

**How to Cite:**

Shewale, S., Undale, V., Kawaste, A., Bhalchim, V., Shelar, M., & Gundecha, S. (2022). On the brink of transformation: Clinical research. *International Journal of Health Sciences*, 6(S3), 4020–4039. <https://doi.org/10.53730/ijhs.v6nS3.6722>

## **On the brink of transformation: Clinical research**

**Mr. Sunil Shewale**

[M. Pharm (QAT), MBA (HR & Marketing), PG-Clinical Trials]-Research scholar, Dr. D. Y. Patil Institute of Pharmaceutical Sciences & Research, University of Pune. Maharashtra (India)

**Dr. Vaishali Undale**

[M. Pharm (Pharmacology), Ph.D. (Pharmacology)]- HOD, Department of Pharmacology, Dr. D. Y. Patil Institute of Pharmaceutical Sciences, & Research, University of Pune. Maharashtra (India)  
Email: [vaishali.undale@dypvp.edu.in](mailto:vaishali.undale@dypvp.edu.in)

**Ms. Akshata Kawaste**

[B. Pharm]- Research Student, Dr. D. Y. Patil Institute of Pharmaceutical Sciences & Research, University of Pune. Maharashtra (India),

**Ms. Vrushali Bhalchim**

[M. Pharm (Pharmacology)]- Research scholar, Dr. D. Y. Patil Institute of Pharmaceutical Sciences & Research, University of Pune. Maharashtra (India),

**Dr. Maruti Shelar**

[M. Pharm (Pharmacognosy), Ph.D. (Pharmacognosy)]- Associate Professor, Dr. D. Y. Patil Institute of Pharmaceutical Sciences and Research, University of Pune. Maharashtra (India),

**Mr. Sachin Gundecha**

[M. Pharm (Pharmaceutics)]- Research scholar, Dr. D. Y. Patil Institute of Pharmaceutical Sciences & Research, University of Pune. Maharashtra (India)

**Abstract**--The research on drug development life cycle and bringing sole new drug to the market is a million dollar question for pharmaceutical organization. Any clinical trial consumes average of 10 to 15 years and USD 1.5-2.0 billion with uncertainty of medications for its effectiveness for human use. Hardly, one out of 10 compounds entering into the clinical trial that reaches to the market rendering a major loss to pharmaceutical or biotech company in case of trial failure. Conversely, with changing time and an increase in the number of medicines approved by regulatory authorities, the regulatory teams are increasing networks for monitoring and

assembling adverse event reports from varied sources. This in turn, has increases annual exponential rise in data volumes and the companies are facing a huge challenge in processing it. To meet such challenges, organizations must sharpen their ability to introduce new wearables for clinical trials and provide advanced cognitive solution to handle large and complex datasets. This has summoned concepts like Artificial Intelligence to expedite medical science and clinical trial and pharmacovigilance attain success.

**Keywords**---artificial intelligence, pharmacovigilance, clinical research, health, ADR.

## **Introduction**

The rapid and efficient movement of drug through drug pipelines poses an enormous challenge especially through time-consuming and costly clinical trial steps. Every delay in the race to market means loss of millions of dollars of revenues for a blockbuster drug (Nayak et al., 2016). More than one third of all Phase III compounds or molecules fail to advance to final approval (Wong et al., 2019). The important key factors responsible for high trial failure rates includes poor study design, insufficient sample size calculation, suboptimal patient cohort selection and recruiting techniques, and inability to monitor patients effectively during trials (Harrer et al., 2019).

In recent year, there are over 3 billion mobile health apps downloaded with most users looking for information on diseases (66%) or treatment (56%) (Sparkes, 2018). The life sciences companies are forecasting a steep surge in medical inquiries and adverse event reports (ADR). The USFDA alone receives more than 150,000 of ADRs every month showing an exponential increase of more than 200 times since 2010. Apart from such official platforms, other sources are facing a similar increase in medical information and ADRs. Though traditional sources of information viz; clinical trials, safety databases are a major part, new sources like health-related media and general social media, wearable devices, academic collaborations, medical forums are pouring in more and more data. Hence, the surge in medical information has put forward two major challenges for pharmacovigilance (PV) industry such as; extraction of Individual Case Safety Reports (ICSR)-relevant information from huge data and processing the resultant cases.

Automation has come a very long way and is well established in the finance and banking sectors since the 1950's (Fisher & McKenney, 1993). Banking is employing artificial intelligence (AI) to protect customer identity data, mimic bank employees, increase digital activities and customer engagement. Similarly, aviation industry also operates through AI and all aircraft use automated landing systems. AI has proven to be useful in the detection and prevention of fraud and in the fight against money laundering. Chatbot is well known in AI application that can stimulate a conversion with a user in a natural language through messaging applications, websites, mobile apps, and telephones. The chatbots

have found their way in various industries including healthcare, and pharmaceuticals.

The AI healthcare market will grow by approximately \$ 8 billion by 2022, rising from USD 667.1 million in 2016 (Glass et al., 2019). AI's key healthcare application program could create \$ 150 billion in annual health care economy savings by 2026 (Fred, 2017). (Figure 1) Despite the sustainable growth and varied applicability of AI, the complexity of ADR case processing, many decision points and adjustments, highly regulated and audited environment, volume and complexity of clinical data to interpret trends to improve and accelerate decision making, identifying potential molecules, finding suitable patient populations might be the reticence in the advancement. This article highlight about, why AI is required to be adapted, its applications, uses and advantages and current technology challenges for adapting AI in clinical research and pharmacovigilance field.

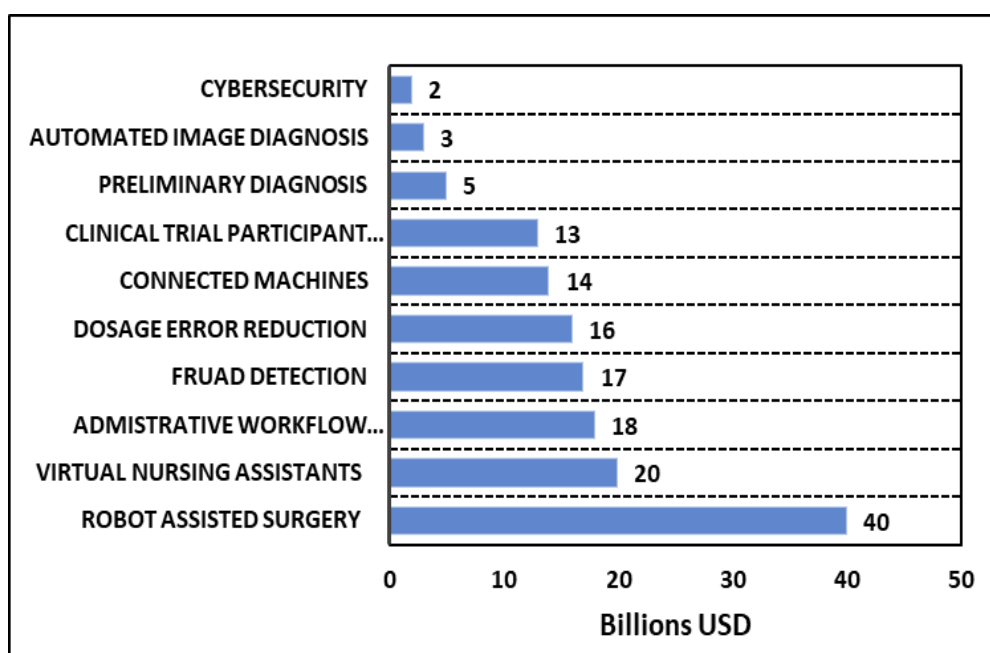


Figure 1. Top AI healthcare application with potential annual growth by 2026  
Source: <https://hitconsultant.net/2017/06/23/artificial-intelligence-healthcare-market-acn/>

