosis (MS), addiction, ADHD, and psychosis [68].

Clinical research's operational obstacles to machine learning

Designing machine, learning algorithms and implementing them for use in critical research involves a number of procedures and discipline. The inaugural episode attempted to put together a team with the necessary machine learning and clinical knowledge to be successful. The danger of constructing a model that uses machine learning approaches that aren't suitable for the current data and research challenges in internal settings increases if such a team is not formed and communication within the team is maintained. [69] For example, if a mortality prediction algorithm is developed without contacting any clinical team members, intubation may be a sign of death. Although this prognosis is undoubtedly correct, it probably has little therapeutic value. Teams, clinical scientists, and data scientist all benefit from collaboration, but for it to completely operationalize, it may need more upfront, disciplinary training, openness, and trust.

For both validation and training, a lot of data is needed from any modern machine, learning models, particularly deep neural networks. Since several recorded examples have

demonstrated that algorithms performed better on invalidation, data sets and training, data sets, developers should use different data sources throughout this procedure, insured generalisability outside of the training dataset. [70] Holders of the information may be reluctant to provide it since clinical research data is frequently patient related or produced at a high expense by organizations (for EHR data) or companies (for clinical trial data). Interoperability may be hampered by various approaches, even while all parties are willing to exchange data. Recent data set such as MIMIC, UK, Biobank and eICU when other reliable data cannot be found, has resources. [71, 72, 73].

However, as training retrospective data (like MIMIC and UK Biobank) rarely transfer smoothly to future applications, any data sources are needed to produce a model that is simply usable. For instance, Gong et al. demonstrated techniques for feature consolidation during significant temporal shift, such EHR transitions, whereas Nestor et al. Highlight the significance of taking Year of care into account in MIMIC, owing to temporal drift. [70, 74].

Furthermore, it is less probable than data produced for clinical therapy would accurately reflect certain patient types and illness stages. For example, even though MIMIC is widely used since it is publicly available, population of intensive care unit cannot be applied to any applications beyond critical care. Achieving repeatability and reproduce ability in ML driven Clinical research is more challenging due to several factors, including computation, model replication , and the importance of variability in many ML techniques. [75] These factors are more closely linked to problems with data quality and availability. Even if a few factors and structures cause issues during the training phase, several machine learning techniques yielded accurate results. The specific method is unable to achieve the required precision and local optimum, particularly when there is insufficient data available during the training phase. [76]

Another difficulty in drug discovery is the transparency problem due to the ambiguity of decision-making in various classification methods. To assess the outcomes, it is necessary to comprehend a number of medication development procedures. Accordingly, it is more helpful in finding novel drug targets, and several factors must be combined to increase interpretability trust (Vamathevan et al., 2019). [77]

Explainability is the idea that algorithm consumers ought to be able grasp the mechanisms that underlie algorithmic output. Techniques to do this, such as maps of salience and attention ratings, have been the subject of much research; yet, questions remain regarding their applicability and performance [78–81]. Despite being a compelling idea, there is much

disagreement about whether explainability actually hinders machine learning's potential to advance clinical practice and research.

Explainability might cause researchers to place undue faith in models that have serious flaws. Rather, the supporters of this approach emphasize reliability. A number of elements of clinical medicine and clinical research, including laboratory, assessment, the full mechanism of medication, the statistical test are not very good, but they are still used because they have been proven to work well and consistently, even though many and uses are still unclear about how or why. [82]. Since the General Data Protection Regulation of the European Union was passed, this conceptual barrier has also turned into an operational one, requiring computational decision-making algorithms to disclose "meaningful details concerning the logic engaged."

Conclusion-

The conventional limits of clinical trials will be expanded as medicine develops in the contemporary era and attempts are made to extract useful insights from medical data. The integration of machine learning technology with the richness of information present in established healthcare settings will necessitate consideration of both privacy and data security.

If ML enabled clinical research can improve the efficacy and quality of biological evidence, there is an ethical obligation to explore the possibility that it may reduce suffering and save lives. It will be necessary to overcome obstacles related to data structure and access, result definitions, establishment and verification process transparency, certification reliability, and possible bias in order to realize this potential. With long-term advantages for patients, physicians, and society at large, responsible growth in this field has the possibility to accelerate scientific discovery.

References-

1. FDA. Novel Drug Approvals for 2018. https://www.fda.gov/drugs/ developmentapprovalprocess/druginnovation/ucm592464.htm (2018).

