Paper RE06

A Predictive Analytics Framework for Diversity-focused Trial Enrolment: Synergistic role of Registry Data, Lived Experiences, and Programming Power

Kapil Khambholja, Genpro Research Inc, Boston, USA
Vatsal Chhaya, Genpro Research, Baroda, India

ABSTRACT

Given the advocacy of clinical trial diversity in the recent past following the FDA guidance, stakeholder involvement has been the buzzword in regulatory decision-making. one of the several effective ways to ensure this includes accelerated use of real-world data, especially in complex disease contexts like rare diseases in oncology. This contains differential effects related to safety and effectiveness in different patient cohorts. Thus, predictive analytics of outcome registry data with stratification of the target population based on age, gender, and ethnicity may prove great trial enrolment insights. The present paper proposes the predictive analytics framework including synergistic use of standardized terms in disease ontologies, AI/ML capabilities, and insights from real-world experiential systems. This paper is the first of its kind to revolutionize current enrolment practices with an opportunity to comply with the latest FDA guidance on a “race and ethnicity diversity plan” for participation in clinical trials.

INTRODUCTION

Clinical trial diversity has gained immense interest from the global clinical research community and regulatory professionals. Underpinning the concepts of bioethics delineated in standard clinical research guidance documents such as Nuremberg’s code and Belmonte’s report, clinical trial diversity considerations ensure that communities of colour or other marginalized populations are appropriately enrolled in any clinical trial. Studies have proven that such populations, due to their intrinsic phenotypical difference from the normal populations, may demonstrate differential treatment responses. For example, the inclusion of communities of colour (CoC) and the oxygen saturation findings in COVID-19 studies, the racial care disparities in multiple sclerosis (MS) patients, and the findings of one of the trials on multiple myeloma patients revealed crucial insights on why DEI considerations in trial enrolment and even overall clinical care are of paramount importance(1–3). Concerningly, the failure in capturing such clinical information influenced by DEI may result in great financial loss and deterioration of the quality of the study design and outcomes.

Propelled by advancements in innovative trial recruitment trends, pharma companies have been inclined towards leveraging predictive analytics capabilities in the clinical trial space. As a result, there are examples of how technology can help achieve this. For example, in past, artificial intelligence and machine learning (AI/ML) have been deployed for predicting the site enrolment period based on a rich set of trial operational data such as patient per site and the number of patients likely to be enrolled(4). However, it is yet to address diversity, equity, and inclusion(DEI) concerns comprehensively. Given the knowledge about real-world data’s breadth traversing all races and ethnicities, a unique DEI-specific framework can be a game changer.

The present paper discusses real-world data’s key characteristics, inputs and potential applications of AI/ML-powered predictive model, challenges ahead, and futuristic action roadmap.

REAL-WORLD DATA: THE “FUTURE”

Unlike clinical trial data, RWD provides pragmatic insights because of several factors: uncontrolled study environment, positive and negative treatment outcomes compared to the trial findings which at times have selective reporting, data collection by patients via medical wearables in some cases, and patient journey touchpoints in real-time on a continual basis.

Ever-increasing collection and use of a large amount of RWD have acted as an impetus for regulatory agencies to develop a data quality framework. In the recent past, European Medicines Agency has released the “first-of-its-kind” two guidance documents explaining the importance of quality in real-world data being collected and describing best practices to use meta-data catalogues from real-world data sources. The data quality framework aims to provide insights on key determinants, the workflows starting from data needs identification to data consumption, data maturity milestones, and quality assessment dimensions and metrics(5).

The future of Real-world evidence can be envisaged with three “V”s: a variety of data from both majority and minority groups, volume capable of providing sufficient insights to decipher intricacies of any disease context, and the velocity at which data acquisition would occur.

INPUTS TO THE MODEL

There will be two types of model input: 1. Quantitative 2. Qualitative. Amongst many, RWD being continuously captured via various outcomes registries include exhaustive information related to patient characteristics and associated clinical outcomes without any restriction based on race or ethnicity. Besides, patient-reported outcomes registries aim to capture the minutest details experienced by patients in real-world settings following the use of any drug or device. Lived experiences captured via newspaper editorials, TV interviews, and open forum discussions-like experiential systems may aid in disease ontology updates and DEI vocabulary updates upon insights from the minority population.

