PhUSE Connect 2020

Paper PP20

Machine Learning-AI is the future – What now?

Priyank H Patel, Accenture, Berwyn, USA

ABSTRACT

The pharmaceutical industry is investing vast amount of money in R&D to bring ML-AI driven solutions to enhance overall End-to-End Clinical Trials experience. The ML-AI based driven solutions will change the approach to the current new drug-biologics application submissions to the regulatory authorities however they are still costly and will still take some time to become industry standard. What's the plan in the interim? Within clinical data operations, the industry invests heavily on the global data collection standards to enhance the data collection experience, bringing pseudo-automation into the current process. This helps foster consistency within the data collection process, expedite CDISC SDTM compliant package creation and ease the analysis component down-stream. There are tools available that highlight pseudo-automation with expediated quality outcomes which have already been proven effective. Using the cost benefit analysis, pharmaceutical companies can realize that these tools are an investment into their future vision of ML-AI solutions.

INTRODUCTION

Overall, Pharma & Biotech companies have spent approx. $179bn in 2018-19 on R&D. The R&D costs have been forecast to grow further by approx. 3% from 2018-2024 due to the rising cost of developing new drugs. The drug discovery and development are a long risky road as only a handful drugs get approval from regulatory agencies and realize commercial success. Pharma and Biotech companies' ROI is declining, and the R&D budget has skyrocketed. The healthcare cost is raising inevitably and there is a tremendous pressure from patients and government on these companies to reduce the cost of drugs. Science is evolving every day and so is technology. The industry is taking new initiatives to invest more on R&D programs in alliance with real world data and new technologies like artificial intelligence, machine learning, and natural language processing to help them stay one step ahead in today's era that challenges a more patient targeted drug development. It will be challenging to imagine the data and patient centric framework without a magnifying focus on the data operations component which can pay a strategic role in supporting the clinical research data transformation and ensuring the integrity of the trial data and its outcome.

PRESENT CHALLENGES WITH R&D OUTCOMES

The R&D process is extremely complex in nature mainly due to the source data complexity and data volumes associated which are the key driving factors for increasing cost. Based on a research by Tufts, the soaring cost is due to multiple data points, and different procedures associated with the trials. More trial procedures are being done per visit with an average 70% increase in distinct procedures, and 65% increase in total procedures. Average number of data points increased by 88% with an increased proportion from

Machine Learning-AI is the future – What now?

Priyank H Patel, Accenture, Berwyn, USA

less traditional sources. [1] Tufts &Veeva survey of 257 pharma, biotech & CROs show that companies expect that in 3 years there would a 40-50% increase in the volume of novel data sources including smart phones, EHR/EMR, eSource (direct from instrument/source), mobile & wearables, with a 70% increase expected in eCOA (clinical outcomes assessments) and a 20% increase in social media; with less than 1% of this data managed in the primary EDC application due to system limitations, data integration issues or resources related demands. [2]

There are two main areas that are full of challenges in the Clinical Data Operations group and has become a priority due to unceasing rise in costs.

1. Increasing pressure on the tools and process:
   - Digital solutions present new challenges for research
   - New and complex study design and data collection
   - Volume and complexity of data types and sources are burdening the system
   - The development timelines and the added costs

2. Existing processes are inefficient and burdensome:
   - Significant time is spent enabling systems
   - Most of the actual data cleaning effort often does not result in data change
   - Manual review is needed to supplement programmable solutions
   - Despite high false positive rate, data issues are caught downstream

FUTURE OF CLINICAL TRIALS

As science continues to evolve, innovative therapies will get introduced based on different adaptive study designs practiced in the clinical trial process. The patient centric clinical development will be the focus of the industry to get more desirable outcomes on complex study designs This will require different approaches to manage the clinical data as the patient-centric model will likely be the most disruptive, especially when vastly decentralized. The industry will have to understand the complex data architectures and scale up the underlying technologies to evaluate the meaningful operational value from the data.