2. John M. and Edward A.F. 1990. In Memoriam: Arthur Samuel: Pioneer in machine learning. AI Magazine. 11(3): 10–11.

3. Zou J., Huss M., Abid A., Mohammadi P., Torkamani A., and Telenti A. 2019. A primer on deep learning in genomics. Nat. Genet. 51(1): 12–18.

4. Husi, H. 2019. Computational Biology. Codon Publications. Brisbane (AU).

5. Kononenko, Igor. Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in medicine 23.1 (2001): 89-109.

6. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, et al. Improved protein structure prediction using potentials from deep learning. Nature. 2020;577(7792):706–10.

7. Yue W, Wang Z, Chen H, Payne A, Liu X. Machine learning with applications in breast cancer diagnosis and prognosis. Designs. 2018;2(2):13. doi:10.3390/designs2020013

8. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA Jr, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Crit Care. 2019;23(1):64. doi:10.1186/s13054-019-2351-7

9. Johnson AEW, Ghassemi MM, Nemati S, Niehaus KE, Clifton DA, Clifford GD. Machine learning and decision support in critical care. Proc IEEE Inst Electr Electron Eng. 2016;104(2):444-466. doi:10.1109/JPROC.2015. 2501978

10. Asan O, Bayrak AE, Choudhury A. Artificial intelligence and human trust in healthcare: focus on clinicians.J Med internet Res. 2020;22(6):e15154. doi:10.2196/15154

11. Wilkinson J, Arnold KF, Murray EJ, et al. Time to reality check the promises of machine learning–powered precision medicine. Lancet Digit Health. 2020;2(12):e677-e680. doi:10.1016/S2589-7500(20)30200-4.

12. C K, KW J, WW T. How artificial intelligence could redefine clinical trials in cardiovascular medicine: lessons learned from oncology. Personalized Med. 2019;16(2):87–92. https://doi. org/10.2217/PME-2018-0130.

13. S H, P S, B A, J H. Artificial Intelligence for Clinical Trial Design. Trends in pharmacological sciences. 2019;40(8):577–591.

doi:https://doi.org/10.1016/J.TIPS.2019.05.005.

14. Goodfellow I, Bengio Y, Courville A. Deep learning. MIT press; 2016.

15. Mohri M, Rostamizadeh A, Talwalkar A. Foundations of machine learning. MIT press; 2018.

16. Delso G, Cirillo D, Kaggie JD, Valencia A, Metser U, VeitHaibach P. How to design AI-Driven clinical trials in Nuclear Medicine. Semin Nucl Med. 2021;51(2):112–9. https://doi.org/10.1053/J.SEMNUCLMED.2020.09.003.

17. Cruz Rivera S, Liu X, Chan A-W, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORTAI Working Group. AI and Machine learning programme in research: the SPIRIT-AI extension. Lancet Digit Health 2020;2(10):e549-e560.

18. Mamoshina P, Volosnikova M, Ozerov IV, Putin E, Skibina E, Cortese F, Zhavoronkov A (2018) Machine learning on human muscle transcriptomic data for biomarker discovery and tissue-specifc drug target identifcation. Front Genet 9:242.

19. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444.

20. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T (2018) The rise of deep learning in drug discovery. Drug Discovery Today 23(6):1241–12.

21. Hinton G (2018) Deep learning–a technology with the potential to transform health care. JAMA 320(11):1101–1102.

22. L. Fiore and P. Lavori, "Integrating Randomized Comparative Effectiveness Research with Patient Care", New England Journal of Medicine, vol. 374, no. 22, pp. 2152-2158, 2016.

23. A. Hernandez, R. Fleurence and R. Rothman, "The ADAPTABLE Trial and PCORnet: Shining Light on a New Research Paradigm", Annals of Internal Medicine, vol. 163, no. 8, pp. 635-636, 2015.

24. Vohora D, Singh G (2018) Pharmaceutical medicine and translational clinical research. Academic Press, London

25. Duch W, Swaminathan K, Meller J (2007) Artifcial intelligence approaches for rational drug design and discovery. Curr Pharm Des 13(14):1497–1508.

26. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. Biostatistics. 2019;20(2):273–86. https://doi.org/10.1093/ biostatistics/kxx069.

27. Schork NJ. Personalized medicine: time for one-person trials. Nature. 2015; 520(7549):609–11. https://doi.org/10.1038/520609a.

28. Glicksberg BS, Miotto R, Johnson KW, Shameer K, Li L, Chen R, et al. Automated disease cohort selection using word embeddings from electronic health records. Pac Symp Biocomput. 2018;23:145–56.