### **Requirement of AI in drug discovery and clinical development phase/s**

AI is a smart technology which play with the data for fruitful outcome. It can be useful to mimic the human intelligence, uncovered the complicated patterns which are difficult to solve or beyond human capacity. This technique has broad range of application in clinical research wherein; it has a capacity to enhances the operational efficiency, reduce time and cost and provide quality benefits. (Figure 2)

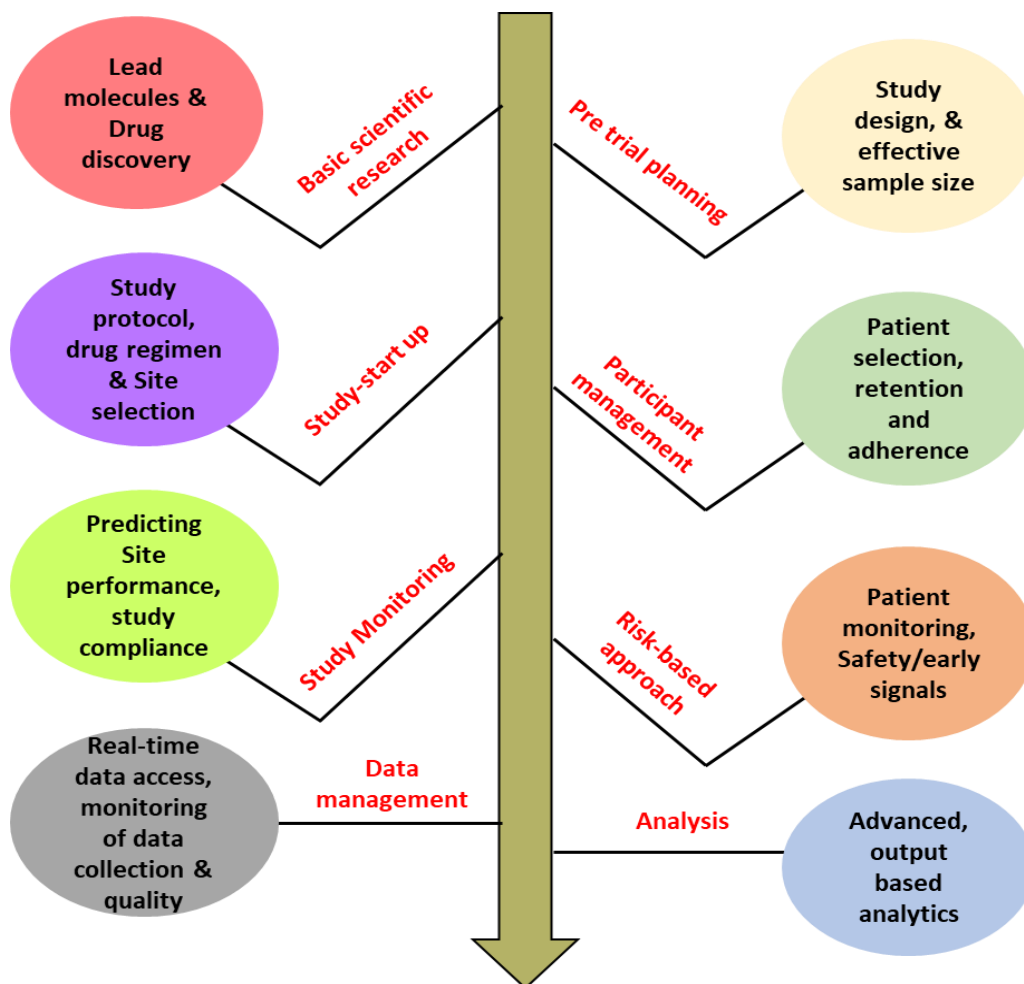


Figure 2. Applications of AI in clinical development

### Drug discovery

The discovery of new drug molecules is an important step in development. Preclinical composite detection, targeted compound testing, and interpretation of lead compounds in clinical trials can be aided by the use of AI, ML, and predictable and predictable thinking methods. This will allow for the selection of promising candidates and the elimination of others who may have a chance of failure before entering the clinical trial. For example, Amyotrophic lateral sclerosis (ALS) is a destructive neurodegenerative disease with no effective treatment and dysfunction of RNA-binding proteins (RBPs) is widely accepted as a contributing factor in ALS pathobiology. AI tool; IBM Watson® has identified additional RBP mutations in ALS that accelerate scientific discovery in ALS and other complex neurological disorders (Bakkar et al., 2018). KnIT, another system, digs for publications to identify new protein kinases containing phosphorylate protein

tumor suppressor p53. A retrospective analysis demonstrated the accuracy of this method and laboratory tests suggest that the kinases identified by the system may actually have phosphorylate p53 (Spangler et al., 2014).

### **Study design**

The study design is the backbone of any clinical trial. If the study is not design well, chances of study success are less. In case of promising study design, lot of time and cost can be saved. It might also result in, improved recruitment rates, fewer non-enrolling sites, and fewer protocol amendments. AI-enabled technologies, can support the selection of trial design through collection, organization and analysis of increasing amounts of scientific and research data, including current and past clinical trials, failed trials, patient support programmes and post-market surveillance. Machine learning, deep learning, natural language processing tools offers selection of optimal primary and secondary endpoints for the trial, which further help in selection of appropriate country and sites for recruitment, enrollment strategies, start-up planning, and execution (Harrer et al., 2019; Bhatt, 2021). Application of AI in this regard, is being further enhanced by development of guidelines involving artificial intelligence: the SPIRIT-AI Extension and CONSORT-AI Extension. These guidelines are being creating transparency and completeness for clinical trial and assist editors and peer-reviewers, as well as the general readership, to understand, interpret and critically appraise the design and risk of bias for a planned clinical trial (Rivera et al., 2020; Liu et al., 2020).

### **Finding clinical trial**

AI can identify publicly accessible web content, which includes digital trial announcements, trial sites, and social media posts, helping to match patients with available clinical trials. AI forums allow companies to market themselves and their opportunities for more effective clinical testing through digital channels such as a social media platform to engage with patients who may be enrolled or responding to a particular treatment. This will help the company to increase the number of people enrolling for clinical trials faster because it allows them to reach these people more directly than just trying to find them through conventional advertising methods which can be costly. Therefore, AI can be an excellent channel for pharmaceutical organizations to publish information about what type of treatment they are offering in a short time and at low cost.