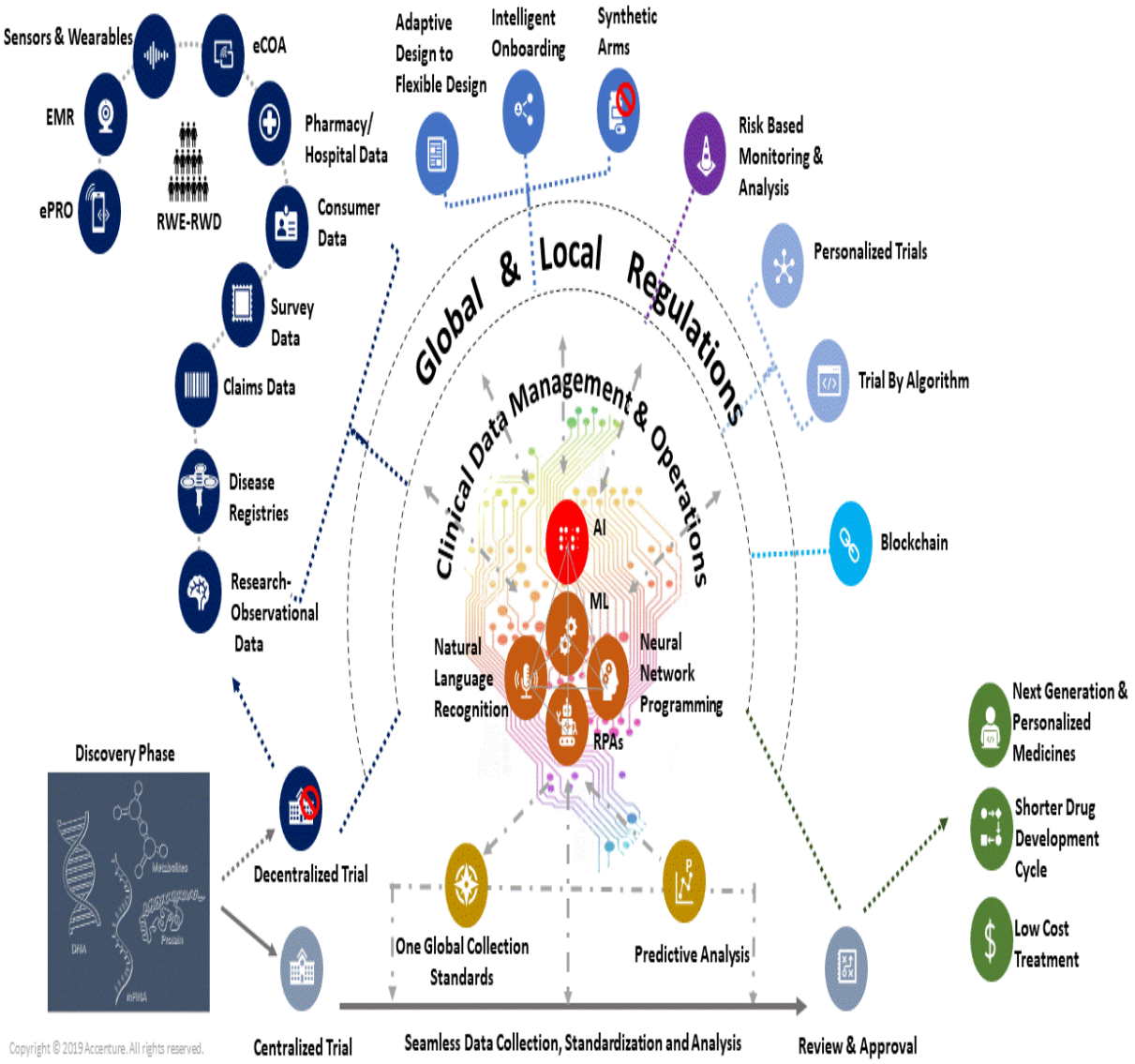
Per Órlaith Burke (Future of Life Sciences Lead at The Dock Senior Principal on Advanced Analytics AI at Accenture), “The research shows that there will be a significant increase in the efficiency and effectiveness of the next wave of clinical trials due to the development of improved technologies, access to real world evidence and data. This should also increase the industry’s capacity to innovate.” [3] This finding was based on a six-week investigation of the topic “The Future of Clinical Trials” with Accenture Life Sciences team from the perspectives of patient and clinician experience, clinical data collection and management, and state of art and technology. The findings were tested in a series of innovation sessions locally and internationally with a wide panel of healthcare experts and numerous clients during the series at The Dock in December 2018. Per the findings and the feedback, it was concluded that the future of the clinical trials

Machine Learning-AI is the future – What now?