MACHINE LEARNING MODEL – DATA PROCESSING TECHNOLOGY

Promising results from the use of advanced AI/ML techniques such as Natural Language Processing, Optical Code Reader, and Social Graph with ease of data structure and searchability of specific information have been shown in previous literature(6). Also, a study has demonstrated the effectiveness of ensemble machine-learning techniques in improving the eligibility screening of patients(7). A study evaluated machine learning prediction models on clinical trial enrolment(4).

POSSIBLE RESULTS AND APPLICATIONS

The successful roll-out of the proposed AI/ML predictive analytics framework may result in 1. Document Development with DEI-specific auto-text suggestions and 2. Data-driven enrolment prediction with DEI focus besides other things. The application areas for the latter can include auto-identification of DEI-intensive advocacy forums and machine-driven steps for minority involvement during trial enrolment planning, system-generated alerts against DEI gaps in trial documents or trial diversity plan, and automated geo-tagging of regions with a higher likelihood of underserved target population.

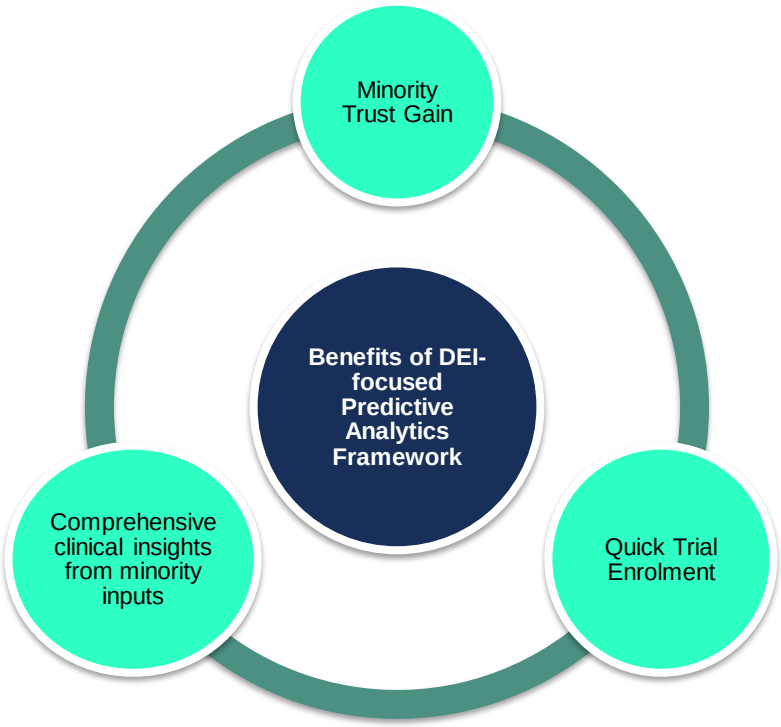


Figure 1 Benefits of DEI-focused Predictive Analytics Framework

Despite their separate agendas and mandates, this predictive analytics framework seeks collective efforts from sponsors, next-generation CRO(tech plus CR), patient Advocacy groups, non-government organizations, academia, for-profit research consultancies, and regulatory agencies.

The below schematic diagram represents the predictive analytics framework in a nutshell.

