ARTIFICIAL INTELLIGENCE IN CLINICAL TRIALS

Abstract

In the era of evidence based medicine thousands of clinical trials are conducted generating huge data from trial participants. Traditional trials highlighted ethical issues leading to patient withdrawal from trials, inadequate enrolments which gave rise to digitalisation with intent to promote authenticity, integrity in clinical research. Artificial intelligence (AI) is a boon in digitalization. It strengthened patient-physician relationships but it also contained some negative aspects which are being resolved. AI creates strong framework for clinical data management.

Keywords: Artificial intelligence, clinical data management, clinical trial, digitalisation, drug development

Authors

Dr. Sonu Gupta
Assistant Professor
Department of Clinical Research
Dasmesh Institute of Research & Dental Sciences
Faridko,.India.

Dr. Ruchika Garg
Professor
Department of Physiology
Adesh institute of Medical Sciences and Research, Hospital
Bathinda, India.

Dr. Ravinder Nath Bansal
Deputy Medical Superintendent
Hospital Administration,
Guru Gobind Singh Medical Hospital
Faridkot, India.

Dr. Sanjana Bansal
Assistant Professor
Department of Anatomy,
Adesh institute of Medical Sciences and Research, Hospital
Bathinda, India.

I. INTRODUCTION AND OBJECTIVE

Alan Turing, the founder of artificial intelligence defined AI as the science dealing with building machinery, chiefly computer programs. In health care sector AI is combined with analytics (AIA). Computers programs which assist in simulating human behaviour are the AI systems. Its clinical applications incorporate clinical decision making, automatic surgery, monitoring of patients and healthcare management etc. [1] Artificial intelligence (AI) uses enormous data from machine learning techniques which are advanced in nature and incorporate manifold artificial neural networks. [2]

Past trials highlighted ethical issues leading to patient withdrawal from trials, inadequate enrolments which gave rise to digitalisation with intent to promote authenticity, integrity in clinical research. Artificial intelligence is a boon in digitalization. It strengthened patient- physician relationships but it also contained some negative aspects which are being resolved. AI creates strong framework for clinical data management.

II. ARTIFICIAL INTELLIGENCE IN THE DRUG DISCOVERY AND DEVELOPMENT PROCESS

AI has its own role in drug discovery and development. It has profound effect on drug development. Evolution of big data and machine learning concepts in AI has increased annual sales in the health care system. [3] AI has promoted phenotypic drug discovery by allowing screening of compounds in cells or animal models without any prior information on biological target. Many pharmaceutical companies have started collaboration with AI companies and are adopting AI approaches for drug candidates, redrafting the new indications for already available drugs and in advanced stage of drug development also. [3]

AI has diverse role in drug discovery and clinical data management. Traditionally clinical trials raised ethical issues resulting in patient withdrawal, inadequate enrolments with a consequent result of digitalisation with intent to promote authenticity, integrity in clinical research. Artificial intelligence is a power boost in digitalization. It has decentralised the clinical trials and strengthened management of clinical trial data.

III. CLINICAL DATA MANAGEMENT (CDM)

Clinical trials are conducted for collecting trial data in response to research question for hypothesis testing. Data is the most important aspect of clinical trial. Research outcome of a trial depends on the quality of generated data. [4, 5] CDM is basically the collection, cleaning, and management of data obtained from trial participants but [4] in compliance with regulatory standards. Clinical data management is primarily conducted to provide authentic and quality clinical data with as minimum as possible errors and missing data. For this purpose good clinical practices are utilized and assurance is made that generated data is complete in all aspects and processed in a correct way. This is accomplished with the help of software which allow audit track, help in provide early detection and clarification of discrepancies in the collected data. In large scale and complex trials sophisticated innovations in CDM ensure easy handling and good data quality. [4, 6]

IV. HIGH-QUALITY DATA

Data which is accurate in terms of statistical analysis and complies with protocol-specified parameters and requirements is called high quality data. For such type of data there are some requirements which are listed below

- Data should have acceptable level of difference without affecting the concluding part of the study.
- In case there is encountered deviation from protocol the participant should be excluded from final database and it may be audited by regulatory authorities.
- Collected data should be free from missing data and minimal acceptable misses.
- Quality of data should comply with applicable regulatory requirements. [4]