### **Investigator and site selection**

One of the most important aspects of clinical research is to select the most effective research sites. The features of the site such as administrative procedures, access to resources, doctors with in-depth knowledge and understanding of the disease, can affect both research times and data quality and integrity. AI technology can help biopharma companies identify target areas, trained investigators, and first responders, and collect and gather evidence to satisfy regulators that the trial process complies with the requirements of Clinical Practice (Taylor et al., 2020).

## Patient recruitment

Patient recruitment is one of the challenging aspects of any trial. Many patients from different therapeutic areas gets benefited due to on-time recruitment in study; which enables patient access to cutting-edge therapies sooner. However, finding out right patient in stipulated time frame especially in complex oncology trials becomes even more difficult. Patient selection models uses patient data obtained from different sources and compare against study protocols. (Figure 3) This will help to identify eligible patient right at the time of patient's diagnosis, reduce the site team burden on pre-screening activity and increases their focus on patient care. It also helps in locating relevant available trials the patient has the potential to successfully match. The accuracy of patient eligibility for breast cancer trial using AI clinical decision support system (CDSS) was found to be 87.6% as compared to manual screening (Haddad et al., 2021). Similarly, in another study of lung cancer, AI matching system evaluated 7252 patient attributes against 11 467 individual trial eligibility criteria and reported 95.7% accuracy for exclusion and 91.6% accuracy for overall eligibility assessment (Alexander et al., 2020).

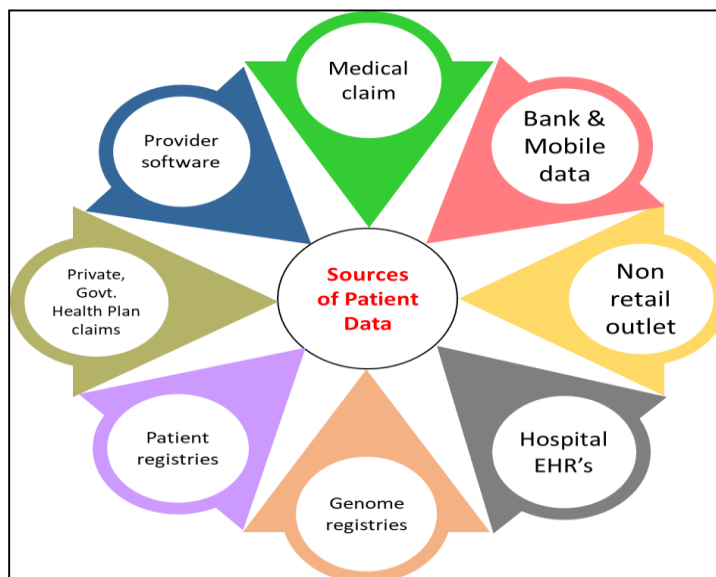


Figure 3. Potential sources of data for patient recruitment in clinical trial

## Patient engagement and adherence

Patient engagement and adherence in clinical study is painful task and its significance rises when considering virtual trials and digital health solutions. It is important that, patient should complete the study in compliance with protocol, other guidelines and regulatory requirements. Increase in patient dropouts or non-adherence in the study may cause additional recruitment to maintain statistical power of the study. This also causes exponential rise in cost on one side and trial delays on another side. Current solutions are aiming to harness technology to shorten the start-up periods, less site visits, and to make patients more involved and feel like a part of the process. Various AI based approaches

such as mobile phone applications for medication adherence, reminder systems for observations of daily livings, patient empowerment through interactive platform, and integrated care programs between health care organizations and health professionals, to improve clinical outcomes and patient ordeals are helpful to achieve success (Babel et al., 2021). Use of AI platform AiCure on mobile devices for measuring medication adherence in a Phase-II trial of the  $\alpha 7$  nicotinic receptor agonist (ABT-126) in subjects with schizophrenia demonstrated more than 17% adherence in AI group compared to control group of modified directly observed therapy (mDOT) (Bain et al., 2017).

### **Clinical monitoring**

The clinical study monitoring is tedious but a very important task performed in any trial. Frequent travels to study sites at pre-fix interval, verifying the patient records and other compliances, source data verifications (SDVs) are some of the imperative activities carried out during trial monitoring. Tremendous manual efforts are invested for verifying these processes. However, it has been observed that SDV could be less useful than originally thought and can be reduced by >90% without any measurable impact on data quality (Mitchel et al., 2011; Mitchel et al., 2014; Tantsyura et al., 2015). AI allows monitoring activities which were previously conducted on-site to be conducted remotely, increasing productive benefits such as;

- Increase efficiency and speeding up time;
- Reduce site preparation time;
- Remove the guesswork created by manual to-do lists and/or spreadsheets;
- Increase focus with document statuses and tags for quick TMF and ISF completeness analysis;
- Automate the recording of site visit activity as well as create transparent email reports that delegate actions to other team members as needed.

The innovations of AI have application for assessment of risk, non-compliance, and delivering predictive analytics. It can help in identification of high-risk sites, sites with recruitment and performance issues or even patients having high risk of potential AEs. The risk-based monitoring approach using AI application (on-site, off site, centralized monitoring) reported 28% of the few critical and major findings of each clinic quality control visit, 10% of the most important findings for each study site audit, 29.3% minimum intermediate query resolution time (number of days for closing electronic data capture queries), 36.1% of the maximum rate between monitoring visits, and cost of each monitoring visit as low as 16.2% for trials conducted during the 2013 to 2017 period (Agrafiotis et al., 2018).

Moreover, it could also enable researchers to monitor multiple measurements from patients which would be nearly impossible for a clinician to monitor and analyze across a number of patients on a regular basis. This in turn could be used to flag certain changes, potential issues or anomalies for a particular patient, directing the medical team to take further action if they deem it necessary following review of the data (Dhinakaran et al., 2022). The benefits of AI based monitoring platform is given in figure 4.