Priyank H Patel, Accenture, Berwyn, USA

is arriving in three successive waves: next, near and horizon. [3] Each wave will trigger constructive impact in the certain parts of the eco-system shown below in figure 1.

Figure 1: Future of Clinical Data Management and Operations



Machine Learning-AI is the future – What now?

Priyank H Patel, Accenture, Berwyn, USA

The **NEXT WAVE** is going to bring a big change in the eco-system due to access to improved technologies and real-time clinical data. This wave should boost the industry's capacity to innovate in the space of: [3]

- Immersed in treatment
- Continuous patient data
- Adaptable trials
- Next gen support eco system

The **NEAR WAVE** is going to focus on patient experience, the seamless management of high volume of complex clinical data sources, and data security. Research findings indicate that the virtual trials could become a norm soon. This wave should push the industry's technology to innovate in the space of: [3]

- Intelligent Trial Onboarding
- In Data We Trust
- Seamless Data Management
- Stay at Home Trials

The **HORIZON** indicates a trajectory towards a future where algorithms can simply predict the clinical trial outcomes. With this hope, the regulatory and industry bodies are working together to address to provide one global standard for the data collection and management practices to build sustainable data repositories which can be linked to RWE-RWD to advance in the accuracy of predictive modeling that drives the research of personalized medicine. This wave should push the industry's capability to innovate in the space of:[3]

- Declaration of Data Standards for Global Clinical Trials
- Personalized Trials
- Trial by Algorithm

ACCENTURE CLINICAL DATA OPERATIONS - TECHNOLOGY REVOLUTION

Currently, despite an access to the improved technologies, the industry has witnessed more disparity that leads to bringing more complexity and increasing costs for the data operations due to the lack of standard approaches enabling the efficient development of new processes. The Clinical Data Operations (CDO) group must implement "fit for purpose" and "end to end" data strategies to prevent the critical risks in their current processes. [4] The leaders have acknowledged that even the widely used system such as Electronic Data Capture (EDC) which enables fast data transfer from the sites when compared to paper-based process, is not utilized to its full potential. It was expected that the EDC system will bring more efficiencies in the data management process by reducing the cycle time during the clinical trial run. Research has shown that it has helped reducing the timelines but not per expectations. This raises some valid questions for the industry – should you think of a different approach to conduct the data review

more efficiently? Does the delay really happen due to lack of proper data standards practiced on the research data? Any approach that you consider must involve the viewpoint of all regulatory challenges as well.

The CDO group costs include data management, standardization and analysis effort that signify the tiniest portion of the overall all R&D cost. The CDO group takes the stage when the drug development slowly progresses from the discovery to clinical trial phase and is responsible for managing a pre-clinical, clinical studies' research data in required structure and format based on the study protocol so, it can be effortlessly collected, validated, managed, standardardized and analyzed downstream. With accessibility of new technologies, a data management practice has evolved recently. Pharma and Biotech companies are heavily investing on Metadata Repository (MDR) concept to gain consistency across the trials for their portfolio. This is in a hope to bring solid infrastructure to the research data implementing uniformity and stability by restricting various approaches in handling the same data differently. This approach requires an upfront investment but with a higher expectation to increase the ROI overtime by bringing flexibility to access real time data, consistency in study designing across portfolio, simplifying the demonstration of traceability of data, and providing seamless dataflow that ease the data standardization and analysis effort downstream. This also enables the concept of global clinical data standards which aids in enabling the data aggregation, cross study analysis, efficient regulatory reviews and automizing the end to end data flow.

Let us now focus more on the functional aspects of the CDO group, something every sponsor is working hard to achieve – end to end automation in their current processes. The common submission compliant deliverables are raw data aka source data per trial protocol, study design and other requirements, standardized output in CDISC standards format such as Study Data Tabulation Model (SDTM) outputs, Analysis Data Model (ADaM) outputs and the Table, Listings and Figures (TLF). Significant efforts have been put to automize the submission compliant deliverables. The Pharma and Biotech industries are heavily collaborating with the technology industries to build a better system and a process which allows to bring maximum automation in the business processes. Nowadays, lots of proof of concepts (POCs) are getting conducted in the market to get a better visibility on rate of success.