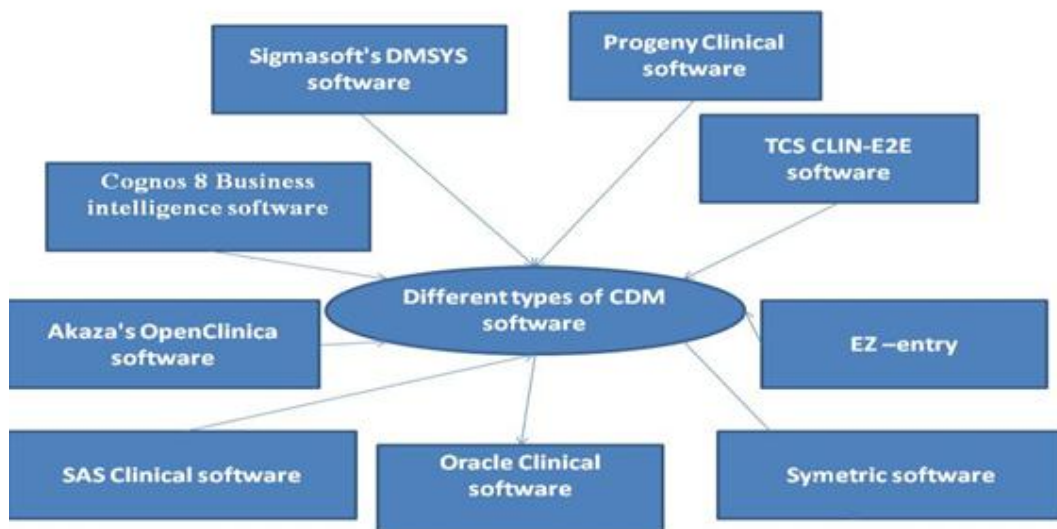## V. DIFFERENT CDM SOFT WARE AND THEIR APPLICATIONS

Figure 1

1. **SAS Clinical Software:** This software addresses unique needs of a new drug compound throughout drug discovery and development process. It generates a simple and easy podium for assessment and management of trial data from several sources like Clinical Data Management system (CDMS), Electronic Data Capture (EDC). It benefits into automation and recognises processes to diminish manual intervention. Appropriate utilization of standards is assured by the correct data usage. Data generated from SAS ensures compliance with quality standards and is less time consuming. [7, 8]

2. **EZ –entry:** It is a modification of EpiData software program in addition to several modules. Actions covered by this software are query management, revision tracking, data entry, import and export of data and finally quality control. It is user friendly and a secure system with user authentication manual. User can access database and revision tracking manual where revision in the original database can be made and a new entry is recorded automatically in the system. Quality of data is ensured by data entry, field value check and query forms. [7, 9]

3. **Oracle Clinical Software:** Information provided by this software is steadfast and protected. Its key benefits are successful team effort, speedy execution, fruitful marketing, industrial compliance and shift from paper to EDC etc. [6]

4. **TCS (Tata Consultancy Services) CLIN-E2E Software:** This software is compliant with 21 CFR part 11 of the GCP. It addresses all 4 phases of the clinical trials. It captures the e-data in addition to integrative approach to sponsors with clinical trial site and the laboratories. It provides the pharmaceutical companies a platform to generate case report forms (electronic and paper) for clinical trial data and monitoring of the site. It generates reusable study templates and CRFs saving cost and time it is in conformity with 21 CFR part 11 of the Good Clinical Practices. [6]

5. **Cognos 8 Business Intelligence Software:** This software has service oriented architecture (SOA) and plays a role in business Intelligence capabilities. This software has made CDM process very easy, simple and accurate. Data quality and performance is determined at personnel level and from collaborators. [6, 8, 10]

6. **Symetric Software:** This software benefits in processes of clinical trial data management such as database sets, quality control of trial data and final export of trial data. This software has complete processes incorporated with discrepancy management, trial data dictionary and coding missing data. It also provides benefit of interactive double-data check, tracking of CRF and any query management. [6]

7. **Akaza's OpenClinica software:** This software acts as a collaborative model having modular architecture flexible for obtaining data for EDC in the clinical data management system. It organizes protocol compliant trials. It is secure and supports in appropriate management of the clinical data with frequent visits. It provides tools for the importation and exportation of trial data across excel sheets. Other applications are data query interfaces, recovery of data, compliance with Health Insurance Probability and Accountability guidelines. [6]

8. **Progeny Clinical Software:** It tracks data from family history. Its main characteristics are family background, sub-pedigrees, custom display pedigrees, displays Haplotypes etc. [6]