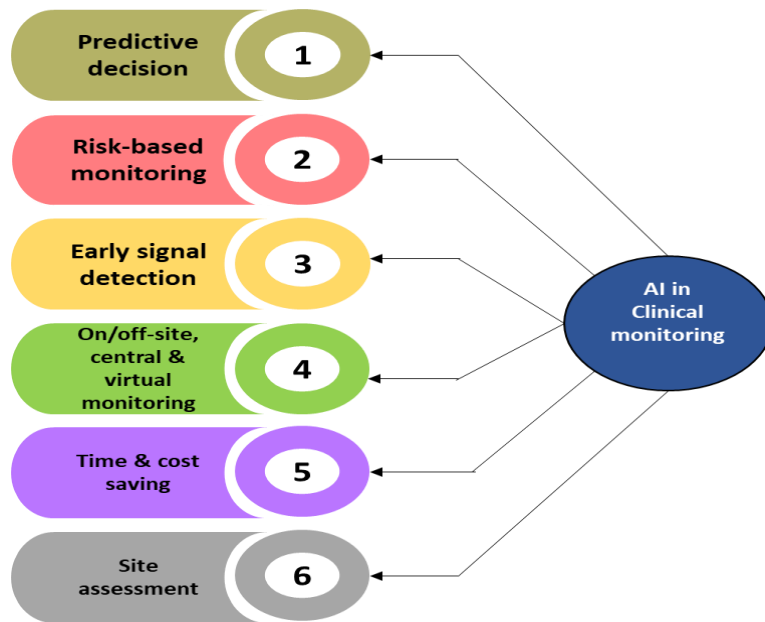


Figure 4. Benefits of AI based monitoring platform

### Virtual clinical trial

A virtual clinical trial refers to digitally empowered clinical trial processes. To overcome the clinical trial challenges, many pharmaceutical companies have started using an emerging concept of virtual clinical trials or decentralized trial. The various components of virtual trial are given in figure 5. This is a relatively new model taking full advantage of digital technology by including apps, virtual patient monitoring, wearable medical devices, web platforms for recruitment, informed consent, counselling, measurement of endpoints, and any adverse reactions which allow the patient to be home-based at every stage of the clinical trial (Ali et al., 2020). These virtual trials are patient-centric, cost-effective, and easy to manage. The trials are operationally feasible with increase recruitment rate, better study compliance, lower drop-outs and faster completion compared to routine studies. For example, Sanofi conducted a virtual diabetes trial (VERKKO) remotely in Europe. It was not drug trial but, Sanofi has tested a 3G-capable, wireless glucose meter. This study represents an important step forward in the clinical trial community, as it is the first clinical trial using an electronic consent approved by European regulatory authority (Narayanasetty & Jallu, 2021).

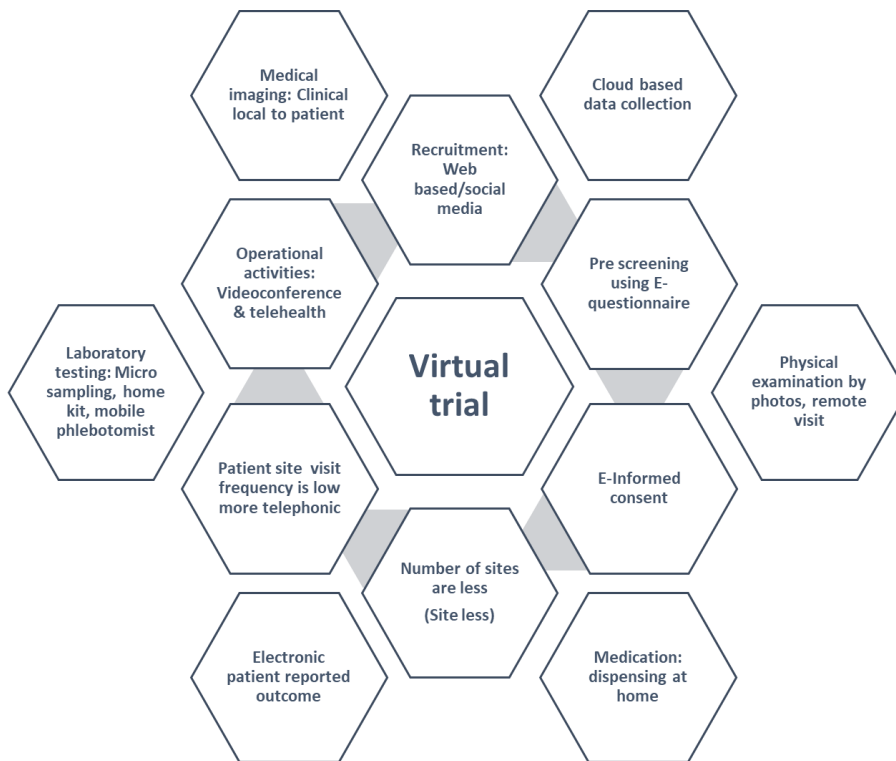


Figure 5. Components of virtual clinical trial

### Requirement of AI in pharmacovigilance

AI and machine learning are supplying capability insights to enhance patient safety with cognitive automation. While nonetheless early in usage, AI is demonstrating the feasibility to expect capability issues. AI is capable of intake a variety of structured and unstructured data formats, advanced OCR/ICR, robust and feature-rich Natural Language Processing (NLP) engine, and advanced analytics and other fields. The full landscape and capabilities of AI are displayed in figure 6. The rapid development of technology has brought opportunities to harvest massive amounts of data, both from social media and emerging biomedical technologies to allow prompt signal detection, develop early warning systems to enrich PV, and patient safety.

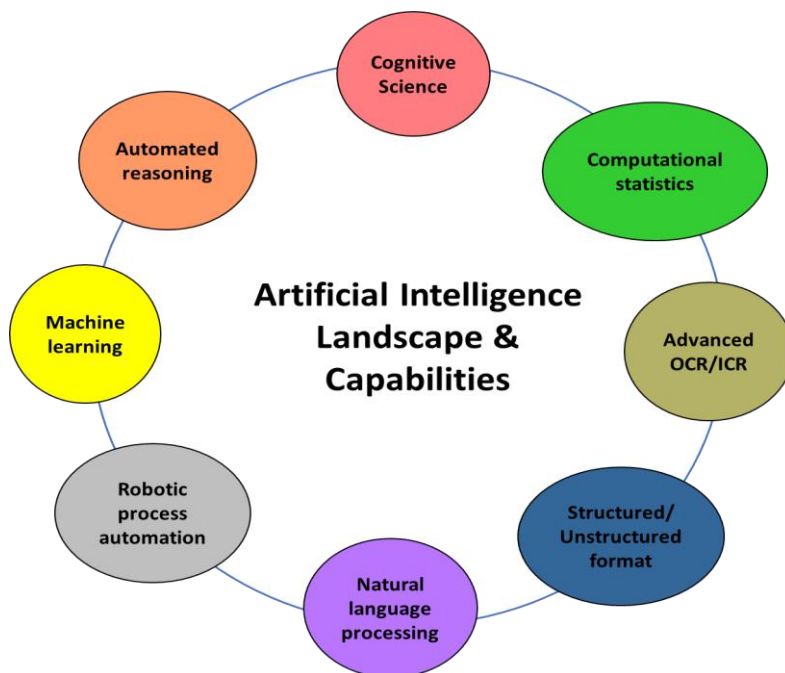


Figure 6. Full landscape and capabilities of Artificial Intelligence

Most of regulatory authorities across the world are necessitating submitting ADRs of international cases. Hence ICSRs are required to be translated into a local language before submission. Writing a safety narrative itself is a challenging task and translating it manually is an equally tedious process requiring a high degree of accuracy with the original narrative. Understanding the local colloquialisms, abbreviations, and source documents are important for translating safety narrative accurately. Considering the fact that same safety report needs to be submitted to regulatory authorities of different countries each with a different language, a multi-lingual support system inclusive of medical coding is desirable to enter cases in respective languages within the timelines.