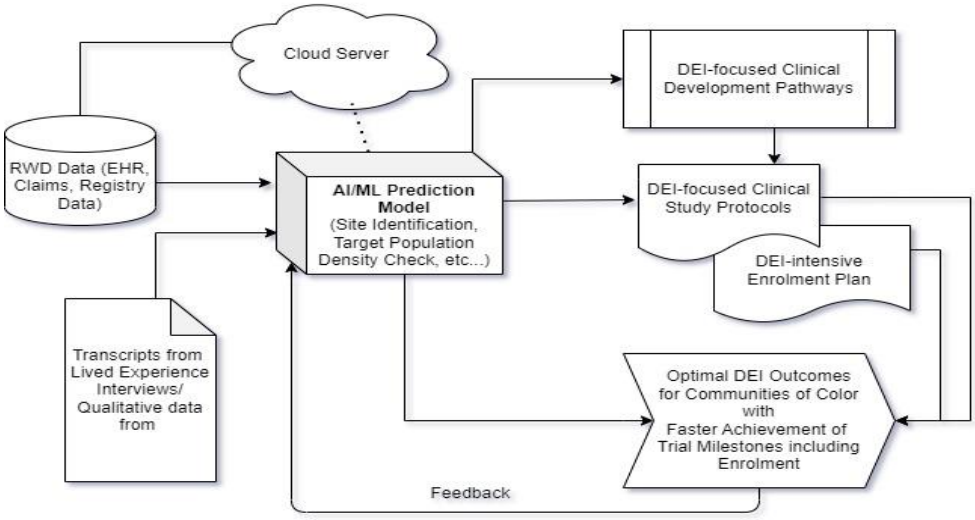


Figure 2 Schematic Diagram of Proposed DEI -focused Predictive Analytics Framework

CHALLENGES AHEAD

Seemingly exciting, the implementation pathway of this model can be tricky and full of challenges. As they say, “easier said than done,” the major issue lies in the core technicalities of the model from a software developer’s angle and overcoming the mental blocks about adopting new prediction-informed enrolment pathways rather than following the conventional enrolment practices.

KNOW-DO GAP

Most of the RWD like healthcare insurance claims data and premium RWD from several organizations is difficult to access and thereby compelling RWD scientists to may need to rely on crude RWD data from hospitals and public registries. This takes a lot of time and effort in data cleaning before it is converted into a usable form. Secondly, even if the data is available, the know-how of identifying and curtailing DEI-specific insight demands an exceptional skill set and an empathetic mindset, making it an even more complex goal to achieve.

PARADIGM SHIFT

Developing a mindset adaptable to the need of the hour is a very tough job. Be it the conceptualization, development, or execution of this model, it requires a lot of patience and a trial-and-error approach with numerous iterations in the coding process. Failure in achieving desired prediction results may lead to “burn-out” and abandonment of the efforts. Besides, some prejudices especially influenced by bias about specific communities may impact the overall enrolment prediction model development and enrolment process itself. Over and above all, different mandates and agendas of stakeholders may result in inter-stakeholder conflicts and differences of opinion.

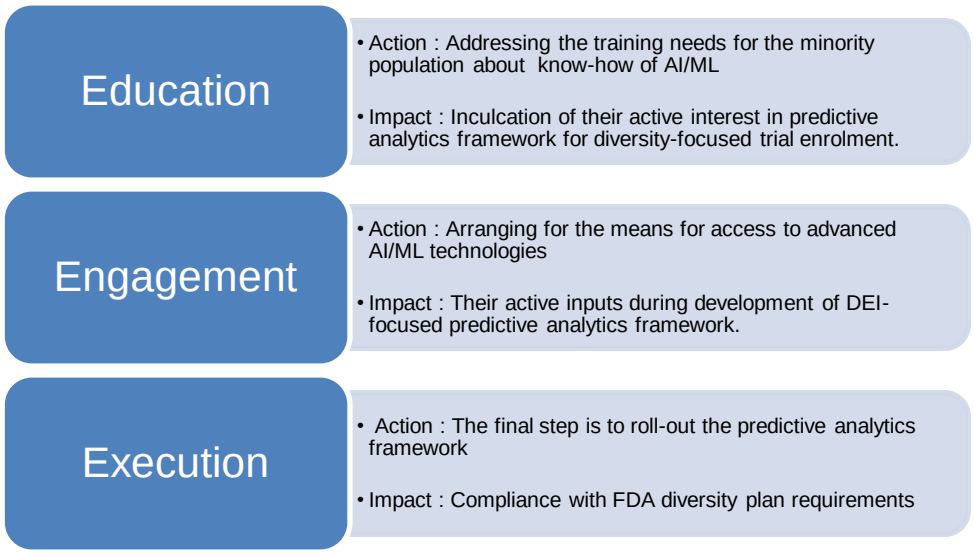
STIGMA AND LACK OF ACTIVE PARTICIPATION FROM MINORITY GROUPS

With an ingrained sense of rejection and seclusion as a result of earlier discriminatory practices, most of the time minority groups remain silent and show reluctance to contribute to the trial enrolment planning. Although careful and compassionate communication may help to break the silos, it is a demanding task.