9. **Sigmasoft's DMSYS Software:** It is affordable and completely validated software. It provides high quality data management services for large scale as well as small scale clinical trials. Its unique application of short cuts allow the operator to replicate clinical data entry sheet, forms, logic checks and importation/exportation of trial data from spread sheet. It serves quick cleaning of data and supports validation by its user validation project. [6]

VI. PROCESS OF CLINICAL DATA MANAGEMENT

1. **Data Collection:** Data is collected either on paper CRF or an electronic CRF (e CRF). Paper CRFs data is filled by an investigator as per the guidelines and data is fed into database by translated data entry. In case of e-CRF the first step is log in the CDMS by the trial investigator and data entry at the clinical trial site. Data errors are minimally reported when data is collected by e-CRF mode and in case there is any data discrepancy

it can be resolved earlier. Pharmaceutical companies opt for speedy drug development by preferring electronic CRF options or distant clinical data entry. [4, 11]

2. **CRF Tracking:** Once the data is fed in the data base, tracking is assured to completeness and authenticity by Clinical Research Associate (CRA) under the guidance of CDM team. This team tracks CRFs for any missing, illegible data and maintains documentation. If any missing data or illegible data exist then the investigator is answerable. [4, 11, 12]

3. **Data Entry:** Double data entry is recommended for paper CRF as such data entry facilitates cleaner database and ensures better compliance due to minimal errors. Double entry of trial data is done by two operators as entry by the 2nd operator facilitates verification and reconciliation of unreadable data by recognizing transcription mistakes. [13, 14, 15]

4. **Data Validation:** Next step after clinical data entry is the validation of data for compliance as per protocol specifications. This step promotes identification of any pitfalls in the entered data. Data validation program (DVP) consists of edit checks as per logic condition mentioned to ensure data validity. Discrepancies may arise from contradictory data, deficient data, protocol deviations, range checks. This can be sorted by investigator after login into CDMS. Throughout the entire process of clinical data management quality control (QC) is a continuous process and is performed at regular intervals. [10]

5. **Discrepancy Management:** Discrepancy management is most crucial and attention seeking step. It includes review of discrepancies, assessing reason, followed by either resolving or announcing as irresolvable. Management of discrepancy facilitates data cleaning and provides sufficient evidence regarding deviation of data. Approximately all CDMS have a discrepancy database. Discrepancies can be either flagged to the investigator site that require clarifications or closed in-house and reorganized in the system. In case of electronic case report form, discrepancies are resolved by the investigator online. Discrepancies are reviewed by CDM team at regular intervals to ensure resolved and declared 'closed' but sometimes this is impossible. [11, 12]

6. **Medical Coding:** Medical coding is the recognition and properly classification of clinical trial associated medical terminologies. Medical coding classifies medical terms reported on the CRF to avoid data replication. This is possible with technical expertise, skills understanding of medical terms, knowledge of disease, medicines used. Coding of adverse events and other diseases is done by Medical Dictionary for Regulatory Activities (MedDRA). World Health Organization–Drug Dictionary Enhanced (WHO-DDE) is the dictionary used for the coding of medications. WHO-ART is the dictionary for terminology of adverse reactions. Customized dictionaries are also used by some pharmaceutical companies. [4]

7. **Database Locking:** Post quality check and quality assurance finalized data is validated. The SAS datasets are finalized under guidance of statistician. Before data base lock a checklist is used to check completion of all activities. After approval from all stakeholders and extraction of clean database is locked and it cannot be amended in the database further, except by privileged users. After locking, data is extracted from the final database. [4]

8. **Players in CDM:** Team members in clinical data management should be science graduate and skilled in computers. Medical and paramedical graduates can be recruited as medical coders. Key roles and responsibilities for clinical data management team are shown in table1.

Player	Role and responsibilities [4, 16]
Data Manager	- supervises entire CDM process - approves the clinical data management procedures, prepare the DMP and documentation of clinical data management - Controls & allocates team access to database
Database Designer / Programmer	- formulating study database, CRF annotation - edit checks for data validation - in the database designing data entry screens and validates the edit checks
Medical Coder	Codes for AE, history related to any medical illness and related medication given at some stage in the conduct of study
Clinical Data Coordinator	- CRF designing, CRF filling directions - develops the DVP & conducts paperwork for discrepancy management, checklists etc
QC associate	- checks for correctness of data entry and data audits - document verification
Data entry operator	data entry & CRF tracking

