International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 7.2 Website: https://ijarmt.com ISSN No.: 3048-9458

Artificial Intelligence–Driven Risk-Based Monitoring Framework for Phase I–III Clinical Trials

Dr. Mihir Patel
Master of Public Health, Monroe University, NY, USA
drmihirpatel19@gmail.com
Dr. Noynika Shukla
(M.B.B.S.), M.S. Ramaiah Medical College, Bengaluru, India
noynika29@gmail.com
Dr. Hemanshi Dhaduk
M.S. in Health Care Administration, Monroe University, NY, USA
hemanshi624@gmail.com

Abstract—As have rapidly developed electronic health records and artificial intelligence, we are presented with new opportunities for enhancing the efficiency and accuracy of clinical trial monitoring. This research presents a Recurrent Neural Networks (RNN)-based, AI-enabled risk-based monitoring system using MIMIC-III data. This section comprises various preprocessing stages, such as handling missing values, one-hot encoding, and min-max scaling, followed by train-test splitting and model implementation. The RNN model is structured to learn the time-dependent relationship of clinical data to predict risk. Experimental results show improved performance, with accuracy (acc) of 92.08, precision (prec) of 94.8, recall (rec) of 96.6, F1-score (F1) of 95.7, and AUC of 0.98. The comparative analysis reveals that the proposed model outperforms conventional machine learning techniques, including MLP, XGBoost, and SVM. The results indicate that the suggested method can be successfully used to enhance the efficiency of early risk identification and patient safety during clinical trials and could serve as a potential solution for real-time AI-based monitoring systems.

Keywords—Artificial Intelligence (AI), Risk-Based Monitoring (RBM), Clinical Trials, Healthcare Analytics, MIMIC-III, Predictive Modeling.

Introduction

The growing use of Electronic Health Records (EHRs) has revolutionized the healthcare systems through the massive digitization of patient records. For seriously ill patients who may be able to recover or who are in danger of dying, the intensive care unit (ICU) provides continuous monitoring and delivers targeted interventions and appropriate treatment. The ICU has received a lot of attention from the medical world as a result of its high fatality rate and pricey resources [1][2]. The gold standard for assessing a treatment's effectiveness and safety is a clinical study, particularly a randomized controlled trial (RCT) [3]. Regardless of the methods employed for drug development and discovery, therapeutics and vaccines must undergo three stages before being authorized for public use and mass production. For example, clinical trials are carried out by the US Food and Drug Administration [FDA] to assess their safety and effectiveness [4].

Given diversity of clinical trial outcome targets explored in prior scientific work, and the inherently sequential and interdependent nature of stage-specific success metrics across the clinical drug development lifecycle, this study first establishes a structured definition of clinical development success by explicitly delineating its core metrics, intended to serve as predictive targets within an overall structured modeling framework [5][6]. Specifically, as shown in Figure 1, success is decomposed into (i) operational success, reflecting the ability to complete the execution of the clinical trial as planned, (ii) scientific success, capturing whether the trial meets its primary endpoint(s) with statistically significant and clinically relevant results, (iii) phase transition, representing advancement of pharmaceutical programs to the next clinical stage, and

(iv) regulatory approval, corresponding to successful market authorization

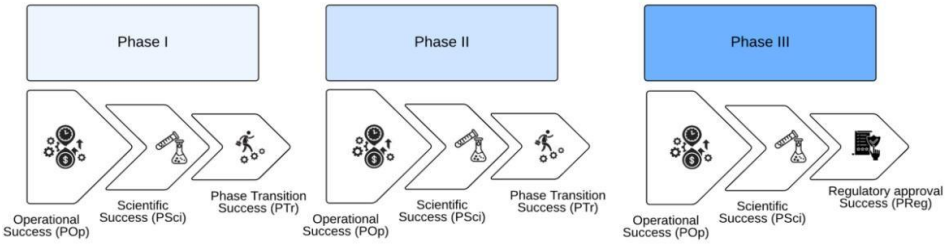

Fig 1: Clinical Drug Development Lifecycle (Phase I–III).

Schematic overview of the sequential and interdependent success metrics considered in this study. Clinical development success is decomposed into four phase-specific predictive probabilities: the probability of operational success (POp), the probability of scientific success (PSci), the probability of phase transition (PTr), and the probability of regulatory approval (PReg), enabling phase-specific modeling and longitudinal risk assessment from early development through downstream milestones [7]. Researchers may either reduce AI risks in clinical trials by acknowledging that the model may do worse than human readers, or they might aim to enhance model performance to equal human performance and throw all risk on the AI model alone [8][9]. The aim of the scientific and technology discipline of AI is to build intelligent machines. AI techniques include Deep Learning (DL), Optical Character Recognition (OCR), Machine Learning (ML), and Natural Language Processing (NLP) [10]. Clinical trials are changed by this cutting-edge technology, which is among the most current. AI development in the healthcare industry has a solid technological foundation due to rapidly development of information technology and the enormous amount of biological data that has been gathered [11]. Researchers are investigating Al applications to reduce the complexity and risk of clinical trials while improving the efficacy of medical diagnostics and service quality.