In any business processes, to leverage benefits of new technologies, the scale and stability of the foundational technologies matter. These become guiding principles to envision the success of the new innovations those are being brought to enable better technology driven solutions and processes.

Paper PP20

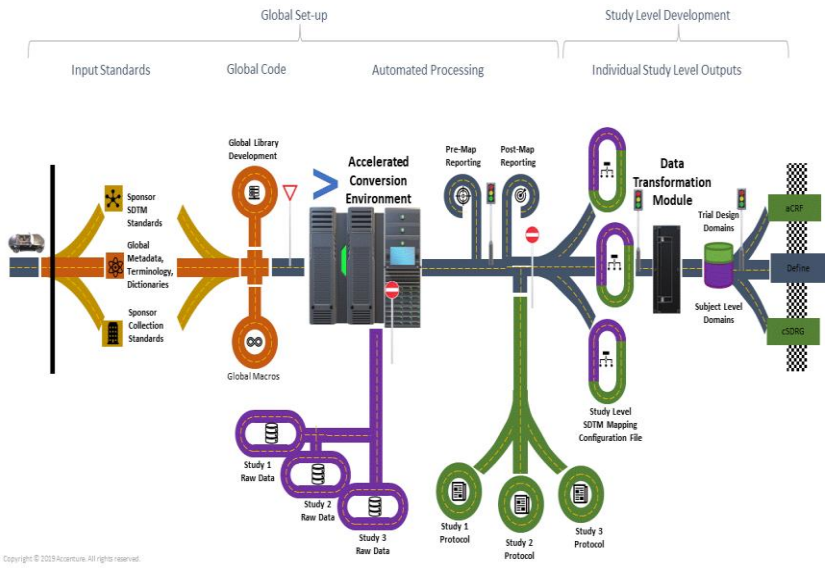
Machine Learning-AI is the future – What now?

Priyank H Patel, Accenture, Berwyn, USA

NEXT WAVE: A GIANT LEAP TOWARDS NEAR WAVE

The Accenture CDO group has been able to set a foundation bringing innovative business solution by helping sponsors to bring pseudo-automation using their global standards to expedite CDISC SDTM submission compliant deliverables with optimum quality and consistency. The CDISC SDTM provides a standard for organizing and formatting clinical data to modernize the collection, management, analysis and reporting processes. CDISC SDTM is one of the required standards for sponsors to implement to electronically submit the trial data to the FDA(USA) and PMDA (Japan). The solution assists conversion activities and provide beneficial tools for developers, to help guide large scale projects for the sponsors. The solution facilitates a proactive approach towards development of CDISC SDTM deliverables. Efficiency, consistency and maintenance via a centralized design control center are key elements to those development benefits. Accenture utilizes global conversion standard libraries provided by the sponsor or produced by Accenture to process and create desired SDTM outputs. Large-scale conversion activities across multiple therapeutic areas require complex data handling, delta reporting abilities for version maintenance and control management to ensure quality sponsor expectations.