China joined the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) in 2017 and adopted several ICH guidelines for safety (Yang et al., 2021). The drug regulatory authority of China, the National Medical Product Administration (NMPA) has set timelines for drug manufacturers to comply with electronic standards for safety report submissions in ICH E2B (R3) and in ICH E2B (R2) format in Chinese. To comply with these requirements, the safety database must meet the E2B (R3)'s XML format. Here we can clearly see the requirement to comply with the regulatory format, timeline, and language mandates. To meet the regulatory requirements, a system that allows accurate translation in applicable format within the stipulated timeline is a must. Employing manual task force for such repetitive work is resource-intensive and time-consuming.

Similarly, the new EudraVigilance system requires marketing authorization holders (MAHs) to log onto a common portal, download ICSRs, review available files, identify relevant reports, and import those cases into their safety database

(EMA, 2022). The ICSRs are delivered in the E2B (R3) XML format which is not easily readable in its raw form. To determine relevance, the PV staff must manually review each ICSR. The overall process is resource-intensive and inefficient.

### **Reasons for effective implementation of AI in pharmacovigilance**

#### **Changes in sources of information**

Internet expansion has deployed exponential information through novel sources influencing every bit of life. The new sources of information influencing PV industry are;

#### **Digital media**

Computer-readable format as websites, web pages, blogs, vlogs, social networking sites, internet forums, chat rooms, health portals.

#### **Wearable technology**

That collects biomedical data (e.g., heart rate, physical activity and sleep pattern, dietary patterns).

#### **Social media**

It is considered a sub-set of digital media and defined by European Commission's Digital Single Market Glossary as *a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0 and that allow the creation and exchange of user-generated content (EU, 2019). It employs mobile and web-based technologies to create highly interactive platforms via which individuals and communities share, co-create, discuss, and modify user-generated content.*

The communication platform is a large collection of PV sensors, each with its own level of uncertainty and bias. Data was not requested and processed in real time. Cell phones already have the ability to monitor local location, heart rate, temperature, blood glucose uptake, electrocardiography, respiratory rate, oxygen saturation, and sleep hours (Pandian et al., 2008; Yilmaz et al., 2010). This can be extended to wearable electrochemical sensors, where sweat, tears or pharmacy data can be collected (Bandodkar & Wang, 2014). Additionally, the collected geolocation data may be used to evaluate trends in AE emergence and drug use. For example, Google's volume of drug-related keywords provides an accurate measure of the actual use of those drugs (Simmering et al., 2014).

Also, after the 2010 earthquake in Haiti, Twitter was shown identifying patterns of cholera outbreaks weeks before traditional reports (Chunara et al., 2012). To comprehend such opportunities, PV teams need an equally able *modus operandi*. The volume and velocity of data generated from digital media sources potentially provide exciting opportunities for advances in PV. However, to yield the potential benefits from social media technical, regulatory and ethical challenges need careful consideration. The technical challenges refer to approaches to process the colossal data produced and extraction of relevant information to generate signals.

### **Changes in quality and quantity of information**

ICSRs require specific information in order to be reported and studies have shown that only a small quantity of social media posts have relevant information that can be reported (Davies, 2008). According to Web-RADR, despite a major contributor to the unstructured data, only 0.09% of signals were identified from social media sources in 2017 (Raizada, 2018). A study conducted to evaluate the potential of digging social media for medicines safety surveillance included 2537 posts related to rosiglitazone/cardiovascular events and 2236 posts related to HPV vaccine/infertility in the analysis. While only ten posts described personal accounts of rosiglitazone/cardiovascular AE experiences, and nine posts described HPV vaccine problems related to infertility, 21 % and 84 % of posts, respectively referenced other web pages and 72 % and 79% posts showed affirmation of the association (Coloma et al., 2015). In another study, 26 months of retrospectively collected social media data from a third-party vendor was used to determine how analysis of social media data could help in signal detection. The results showed that social media data analysis cannot identify new safety signals for selected products but provided a unique insight into the patient perspective (Bhattacharya et al., 2017). These studies show that a vast amount of data is available through new sources, but relevant data is insubstantial, which still cannot be overlooked.

### **Resource constraint**

Case processing is a repetitive and manual task. Case processing consumes a significant amount of resource; almost two-thirds as per PVNet benchmark data (Raizada, 2018). During one of the recent surveys regarding challenges in triaging reports, high volumes of AE and time spent downloading the files were concerns of the majority of responders (both 37%) than the reconciliation of applicable files (15%) and the manual segregation of files (11%) (ArisGlobal, 2018). Hence, the health care industry requires to converge efforts to reduce the case processing burden, and time required.

### **Changing times, changing demand**

The new drugs approved over recent years have limited long-term safety data raising concerns and demanding more real-time data to satisfy the need to know the possible consequences. To meet the demands of safety information, PV has to see beyond technical, ethical and regulatory challenges.

### **Efficient and progressive process**

Case processing in PV involves data verification and validity check, data entry, coding the AE and drugs, assessment of causality and expectedness, and writing case narrative. It is similar to sorting through the haystack for relevant and valid information. PV professionals perform many monotonous, recurring activities before analyses and evaluation of the cases. The AI has seamless applicability in ADR processing, allowing PV resources to focus on strategic activities to realize better outcomes.

## **AI applications in pharmacovigilance**

The repetitive, rule-based nature of the PV process makes it a suitable candidate to apply AI using NLP, ML, and robotic process automation. The AI application will build a thesaurus from the data sources to identify qualifying AEs which will be used to detect signals and prioritize. An application for efficient translation of narratives is a necessity to ensure high consistency and accuracy. A multi-lingual technology with an integrated arrangement for review and quality check of the translated case reports to ensure error-free content contemplating the regulatory requirement to submit an ADR in different countries would add ease. The audit trails of such system should serve as an important source to demonstrate compliance with quality processes during inspections. Some recent studies have shown applicability and advantages of AI in PV processing. A study, compared performance of a rule-based and machine learning models for classifying ICSRs from social media data with that of human PV experts. The investigators selected 311,189 social media posts with mentions of products and brands in combination with common medical and scientific terms. The model showed an accuracy of 65% and took 48 hrs to complete a task that would have taken an estimated 44,000 hrs for human experts (Comfort et al., 2018).

In another study, a consortium of ten cognitive services was trained through deep-learning. The input was 20,000 ICSRs received by Celgene drug safety over a 2-year period to evaluate the consortium to identify key characteristics of spontaneous ICSRs to an acceptable level of accuracy and effective use in a real-world setting. The consortium reached an evaluative score of  $\geq 75\%$  for spontaneous ICSRs (Abatemarco et al., 2018). These studies indicate that cognitive services are able to augment the PV process and free up professionals for decision-making activities to improve signal detection. Web-RADR; the joint venture between European agencies and public and private organizations, agrees that social media overall contains less safety information compared to VigiBase but it does provide insights into areas which is harder from traditional methods of reporting (WEB-RADR, 2019). The aim of Web-RADR is conducting scientific research on the use of social media networks and developing dedicated ADR reporting systems (Applications) to competent National Authorities in Europe.