The rising complexity of clinical trials and the growing volume of electronic health records have rendered older monitoring methods ineffective, time-intensive, and more error-prone. Patient safety and the ability to identify risks as early as possible are essential issues, particularly in multi-phase clinical trials. Thus, there is a great need for intelligent, automated, and datadriven solutions capable of analyzing temporal clinical information. The proposed study is inspired by the potential of AI, specifically DL, to improve risk prediction and monitoring accuracy and to facilitate timely decision-making in a clinical trial setting, using datasets such as MIMIC-III. The following are the main contributions to this study:

- Introduces a clinical trial risk-based monitoring system based on AI. Learns an RNN to learn temporal relationships in clinical records.
- Conducts extensive data preprocessing, such as missing values, one-hot encoding, and normalization.

- Provides excellent results in terms of acc, prec, rec, F1, and AUC. compares the results against traditional models such as MLP, XGBoost, and SVM.
- Early risk detection and patient safety are shown to be improved through clinical trial monitoring.

The remainder of this paper is organized as follows. Section II reviews the related work on AI and risk-based monitoring in clinical trials. Section III describes proposed methodology, using MIMIC-III dataset. Section IV provides experimental findings, performance assessment, and a comparison with current techniques. Finally, Section V concludes paper with key findings and outlines directions for future work. The limitations of traditional clinical trial monitoring methods, which rarely provide insight into the complex temporal patterns and dynamic risk factors in patient data, justify the proposed study. Using an RNN trained on MIMIC-III data, the proposed framework enables more accurate and timely risk prediction. This work is novel in that it incorporates DL-based temporal modeling, a well-structured preprocessing pipeline, and a risk-driven monitoring policy, achieving better forecasting than traditional ML methods. Also, the framework focuses on early risk identification and real-time applicability, which is why it can be considered an effective and scalable solution for contemporary AI-driven clinical trial monitoring systems.

Literature Review

The analysis and monitoring of clinical trials have been significantly enhanced by current advancements in ML and AI.

Y. Yuan et al. (2023) conducted a multi-center clinical trial to collect infant videos for investigating the feasibility of proposal. A benchmark including four machine learning methods of SVM, KNN, MLP, and CNN (ResNet18) was set up to classify the sleep/awake stage of infants. To alleviate the overfitting issue caused by over-sampling of a sleeping infant, propose to integrate ResNet18 with the contrastive learning strategy to strengthen the consistency of facial features learned from different infants. The clinical evaluation shows that all benchmarked methods obtained an accuracy above 75% while the best accuracy of 86% was attained by suggested approach [12].

Similarly, X. Wang et al. (2023) assessed the ability of seven benchmark transformer-based models—BERT, ALBERT, BioBERT, BioClinicalBERT, RoBERTa, ELECTRA, and RoBERTa-MIMIC-Trial—to recognize SDoH terms using a corpus of MIMIC-III clinical notes that had already been annotated. According to both strict and flexible criteria, research shows that the BioClinicalBERT model scores best on F1 (0.911, 0.923) [13]. In another study, F. Di Martino et al. (2021) estimated daily intake of the main dietary components, meal completeness, and variability using feature engineering, while accounting for physiological data. Next, using ground-truth values from Mini Nutritional Assessment rating scale, trained several ML models. The dataset imbalance was addressed by using SMOTE and cost-sensitive learning. The findings indicate that median rec and acc values of top ML models for predicting risk of malnutrition were 92% and 94%, respectively [14]. Furthermore, X. Liu et al. (2020) suggested utilizing DL to extract eligibility-criteria variables from COVID-19 trials, trial design and optimization may be quantitatively analyzed. To extract entities (i.e., variables) from COVID19 trial eligibility requirements, train an Att-BiLSTM model using the top-performing model. It evaluates Att-BiLSTM's performance against a conventional ontology-based approach. Using a benchmark dataset, Att-BiLSTM performs better than the ontology model. Att-BiLSTM achieves prec of 0.942, rec of 0.810, and F1 of 0.871, whereas ontology model only achieves an accuracy of 0.715, rec of 0.659, and F1 of 0.686 [15].

E. R. Hutchison et al. (2020) As part of an effort to harmonize the Aggregate Analysis of ClinicalTrials.gov (AACT) drop-withdrawal reasons to a controlled vocabulary, explored two solutions. Elastic's fuzzy matching capability matched entries in the AACT Drop-Withdrawal table to a list of user-specified terms (74.6% coverage). The second approach was a custom pipeline employing NLP preprocessing, Levenshtein Distance (Fuzzy Matching), and semantic similarity mapping using a pre-trained FastText Model (98% coverage). Although manual oversight is still required, the amount of effort to harmonize with a controlled vocabulary is notably reduced [16].