RECOMMENDED ACTION ROADMAP

An effective roadmap can be proposed in the form of a three-pronged strategy. First and foremost, the requirement is to develop an ecosystem that facilitates the integration of Humans & Technology via 3 “E”s.

.



Another effective step for developing efficient prediction models is to promote interdisciplinary research so that technocrats and DEI champions can join hands to produce optimal solutions. Substantiating the existence with a sense of gratitude and respect is a great gesture to win people’s trust. Instead of keeping them aloof, if the minority advocates are given the role of “clinical trial diversity ambassadors,” they feel motivated and provide meaningful insights for the DEI-centric enrolment strategy. Lastly, the involvement of minority representatives during prediction technology development, its pilots, and its final implementation will help make the product more robust, reliable, and sustainable.

CONCLUSION

To summarize, DEI is a crucial consideration for next-gen trial planning & recruitment. RWD possesses the immense potential to revolutionize the existing trial enrolment landscape. We proposed a multidisciplinary framework using input data from registries and lived experiences and AI/ML techniques for trial enrolment prediction and allied applications. Addressing AI/ML training needs and engagement issues about minority groups will envisage increased participation to improve existing trial enrolment practices. Although existing know-do gaps and paradigm shifts from conventional to prediction-driven enrolment may act as bottlenecks against effective framework implementation, the promotion of interdisciplinary research efforts and offering of ambassadorships to minorities may identify optimal solutions in times to come. We suggest an industry-wide coalition of stakeholders from academia, pharma sponsors, patient groups, data providers/aggregators, and regulators work together on a mission mode.

REFERENCES

1. Kanapuru B, Fernandes LL, Fashoyin-Aje LA, Baines AC, Bhatnagar V, Ershler R, et al. Analysis of racial and ethnic disparities in multiple myeloma US FDA drug approval trials. Blood Adv. 2022 Mar 14;6(6):1684–91.

2. Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. Anesthesiology. 2005 Apr;102(4):715–9.

3. Ca P, Ja L. Racial and ethnic disparities in treatment response and tolerability in multiple sclerosis: A comparative study. Mult Scler Relat Disord [Internet]. 2021 Nov [cited 2022 Oct 21];56. Available from: https://pubmed.ncbi.nlm.nih.gov/34536772/

4. Liu J, Allen PJ, Benz L, Blickstein D, Okidi E, Shi X. A Machine Learning Approach for Recruitment Prediction in Clinical Trial Design [Internet]. arXiv; 2021 [cited 2022 Oct 21]. Available from: http://arxiv.org/abs/2111.07407

5. EMA. High-quality data to empower data-driven medicines regulation in the European Union [Internet]. European Medicines Agency. 2022 [cited 2022 Oct 21]. Available from: https://www.ema.europa.eu/en/news/high-quality-data-

empower-data-driven-medicines-regulation-european-union

6. How NLP, OCR, and Social Graphs Accelerate Subject Recruitment for Clinical Trials [Internet]. Applied Clinical Trials Online. [cited 2022 Oct 21]. Available from: https://www.appliedclinicaltrialsonline.com/view/how-nlp-ocr-and-social-graphs-accelerate-subject-recruitment-for-clinical-trials

7. Cai T, Cai F, Dahal KP, Cremone G, Lam E, Golnik C, et al. Improving the Efficiency of Clinical Trial Recruitment Using an Ensemble Machine Learning to Assist With Eligibility Screening. ACR Open Rheumatol. 2021 Jul 23;3(9):593–600.

CONTACT INFORMATION

Author Name: Kapil Khambholja
Company: Genpro Research Inc
Web: www.genproresearch.com