There are a few examples of AI applications developed specifically for PV for tasks like auto-narrative generation; narrative analysis, QC assessment; causality assessment. The approach is advanced further with Intelligent Augmentation: supporting the human intelligence through technology to verify manual PV activities. 'Touchless' case processing is an forecasted AI application intended to receive, verify, code, process and submit non-serious cases without human intervention. The AI is useful tool in pharmacovigilance having multiple application which are highlighted below.

### **AI for data collection**

ICSR information needs to be extracted from the structured and unstructured data in a regulatory compliant manner from various sources such as drug labels, scientific publications, patient records, biomedical literature, emergency department, and information exchange on social media. NLP and text mining

together act as effective medical text classifiers to gather relevant facts and insights from the sources. Extracted information is not limited to patient safety events but include drug-drug interactions, medication, and disease as well.

### **AI for Data Enrichment**

The AI solution provides support dictionary for WHODD/MedDRA terms, product identification, PV field terms etc. Some of the research in this area has been devoted to creation of the annotated source data that are used to develop and test new machine-learning natural language processing algorithms. Cognitive Automation imparts ontologies that helps create lexicons of various biological concepts.

### **AI for Analysis and decision making**

The cutting-edge rich visualization framework enables in-place browser based, real-time analytics. It will help in clustering data and discovery of associations and early signal detection. Benefit-risk assessment. AI-based system learns from the historical data and can detect the risk and predictive algorithms could estimate burden. Also help in predict effectiveness of risk-minimization measures. Supervised or unsupervised learnings play a major role in devising hypothesis For example; building Unlisted Events and Drugs Correlation, Causality Classifiers etc., specific types of Neural Networks are built and improvised with training over a period. These are faster and more accurate compared to other methods.

### **AI for Output Generation**

The solution enables standardization of data and integration of output data with safety databases. The scope of pharmacovigilance is to develop a better understanding of drugs and risk associated with them. The scope continues to broaden with growing number of medicinal products and congruent problems caused by irrational drug use, drug overdose, polypharmacy, increased use of traditional and over-the-counter medications and drugs, illicit drug sales, increased self-medication, substandard medications, drug errors, and ineffectiveness. Drug safety is more than just the monitoring, detection and testing of ADRs that occur under clearly defined conditions and within a certain dose range. Increasing drug use worldwide on a large scale in the short term attracts the best international drug monitoring and efficiency. Current systems need flexibility to accommodate a wider scope with help of artificial intelligence.

### **Current challenges for AI in clinical research**

The clinical research is multi-stage and multi-disciplinary process hence there are various operational barriers for use of AI in clinical research. Scaling AI for long term use require professionalization of the industry. The clinical and AI domain experts are necessary for getting the success. Failing to have a trained team will land in the risks of either developing a model that distorts clinical reality or using an ML technique that is inappropriate to the available data and research question at hand (Wiens et al., 2019).

AI models require data at scale, hence there is concern on privacy of the large volume of patient data. Moreover, the data in clinical trial is generated by researchers, doctors, institution and they are reluctant to share with others. The biggest challenge to the development of novel AI-based tools is the curation of high-quality, representative data sets that are sufficient for training and validation of AI models to perform the desired tasks. It has been observed that, algorithms performed significantly differently in validation data sets compared with training data sets (Nestor et al., 2018). Additionally, a lack of data standards often causes problems with the analysis and effective use of the data.

Many geriatric patients are not familiar with newer technology and its usage. Therefore, the recruitment of such patients in clinical trials will be difficult; as they need a personal attention and relationship to get participated in study. Many philosophical assumptions about laboratory methods, statistical tools, or certain medications are the basis of barriers without its integrity and reliability. Thus, philosophical barrier has more recently become an operational barrier as well with the passage of the European Union's General Data Protection Regulation, which requires that automated decision-making algorithms provide meaningful information about the logic involved (Weissler et al., 2021). In case of Phase-I clinical trial of new molecules in healthy subjects, remote clinical monitoring is problematic. These kinds of study include assessment of early safety signals where patients need to be closely observed and located near a clinical site in case there is a reaction. Technological challenges in employing AI in PV include the complexity of the translation algorithms to interpret layperson's terminology. Layperson terminology refers to the following:

- Drugs are referred with their brand names, active ingredients, colloquialisms or common drug names.
- Use of artistic expressions or terms not found within existing medical dictionaries to describe Adverse Event. The terminology is not always based on the medical lexicon.
- Poor grammar, spelling errors, abbreviations, and sarcasm.
- Discussion about drug rather than the actual AE
- Other problems are lack of specificity. Also, it is difficult to evaluate the exact nature of the AE and assess causality, and most importantly, the ethical dilemma if such patients should be contacted to obtain further information.

This limits the effectiveness of the lexicon-based approach and reduces a system's ability to automatically extract relevant information, code it accurately for further analysis (Leaman et al., 2010; Nikfarjam & Gonzalez, 2011). Supervised machine learning requires training data which is human-annotated data. Only a small proportion of data collected from social media may contain relevant information, hence large volumes of human-annotated data are not possible. Traditional NLP methods have proven to be inadequate when applied to short texts, like on Twitter (Yates & Goharian, 2013). Discovering patterns in texts indicative of ADRs instead of lexicon-based approaches seems beneficial as it can detect inexact matches (Nikfarjam & Gonzalez, 2011; Yates & Goharian, 2013).

## Conclusion

The conduct of any clinical trial using randomized, double blind, active control design is more acceptable approach. Though, the use of AI techniques offers the potential to improve the success and efficiency of study. AI is being recognized for transforming the clinical development landscape in positive way. In general, it reduces the costs and accelerate every stage of clinical research and drug development from matching patients with clinical trials and handling data to discovering drugs themselves. These technologies hold huge promise, but at the same time, it cannot replace human expertise inclusive of chemical synthesis, clinical trials, regulatory approvals and production stages. Nevertheless, the requisiteness to start implementing AI technologies in the medicines research can be a breakthrough for the future market sustainability and revolution of biopharmaceutical organizations. Augmentation and maturity of AI tools will benefit patients, sponsors, payers and physicians alike. For such a broader advancement; reshaping of traditional clinical development approaches and effective use and implementation of AI-based techniques is inevitable.

## Acknowledgments

We are thankful to Mr. Shubham Padole, Ms. Sweta Lembhe & Ms. Shital Satone; Research Student, Department of Pharmacology, Dr. D. Y. Patil Institute of Pharmaceutical Sciences & Research, Pune-411018, Maharashtra. India collection of data during the manuscript writing.