Figure 2: Foundational CDISC SDTM Data Conversion Model

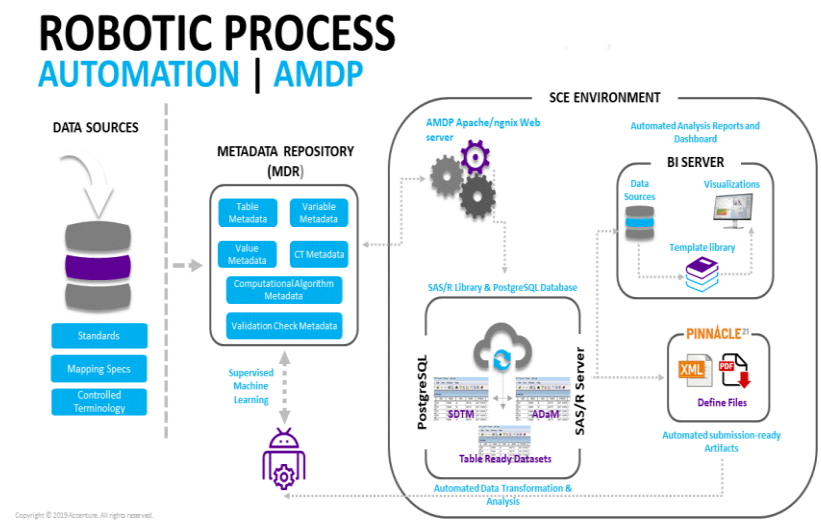


The solution can provide a more fine-tuned large global capacity model that can garner greater consistency and establish groundwork for future development and streamlined review and update methodologies. Developing a global MDR lifecycle system to sustain growth, create efficiencies, increase production and produce a turnkey approach to SDTM conversion services. Furthermore, to build a system that can handle development activities, which are beyond the simple scope of an MDR. Figure 2 provides a high-level MDR foundation that will be the basis for continuous evolutions of next generation solutions to enable access to better technology landscape with greater data connectivity and efficiencies. One of the conceptual models is shown in figure 3. This conceptual model has been prioritized to become a

Machine Learning-AI is the future – What now?

Priyank H Patel, Accenture, Berwyn, USA

Figure 3: Automated Metadata Driven Processing Model



reality. The primary focus of this model is to allow sponsors to build collection and reporting standards across organizational portfolio and bring stability on the continuous data transformation requirements. This can allow sponsors to lower the cycle times, repeated use of submission build artifacts by putting minimum effort and maximum efficiency so the user can quickly make critical decisions based on the science. The POCs are being conducted to analyze the effectiveness in comparison to current era's technology driven outputs. The model conceptualized is to bring end to end process automation where the managed standards library and the code generating engine create fully validated SDTM, ADaM and TLF artifacts.

CONCLUSION

It costs approx. $2.9bn. for Pharma and Biotech industry to successfully develop a new drug and bring it to the market for its commercial use. The current research data suggests that ROI is declining, and the R&D budget is skyrocketing. The next wave, near wave and horizon of the clinical trials, can only be achieved if the Pharmaceutical industry leaders recognize the importance and advantages of pivoting to this exciting future. The accessibility of technology will have a huge impact on everyone involved in the clinical trial process, including the data management and operations group. Data complexity and data volumes are challenging today's business models and driving up the cost of research and development. The AI, ML-driven solutions provide a powerful mechanism that can improve business performance and strengthen the decision-making process. It is critical for the industry to carefully determine the types of tasks, strengths, and limitations of each in order to go through technology upgrades. The foundation pseudo-automation solution offers sponsors to expedite CDISC SDTM submission compliant deliverables with optimum quality and consistency using their data standards. This foundation now has become the incentive to understand the unknowns of the current process and bring further automation with integrating with other novel technologies to offer a better, more efficient product.

Machine Learning-AI is the future – What now?

Priyank H Patel, Accenture, Berwyn, USA

REFERENCES

1. **Tufts University.** Rising Protocol Complexity is Hindering Study Performance, Cost, and Efficiency. s.l. : Tufts Center for the Study of Drug Development, 2018. July/August, Vol. 20 No. 4.

2. **Tufts University**. Clinical Data Volume and Diversity Pose Increasing Challenges and Delays. s.l. : Tufts Center for the Study of Drug Development, 2018. Jan/Feb, Vol. 20 No. 1.

3. **Burke, Órlaith.** The future of clinical trials. s.l. : Accenture The Dock, 2019.

4. **Science, SCDM-The Evolution of Clinical Data Management to Clinical Data.** https://scdm.org/wp-content/uploads/2019/09/SCDM-Reflection-Paper-Evolution-to-Clinical-to-Data-Science.pdf. [Online] 09 2019.

RECOMMENDED READING

1. Accenture The Dock, The Future of Clinical Trials https://www.accenture.com/_acnmedia/PDF-107/Accenture-The-Dock-Clinical-Trials-LifeSciences-Aug19.pdf#zoom=50
2. SCDM, The Evolution of Clinical Data Management to Clinical Data Science https://scdm.org/wp-content/uploads/2019/09/SCDM-Reflection-Paper-Evolution-to-Clinical-to-Data-Science.pdf
3. Innovation in the pharmaceutical industry: New estimates of R&D costs https://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/12742/DiMasi-Grabowski-Hansen-RnD-JHE-2016.pdf;sequence=1
4. FDA - https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence
5. https://info.evaluate.com/rs/607-YGS 364/images/EvaluatePharma_World_Preview_2019.pdf

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Priyank H Patel

Accenture Life Sciences
1160 W. Swedesford Road
Berwyn, PA 19312
priyank.h.patel@accenture.com