29. Liao KP, Cai T, Savova GK, Murphy SN, Karlson EW, Ananthakrishnan AN, et al. Development of phenotype algorithms using electronic medical records and incorporating natural language processing. BMJ. 2015;350(apr24 11):h1885. https://doi.org/10.1136/bmj.h1885.

30. Li L, Cheng WY, Glicksberg BS, Gottesman O, Tamler R, Chen R, et al. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. Sci Transl Med. 2015;7(311):311ra174.

31. Our Solution 2021 [cited 2021 February 2]. Available from: https://www. bullfrogai.com/our-solution/.

32. Zhang X, Xiao C, Glass LM, Sun J. DeepEnroll: patient-trial matching with deep embedding and entailment prediction. In: Proceedings of the Web Conference 2020. Taipei: Association for Computing Machinery; 2020. p. 1029–37.

33. Calaprice-Whitty D, Galil K, Salloum W, Zariv A, Jimenez B. Improving clinical trial participant prescreening with artificial intelligence (AI): a comparison of the results of AI-assisted vs standard methods in 3 oncology trials. Ther Innov Regul Sci. 2020;54(1):69–74. https://doi.org/1 0.1007/s43441-019-00030-4.

34. How it works 2019 [cited 2021 February 2]. Available from: https://deep6.ai/ how-it-works/.

35. Bain EE, Shafner L, Walling DP, Othman AA, Chuang-Stein C, Hinkle J, et al. Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a phase 2 clinical trial in subjects with schizophrenia. JMIR Mhealth Uhealth. 2017;5(2):e18. https://doi.org/10.2196/mhealth.7030.

36. Labovitz DL, Shafner L, Reyes Gil M, Virmani D, Hanina A. Using artificial intelligence to reduce the risk of nonadherence in patients on anticoagulation therapy. Stroke. 2017;48(5):1416–9. https://doi.org/10.1161/STROKEAHA.116.016281.

37. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela JP. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology. 2018;43(8):1660–6. https://doi.org/10.1038/s41386-018-0030-z.

38. Chaudhuri S, Oudejans D, Thompson HJ, Demiris G. Real-world accuracy and use of a wearable fall detection device by older adults. J Am Geriatr Soc. 2015;63(11):2415–6. https://doi.org/10.1111/jgs.13804.

39. Chen R, Jankovic F, Marinsek N, Foschini L, Kourtis L, Signorini A, et al. Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage: Association for Computing Machinery; 2019. p. 2145–55.

40. Antoniou M, Kolamunnage-Dona R, Jorgensen AL. Biomarker-guided non-adaptive trial designs in phase II and Phase III: a methodological review. JPM 2017; 7 (1): 1.

41. Ginsburg GS, Phillips KA. Precision medicine: from science to value. Health Aff (Millwood) 2018; 37 (5): 694–701.

42. Grainger D. 2016. The multi-omics revolution. J. Pers. Med. 81–86.

43. Khoury MJ, Gwinn M, Bowen MS, Dotson WD. Beyond base pairs to bedside: a population perspective on how genomics can improve health. Am J Public Health. 2012;102(1):34–7. https://doi.org/10.2105/AJPH.2011.300299

44. Wei Z., Wang K., Qu H.-Q., Zhang H., Bradfield J., Kim C., et al. 2009. From disease association to risk assessment: an optimistic view from genome-wide association studies on type 1 diabetes. PLoS Genet. 5(10): e1000678

45. Burns L, et al. Real-world evidence for regulatory decision-making: guidance from around the world. Clin Ther. 2022;44:420–37.

46. Pearl J. Causality. Cambridge university press; 2009.

47. Pearl J, Mackenzie D. The book of why: the new science of cause and efect. Basic books; 2018.

48. Larrouquere L, Giai J, Cracowski JL, Bailly S, Roustit M. Externally controlled trials: are we there yet? Clin Pharmacol Ther. 2020;108:918–9.

49. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol. 2016;183:758–64.

50. Miksad RA, Abernethy AP. Harnessing the power of real-world evidence (RWE): a checklist to ensure regulatory-grade data quality. Clin Pharmacol Ther. 2018;103:202–5.

51. Thorlund K, Dron L, Park JJ, Mills EJ. Synthetic and external controls in clinical trials–a primer for researchers. Clin Epidemiol. 2020; 457–67.

52. Yao L. et al. Representation learning for treatment efect estimation from observational data. Adv Neural Inf Proces Syst. 2018;31.