## Conflict of Interest

No conflict of interest.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## References

- Abatamarco, D., Perera, S., Bao, S. (2018). Training Augmented Intelligent Capabilities for Pharmacovigilance: Applying Deep-learning Approaches to Individual Case Safety Report Processing. *Pharmaceutical Medicine*, 32(6), 391-401. [10.1007/s40290-018-0251-9](https://doi.org/10.1007/s40290-018-0251-9).
- Agrafiotis, D.K., Lobanov, V.S., Farnum, M.A., Yang, E., Ciervo, J., Walega, M., Baumgart, A., Mackey, A.J. (2018 July). Risk-based Monitoring of Clinical Trials: An Integrative Approach. *Clinical Therapeutics*, 40(7), 1204-1212. <https://doi.org/10.1016/j.clinthera.2018.04.020>.
- Alexander, M., Solomon, B., Ball, D.L., Sheerin, M., Dankwa-Mullan, I., Preininger, A.M., Jackson, G.P., Herath, D.M. (2020 July). Evaluation of an artificial intelligence clinical trial matching system in Australian lung cancer patients, *JAMIA Open*, 3(2), 209-215. <https://doi.org/10.1093/jamiaopen/ooaa002>.

- Ali, Z., Zibert, J.R., Thomsen, S.F. (2020). Virtual Clinical Trials: Perspectives in Dermatology. *Dermatology*, 236, 375-382. 10.1159/000506418.
- ArisGlobal. (2018). From Push to Pull – Making the Triage of Adverse Events Downloaded from EMA Hassle Free Available online at: <https://www.arisglobal.com/push-to-pull-making-triage-of-adverse-events-downloaded-from-ema-hassle-free/> (Accessed on 2<sup>nd</sup> August 2019).
- Babel, A., Taneja, R., Mondello Malvestiti, F., Monaco, A., Donde, S. (2021 June). Artificial Intelligence Solutions to Increase Medication Adherence in Patients With Non-communicable Diseases. *Front Digit Health*, 3, 669869. 10.3389/fdgth.2021.669869.
- Bain, E.E., Shafner, L., Walling, D.P., Othman, A.A., Chuang-Stein, C., Hinkle, J., Hanina, A. (2017 February). 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.*, 21, 5(2), e18. 10.2196/mhealth.7030.
- Bakkar, N., Kovalik, T., Lorenzini, I., Spangler, S., Lacoste, A., Sponaugle, K., Ferrante, P., Argentinis, E., Sattler, R., Bowser, R. (2018 February). Artificial intelligence in neurodegenerative disease research: use of IBM Watson to identify additional RNA-binding proteins altered in amyotrophic lateral sclerosis. *Acta Neuropathol*, 135(2), 227-247. <https://doi.org/10.1007/s00401-017-1785-8>.
- Bandodkar, A., Wang, J. (2014). Non-invasive wearable electrochemical sensors: a review. *Trends In Biotechnology*, 32(7), 363-371. 10.1016/j.tibtech.2014.04.005.
- Bhattacharya, M., Snyder, S., Malin, M. (2017). Using Social Media Data in Routine Pharmacovigilance: A Pilot Study to Identify Safety Signals and Patient Perspectives. *Pharmaceutical Medicine*, 31(3), 167-174. 10.1007/s40290-017-0186-6.
- Bhatt, A. (2021). Artificial intelligence in managing clinical trial design and conduct: Man and machine still on the learning curve? *Perspect Clin Res.*, 12(1),1-3. 10.4103/picr.PICR\_312\_20.
- Chunara, R., Andrews, J., Brownstein, J. (2012). Social and News Media Enable Estimation of Epidemiological Patterns Early in the 2010 Haitian Cholera Outbreak. *The American Journal of Tropical Medicine And Hygiene*, 86(1), 39-45. 10.4269/ajtmh.2012.11-0597.
- Coloma, P., Becker, B., Sturkenboom, M. (2015). Evaluating Social Media Networks in Medicines Safety Surveillance: Two Case Studies. *Drug Safety*, 38(10), 921-930. 10.1007/s40264-015-0333-5.
- Comfort, S., Perera, S., Hudson, Z. (2018). Sorting Through the Safety Data Haystack: Using Machine Learning to Identify Individual Case Safety Reports in Social-Digital Media. *Drug Safety*, 41(6), 579-590. 10.1007/s40264-018-0641-7.
- Davies, M. (2008). Listening to consumers in a highly regulated environment: how pharmaceutical manufacturers can leverage consumer-generated media. Available online at: [http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/11/Nielsen-Online-HealthcarePractice\\_Social-Media-Adverse-Event-Reporting\\_nov09.pdf](http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/11/Nielsen-Online-HealthcarePractice_Social-Media-Adverse-Event-Reporting_nov09.pdf) (Accessed on 2<sup>nd</sup> August 2019).
- Dhinakaran, M., Phasinam, K., Alanya-Beltran, J., Srivastava, K., Babu, V., Singh, S.K. (2022). A System of Remote Patients' Monitoring and Alerting Using

- the Machine Learning Technique. *Journal of Food Quality*, Article ID 6274092. <https://doi.org/10.1155/2022/6274092>.
- European Medicines Agency (EMA). (2022). EV reporting process users: EV Gateway, WEB-Trader, EV-post functions training module. Available online at: [https://www.ema.europa.eu/en/documents/presentation/presentation-ev-reporting-process-users-ev-gateway-web-trader-ev-post-functions-training-module-ev\\_en.pdf](https://www.ema.europa.eu/en/documents/presentation/presentation-ev-reporting-process-users-ev-gateway-web-trader-ev-post-functions-training-module-ev_en.pdf) (Accessed on 2<sup>nd</sup> April 2022).
- European Union. (2019). Available online at: <https://ec.europa.eu/digital-single-market/en/glossary> (Accessed on 2<sup>nd</sup> August 2019).
- Fisher, A., McKenney, J. (1993). The development of the ERMA banking system: lessons from history. *IEEE Annals of The History Of Computing*, 15(1), 44-57. 10.1109/85.194091
- Fred, P. Accenture: Artificial Intelligence Healthcare Market to Reach \$6.6B by 2021. (2017 June). Available online at: <https://hitconsultant.net/2017/06/23/artificial-intelligence-healthcare-market-acn/> (Accessed on 15<sup>th</sup> March 2022).
- Glass, L., Shorter, G., Patil, R. (2019). Five Ways to Enhance Clinical Operational Efficiencies Utilizing AI. Available online at: [https://www.contractpharma.com/contents/view\\_experts-opinion/2019-10-21/five-ways-to-enhance-clinical-operational-efficiencies-utilizing-ai/#\\_ftn1](https://www.contractpharma.com/contents/view_experts-opinion/2019-10-21/five-ways-to-enhance-clinical-operational-efficiencies-utilizing-ai/#_ftn1) (Accessed on 23<sup>rd</sup> March 2022).
- Haddad, T., Helgeson, J.M., Pomerleau, K.E., Preininger, A.M., Roebuck, M.C., Dankwa-Mullan, I., Jackson, G.P., Goetz, M.P. (2021 March). Accuracy of an Artificial Intelligence System for Cancer Clinical Trial Eligibility Screening: Retrospective Pilot Study. *JMIR Med Inform.*, 9(3), e27767. 10.2196/27767.
- Harrer, S., Shah, P., Antony, B., Hu, J. (2019 August). Artificial Intelligence for Clinical Trial Design. *Trends in Pharmacological Sciences*, 40(8), P577-591. <https://doi.org/10.1016/j.tips.2019.05.005>
- Leaman, R., Wojtulewicz, L., Sullivan, R. (2010). Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. In *2010 workshop on biomedical natural language processing*, 117-25.
- Liu, X., Cruz Rivera, S., Moher, D., Calvert, M.J., Denniston, A.K., SPIRIT-AI and CONSORT-AI Working Group. (2020 September). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med.*, 26(9), 1364-1374. 10.1038/s41591-020-1034-x.
- Mitchel, J.T., Gittleman, D., Schloss Markowitz, J.M., Cho, T., Kim, Y.J., Choi, J., Hamrell, M.R., Nora, S.D., Carrara, D. (2014). Time to change the clinical trial monitoring paradigm: results from a multicenter clinical trial using a quality by design methodology, risk-based monitoring and real-time direct data entry. *Appl Clin Trials*, Available online at: <http://www.appliedclinicaltrials.com/time-change-clinical-trial-monitoring-paradigm> (Accessed on 31<sup>st</sup> March 2022).
- Mitchel, J.T., Kim, Y.J., Choi, J., Park, G., Cappi, S., Horn, D., Kist, M., D'Agostino, R.B. (2011). Evaluation of Data Entry Errors and Data Changes to an Electronic Data Capture Clinical Trial Database. *Drug Information Journal*, 45(4), 421-430. 10.1177/009286151104500404.