VII. DIGITALISATION IN DRUG DEVELOPMENT

There has been reported a transformation in drug development due to expensive and complex nature of clinical trial and patient centric approach of drug trials. Safety and efficacy of experimental molecule is an important concern in clinical research. During the entire process of clinical trial, sponsors collaborate with clinicians and patients for detecting safety and efficacy of a novel molecule. Trials are usually designed in such a way that it is feasible and easy for drug sponsor to conduct the study as per his convenience but no consideration is given to convenience of patients which increases complexity in trials. But now the pharmaceutical industry is following a patient centric approach in drug development which has its unique advantages like patient is acquainted with disease, trial procedures, drug molecule, site of trial conduct and results from similar studies or trials. This results in quality talk between patients and physician. Contract research organizations (CROs) provide readymade repository of trial patients in the support of pharmaceutical companies by saving cost and effort. E.g. Digital Patient Unit program of quintiles uses real world patients and provide opportunity to pharma sponsors to screen them and recruit them as trial participants. [17]

1. **mHealth** utilizes clinical practice under the umbrella of portable diagnostic devices. Such device has transformed health care delivery system from health-systems generated to remote cum patient centric. This transformation increases opportunities for patient enrolments, outcomes and decreases the cost fcator. According to World Health

Organization mHealth is a triad of e-health, public health and medical practice facilitated by use of devices such as cellular phones, personal digital assistants (PDAs), patient monitoring gadgets etc. It is advantageous over traditional trial programmes in terms of safety, efficacy, real time quality data, real world evidence from continuous data, improved patient adherence and remote monitoring. [18, 19]

2. **Applications and Limitations of mHealth**
   - Self-care and self-monitoring in terms of blood glucose levels, weight, and daily activity.
   - Managed care: Monitoring technology for instance as
     - in Dementia: monitoring patient visits to the refrigerator or bathroom,
     - in COPD: use of inhaler,
     - in Ebola: vitals monitoring,
     - in Diabetes: blood glucose levels,
     - in Vaccines: body temperature,
     - in Hypertension/diabetes: medical and physical activity adherence,
     - Hospital at home: Bluetooth devices can be connected enabling patients to be at home, rather than in hospital. e.g. intravenous (IV) administration, dialysis at home, and webcams using two-way communication between patients and caregivers. The technology can also be used to link small hospitals into central hubs to provide expert care.

   Limitations: lack of regulatory clearance, expensive, security of data issues. [18, 20]

3. **Roche smartphone app:** Roche pharma Research & Early Development provides remote patient monitoring by **smartphone app** for patients suffering from Parkinson disease. This app is complemented with traditional physician assessments using a scale that measures disease, disability, impairment and symptom severity but restricted to the specific times when patients goes to physicians. The active tests include a series of half minute activities like such as balance test, voice test, gait test, dexterity test, postural tremor test, rest tremor test. This is advised to be used for a period of about 32 weeks. [21]

4. **HealthPatch MD** from collaboration of Vital Connect and Medidata is a biosensor which is wearable with ease for remote monitoring of patient in an efficient way. It helps in continuous measurement of heart rate, skin temperature, physical activity, respiratory rate, electrocardiogram grading etc. It captures real-time client's physical health parameters.

5. **Iodine** – is an app for flu & cold season from health information website. The app helps its consumers to find and compare a lot of medication for cold and flu to cure their symptoms. [17]

6. **Reg4all or Registries for All** is matchmaker between clinical trials and patients. This has a patient centric approach with decision making for patients to whom they can allow to access their clinical data. [17]

7. **Treato- treato** captures original information from the social media like Facebook, Twitter, and patient forums using patient language dictionaries, big data analytics and

natural language processing algorithms. This captured information from app guides the pharmaceutical companies in drug development by gaining insight in patient lives.

8. **Electronic Health Records (EHRs) via HealthIT.gov:** In healthcare sector it is most crucial to improve patient satisfaction, clinical outcomes and generate income. US government has encouraged the use of EHR via HealthIT.gov and provides benefit in the form of incentives to doctors who are minimizing medical errors and give quality by meaningful use of EHRs. [17]

9. **Patient education tools:** There are the websites available for educating the patients for diseases. For example **Krames patient education** by StayWell, **ExitCare OnScreen™** video solutions for education of clients, **WELVU** – Mobile based patient education tool providing medical images, scores, and health outcomes, **MediGuard™** a dose compliance tracking tools reminds for dose intake. In clinical trial retention of patients is crucial for the project success. Acurian, a recruitment and retention service source to the patients uses social media like Facebook and Myspace for patient retention purposes. [17]