53. Johansson F, Shalit U, Sontag D. In International conference on machine learning. 3020–29 (PMLR).

54. Shalit U, Johansson FD, Sontag D. In International Conference on Machine Learning. 3076–85 (PMLR).

55. San O. The digital twin revolution. Nature Computational Science. 2021;1:307–8.

56. Björnsson B, et al. Digital twins to personalize medicine. Genome medicine. 2020;12:1–4.

57. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. Nat Med. 2022;28:1773–84.

58. Wyner Z, et al. The FDA MyStudies app: a reusable platform for distributed clinical trials and real-world evidence studies. JAMIA open. 2020;3:500–5.

59. Beaulieu T, Knight R, Nolan S, Quick O, Ti L. Artifcial intelligence interventions focused on opioid use disorders: a review of the gray literature. Am J Drug Alcohol Abuse. 2021;47:26–42.

60. Waltz E. Drugs go wireless. Nat Biotechnol. 2016;34:15–9.

61. Kadakia KT, et al. Virtual clinical trials in oncology—overview, challenges, policy considerations, and future directions. JCO Clin Cancer Inform. 2021;4:421–5.

62. Teng M. In 2010 IEEE International Conference on Progress in Informatics and Computing. 603–08 (IEEE).

63. Au R, Kolachalama VB, Paschalidis IC. Redefning and validating digital biomarkers as fuid, dynamic multi-dimensional digital signal patterns. Front Digit Health. 2022;3:208.

64. Bent B, et al. The digital biomarker discovery pipeline: an open-source software platform for the development of digital biomarkers using mHealth and wearables data. J Clin Transl Sci. 2021;5:e19.

65. Ricotti V, et al. Wearable full-body motion tracking of activities of daily living predicts disease trajectory in Duchenne muscular dystrophy. Nat Med. 2023; 1–9.

66. Servais L, et al. Stride velocity 95th centile: insights into gaining regulatory qualifcation of the frst wearable-derived digital endpoint for use in Duchenne muscular dystrophy trials. J Neuromuscul Dis. 2022;9:335–46.

67. Sim I. Mobile devices and health. N Engl J Med. 2019;381:956–68.

68. Kennedy C. Pear approval signals FDA readiness for digital treatments. Nat Biotechnol. 2018;36.

69. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med. 2019;25(9):1337–40. https://doi.org/10.1038/s41591-019-0548-6 Epub 2019/ 08/21. PubMed PMID: 31427808.

70. Nestor B, McDermott M, Chauhan G, et al. Rethinking clinical prediction: why machine learning must consider year of care and feature aggregation. arXiv preprint 2018;arXiv:181112583.

71. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data. 2016;3(1):160035. https://doi.org/10.1038/sdata.2016.35.

72. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Sci Data. 2018;5(1):180178. https://doi.org/10.1038/sda ta.2018.178.

73. UK Biobank. www.ukbiobank.ac.uk. Accessed 22 Mar 2021.

74. Gong JJ, Naumann T, Szolovits P, Guttag JV. Predicting clinical outcomes across changing electronic health record systems. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Halifax: Association for Computing Machinery; 2017. p. 1497–505.

75. Beam AL, Manrai AK, Ghassemi M. Challenges to the reproducibility of machine learning models in health care. JAMA. 2020;323(4):305–6. https:// doi.org/10.1001/jama.2019.20866.

76. Ghasemi F, Mehridehnavi A, Fassihi A, Pérez-Sánchez H (2018) Deep neural network in qsar studies using deep belief network. Appl Soft Comput 62:251–258.

77. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, Li B, Madabhushi A, Shah P, Spitzer M et al (2019) Applications of machine learning in drug discovery and development. Nat Rev Drug Discovery 18(6):463–477

78. Adebayo J, Gilmer J, Muelly M, Goodfellow I, Hardt M, Kim B. Sanity checks for saliency maps. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montréal: Curran Associates Inc.; 2018. p. 9525–36.

79. Wiegreffe S, Pinter Y. Attention is not not explanation. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Hong Kong, China: Association for Computational Linguistics; 2019.

80. Jain S, Wallace BC. Attention is not explanation: NAACL-HLT; 2019.

81. Serrano S, Smith NA. Is attention interpretable? In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 2931– 2951, Stroudsburg, PA, USA, 2019. Association for Computational Linguistics.

82. Sendak M, Elish MC, Gao M, Futoma J, Ratliff W, Nichols M, et al. “The human body is a black box”: supporting clinical decision-making with deep learning. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. Barcelona: Association for Computing Machinery; 2020. p. 99–109.