- Narayanasetty, S., Jallu, R. (2021). A Review on Virtual Clinical Trials: The Future. *Int. J. Pharm. Sci. Rev. Res.*, 68(1), 111-116. <http://dx.doi.org/10.47583/ijpsrr.2021.v68i01.019>.
- Nayak, V.S., Khan, M.S., Shukla, B.K., Chaturvedi, P.R. (2016). Artificial intelligence in clinical research. *Int J Clin Trials*, 3(4), 187-93.
- Nestor, B., McDermott, M.B.A., Chauhan, G., Naumann, T., Hughes, M.C., Goldenberg, A., Ghassemi, M. (2018 November). Rethinking clinical prediction: Why machine learning must consider year of care and feature aggregation. 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada. <https://doi.org/10.48550/arXiv.1811.12583>.
- Nikfarjam, A., Gonzalez, G.H. (2011). Pattern mining for extraction of mentions of Adverse Drug Reactions from user comments. *AMIA Annu Symp Proc.* 2011, 1019-1026.
- Pandian, P., Mohanavelu, K., Safeer, K. (2008). Smart Vest: Wearable multi-parameter remote physiological monitoring system. *Medical Engineering & Physics*, 30(4), 466-477. 10.1016/j.medengphy.2007.05.014.
- Raizada, M. (2018). Digitally transformed pharmacovigilance. Available online at: [http://www.navitaslifesciences.com/collaterals/Whitepapers/WP-Digitally\\_transformed\\_pharmacovigilance.pdf](http://www.navitaslifesciences.com/collaterals/Whitepapers/WP-Digitally_transformed_pharmacovigilance.pdf). (Accessed on 2<sup>nd</sup> August 2019).
- Rivera, S.C., Liu, X., Chan, A.W., Denniston, A.K., Calvert, M.J., SPIRIT-AI and CONSORT-AI Working Group. (2020 September). Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. *BMJ*, 9, 370, m3210. 10.1136/bmj.m3210.
- Simmering, J., Polgreen, L., Polgreen, P. (2014). Web search query volume as a measure of pharmaceutical utilization and changes in prescribing patterns. *Research In Social And Administrative Pharmacy*, 10(6), 896-903. 10.1016/j.sapharm.2014.01.003.
- Spangler, S., Wilkins, A.D., Bachman, B.J., Nagarajan, M., Dayaram, T., Haas, P., Regenbogen, S., Pickering, C.R., Comer, A., Myers, J.N., Stanoi, I., Kato, L., Lelescu, A., Labrie, J.J., Parikh, N., Lisewski, A.M., Donehower, L., Chen, Y., Lichtarge, O. (2014 August). Automated hypothesis generation based on mining scientific literature. Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 1877-1886. <https://doi.org/10.1145/2623330.2623667>.
- Sparkes, S. (2018). The Role of Artificial Intelligence within Pharmacovigilance and Medical information. Aris Global. Available online at: <https://www.arisglobal.com> (Accessed on 2<sup>nd</sup> March 2019).
- Tantsyura, V., McCanless Dunn, I., Fendt, K., Kim, Y.J., Waters, J., Mitchel, J. (2015). Risk-based monitoring: a closer look at source document verification, queries, study size effects and data quality. *Ther Inno. Regul Sci.*, 49, 903-910. <https://doi.org/10.1177/2168479015586001>.
- Taylor, K., Properzi, F., Cruz, M.J. (2020). Intelligent clinical trials Transforming through AI-enabled engagement. Available online at: <https://www2.deloitte.com/us/en/insights/industry/life-sciences/artificial-intelligence-in-clinical-trials.html/#endnote-sup-5> (Accessed on 31<sup>st</sup> Mar. 2022).
- WEB-RADR. (2019). WEB-RADR: Recognising Adverse Drug Reactions. Available online at: <http://web-radr.eu/> (Accessed on 2<sup>nd</sup> August 2019).
- Weissler, E.H., Naumann, T., Andersson, T., Ranganath, R., Elemento, O., Luo, Y., Freitag, D.F., Benoit, J., Hughes, M.C., Khan, F., Slater, P., Shameer, K.,

- Roe, M., Hutchison, E., Kollins, S.H., Broedl, U., Meng, Z., Wong, J.L., Curtis, L., Huang, E., Ghassemi, M. (2021). The role of machine learning in clinical research: transforming the future of evidence generation. *Trials* 22, 537. <https://doi.org/10.1186/s13063-021-05489-x>.
- Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V.X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P.N., Thadaney-Israni, S., Goldenberg, A. (2019 September). Do no harm: a roadmap for responsible machine learning for health care. *Nat Med.*, 25(9), 1337-40. <https://doi.org/10.1038/s41591-019-0548-6>.
- Wong, C.H., Siah, K.W., Lo, A.W. (2019 April). Estimation of clinical trial success rates and related parameters. *Biostatistics*, 20(2), 273-286. <https://doi.org/10.1093/biostatistics/kxx069>.
- Yang, F., Heng, J., Li, K., Wang, J. (2021 March). The new Good Clinical Practice-2020 in China: Views from ethical perspective. *The Lancet*, 8, 100117. <https://doi.org/10.1016/j.lanwpc.2021.100117>.
- Yates, A., Goharian, N. (2013). ADRTrace: Detecting Expected and Unexpected Adverse Drug Reactions from User Reviews on Social Media Sites. In: ,et al. *Advances in Information Retrieval. ECIR 2013. Lecture Notes in Computer Science*, vol 7814. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-36973-5\\_92](https://doi.org/10.1007/978-3-642-36973-5_92)
- Yilmaz, T., Foster, R., Hao, Y. (2010). Detecting Vital Signs with Wearable Wireless Sensors. *Sensors*, 10(12), 10837-10862. 10.3390/s101210837.