10. **Application of Digitalisation tools in CDM:** Digital technologies can transform the future of clinical trials. Digital technologies can facilitate participation by clinical research staff and yield quality data. [22, 23]

VIII. FACILITATORS AND BARRIERS IN DIGITALISED CDM

1. **Enrolment in Trial:** Recruitment in a clinical trial is a most critical point of consideration since more than 2/3 clinical trial sites don't succeed to accomplish original target. Digital approaches such as websites and online patient communities and EHRs, laboratory results can facilitate in recruitment procedure with minimal cost and efforts. One suitable example of enrolment tool is Antidote which retrieves data from website (clinicaltrials.gov) and machine learning generates structured clinical trial criteria. It fabricates pre-feasibility questionnaire into easily understandable language for patients. By completing such questionnaire they can easily sift through several studies and choose one of their interest. [22, 23]

2. **Participants to Collaborators:** Digital technology has transformed Participants from subjects to collaborators while increasing patient engagement in the clinical research minimizing research mistrust and addressing patient-centric concerns. Digitalization measures patient outcomes like performance in specific daily activities and feedback via online surveys and focus groups to reshape the final treatment. [22, 23]

3. **Multiple Visits:** Mutilple visits to clinical trial sites for assessment are somewhat problematic, inconvenient and increase cost factor. Digital technologies has reduced travel time of trial participants and increased convenience. This factor affects willingness and ability to participate. Virtual trials with e-consent, telemedicine, apps and biosensors for patient communication make data collection easy from homes with no trial site visits. It was found that about 50% of clinical trials can be conducted virtually. One example is Roche in multiple sclerosis (MS) study where smartphone sensor connected app was used to monitor and compare findings with clinical evaluation. It helped in assessment of patient's neurological activity by performances such as gait, balance exercises, cognitive tests and hand and wrist turning. Assessments were depicted in the form of disease trend.

Results from remote study were found more sensitive as compared to in-clinic assessments. [22, 23]

4. **Improved Treatment Adherence by Digital Tools**: Digital technologies improve treatment adherence in the form of self-reporting and assurance about consumption of prescribed drugs which is verified by blood investigations but sometimes may require extra visits to trial site. Digital adherence tools such as smartphone apps and text messaging can record patient health data, answer their queries in real time, remind them to take medication and plan site visits. Even facial recognition tools can confirm adherence. [22, 23] Clinical data management entails utilization of many software tools and like Rave, Oracle Clinical and eClinical Suite etc. (table. 2) where huge amount of data is generated as in case of multicentric trials, CDMS is most crucial to handle such data. [4]

Table 2: Commonly used Clinical Data Management (CDM) tools

Software	Applications	Reference
BRAAN, DataLabs, Fast Track	at the stage of Clinical Trial Development for designing Protocol and execution	[7]
Oracle clinical	for study monitoring and reporting, AE reporting, document management	
Nextrials, DataLabs, Oracle, Phase Forward, DataTrack, Parexel, ClinPhone, CRF, eResearch Technology	invivo data, electronic data, electronic patient diaries, capture, case report forms	
SyTech, Wimmer systems	during completion of study and regulatory filing, analysis of data	

Every information obtained from study participants is sensitive and should be protected. Security of data is most essential aspect of clinical research. Privacy and confidentiality of the data should be maintained in compliance with the informed consent. With AI there are chances of lack of transparency also which need to be resolved. [24]

IX. CONCLUSION

AI tools have proved their performance in drug discovery and development. AI tools have similar fuctions and demand sophisticated information technology and infrastructure. Some softwares can be downloaded free of cost from respective websites. Pharmaceutical companies are in collaboration with AI companies. This collaboration is a boon for clinical research industry but this has generated some ethical issues also. Ethical guidelines in the medical field regarding the use of AI are being established by many international authorities.
Conflict of Interest: Nil

REFERENCES

[1]   Azzi S, Gagnon S, Ramirez A etal. Healthcare Applications of Artificial Intelligence and Analytics: A Review and Proposed Framework Appl. Sci. 2020, 10, 6553; doi:10.3390/app10186553 page 1-21

[2]   Park SH, Kim YK, Lee JY et al. Ethical challenges regarding artificial intelligence in medicine from the perspective of scientific editing and peer review. Sci Ed 2019; 6(2):91-98. https://doi.org/10.6087/kcse.164

[3]   Lamberti MJ, Wilkinson M, Donzanti BA etal. A Study on the Application and Use of Artificial Intelligence to Support Drug Development Clinical Therapeutics.2019:41(8); 1414-1426.

[4]   Krishnankutty, B., Bellary, S., Kumar, N. B., & Moodahadu, L. S. (2012). Data management in clinical research: An overview. Indian journal of pharmacology. 2012: 44(2); 168–172. https://doi.org/10.4103/0253-7613.93842

[5]   Gerritsen MG, Sartorius OE, vd Veen FM, Meester GT. Data management in multi-center clinical trials and the role of a nation-wide computer network. A 5 year evaluation. Proc Annu Symp Comput Appl Med Care 1993:659-62

[6]   Lu Z, Su J. Clinical data management: Current status, challenges, and future directions from industry perspectives. Open Access J Clin Trials 2010;2:93-105.

[7]   Gazali, Kaur S, Singh I. Artificial intelligence based clinical data management systems: A review. Informatics in Medicine Unlocked. 2017: 9 ; 219–229

[8]   Litchfield J, Freeman J, Schou H etal. Is the future forclinical trials internet-based? A cluster randomized clinical trial. Clin Trials 2005;2(1):72–9.

[9]   Gao QB, Kong Y, Fu Z et al. EZ-entry: a clinical datamanagement system. Comp Biol Med 2008;38(9):1042–4.B.

[10]   Almenoff JS, Pattishall EN, Gibbs TG etal. Novelstatistical tools for monitoring the safety of marketed drugs. Clin Pharmacol Ther2007;82(2):157–66.

[11]   Argraves S, Brandt CA, Money R etal. Informatics tools to improve clinicalresearch study implementation. Contemp Clin Trials 2006;27(2):112–22.

[12]   Gerritsen MG, Sartorius OE, vd Veen FM etal. Data management in multi-center clinical trials and the role of a nation-wide computer network: a 5 year evaluation. Proc Annu Symp Comput Appl Med Care. 1993:659–62.

[13]   Cummings J, Masten J. Customized dual data entry for computerized data analysis. Qual Assur 1994; 3:300-3.

[14]   Reynolds-Haertle RA, McBride R. Single vs. double data entry in CAST. Control Clin Trials 1992;13:487-94

[15]   Lu Z. Information technology in pharmacovigilance: benefits, challenges, and futuredirections from industry perspectives. Drug Healthc Patient Saf . 2009;1:35–45.

[16]   Ottevanger PB, Therasse P, van de Velde C etal. Quality assurance in clinical trials. Crit Rev Oncol Hematol 2003;47:213-35.

[17]   Sharma NS. Patient centric approach for clinical trials: Current trend and new opportunities. Perspect Clin Res. 2015;6(3):134-138. doi:10.4103/2229-3485.159936

[18]   WHO, New horizons for health through mobile technologies Global Observatory for e-Health series. 2011. Vol.3. Available from:www.who.int/goe/publications/goe_mhealth_web.pdf. Last accessed 6 May 2021.

[19]   Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. Eur Heart J. 2016;37(18):1428-1438. doi:10.1093/eurheartj/ehv770

[20]   King L, Agarwal S, Bhardwaj S et al. "e Source in Clinical Research: A Data Management Perspective on the Use of Mobile Health Technology," the Society for Clinical Data Management.. Available from http://scdm.org/publications/white-papers/. Last accessed 07 May 2021.

[21]   Pennic F. New Roche App Measures Parkinson's Disease Fluctuations. 2015 https://hitconsultant.net/2015/08/13/new-roche-app-measures-parkinsons-disease-fluctuations/ Last accessed 07 May 2021.

[22]   Inan OT, Tenaerts P, Prindiville SA et al. Digitizing clinical trials. npj Digit. Med. 2020: 3; 101 (2020). https://doi.org/10.1038/s41746-020-0302-y

[23]   Anderson D, Hefner M. Digital RD 4 Ways To Maximize Patient Engagement In Clinical Trials. Clinical leader. 2018 https://www.clinicalleader.com/doc/digital-r-d-ways-to-maximize-patient-engagement-in-clinical-trials-0001 page 1-5. Last accessed 07 May 2021.

[24]   Filkins BL, Kim JY, Roberts B, et al. Privacy and security in the era of digital health: what should translational researchers know and do about it?. Am J Transl Res. 2016;8(3):1560-1580.