L. Tong et al. (2019) concentrated on death occurrences as the study's predictive aim. It gathered 28,340 reports from Clinicaltrial.gov and converted information into vectorized ML features. These features were standardized across investigations using semantic mapping and feature selection techniques. Five ML models for prediction were constructed using the chosen clinical trial variables. Evaluated and contrasted each model's performance on the prediction challenge. According to results, logistic regression approach got highest overall receiver operating characteristic score (0.7344) [17]. The existing literature is primarily concerned with individual clinical tasks as opposed to the creation of an integrated risk-based monitoring model of clinical trials. The majority of methods are based on the use of static models and cannot be used to detect the temporal relationships within the patient data, which restricts their effectiveness in constant monitoring. Also, the problems of absence of real-time prediction, less generalizability, and inability to integrate heterogeneous clinical data are still unsolved. Thus, an AI-based, time-conscious model capable of making precise and scalable risk forecasts based on data such as MIMIC-III is required (ICU type). **Methodology**

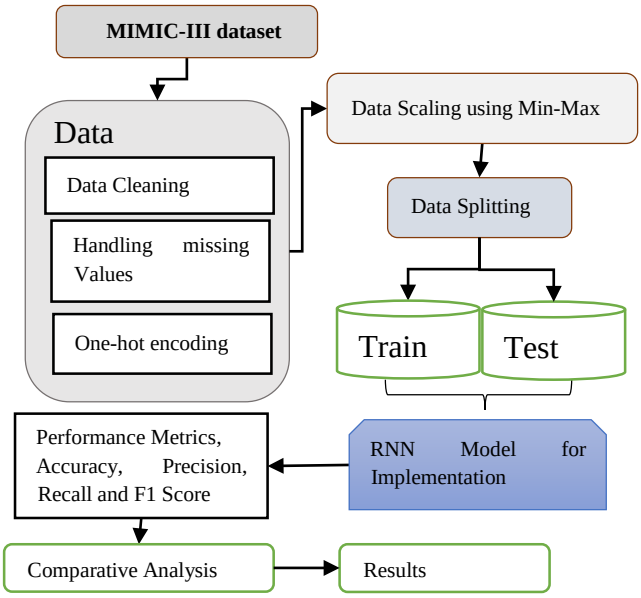

Fig 2: Proposed Flowchart for Clinical Trial

The methodology begins with data extracted from the MIMIC-III dataset, which is then preprocessed through data cleaning, missing-value handling, and encoding one-time. The data is then normalized using min-max scaling and divided into training and testing sets. An RNN model is used to capture temporal patterns and predict risk. Performance measures (acc, prec, rec, and F1) are used to evaluate model and then compared with others to confirm the functionality of proposed method. The whole methodology is illustrated in Figure 2.

Data Collection

The experimental setup utilizes Medical Information Mart for Intensive Care MIMIC-III collection, which includes de-identified medical records for more than 40,000 patients in critical care. The present dataset is very comprehensive and comprises time-series vitals,

unstructured clinical notes and sensitive demographic attributes. The final processed dataset consists of 30,000 samples, with each sample containing 20 temporal features (e.g., heart rate, blood pressure, and oxygen saturation), tokenized embeddings of clinical notes, and 5 demographic attributes (age, gender, ethnicity, and admission type).

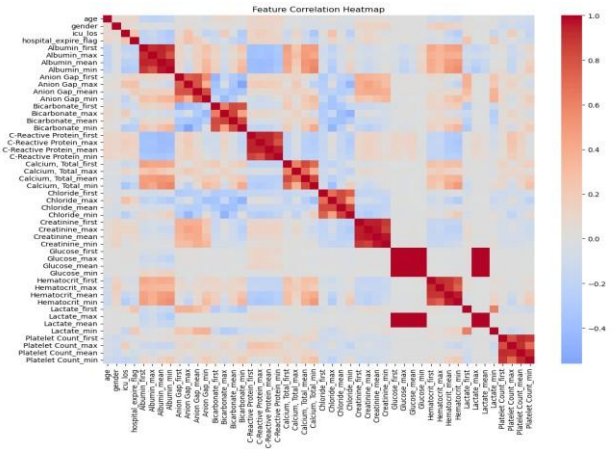

Fig 3: Feature Correlation Heatmap of the Dataset

Figure 3 shows a correlation matrix of features of MIMIC-III dataset. The cells indicate correlation coefficient of pairs of clinical variables, with darker colors depicting the stronger relationships of positive or negative associations. The heatmap identifies strong correlations between some physiological and laboratory characteristics that can be used to affect model learning. This analysis facilitates storing the redundant features, comprehending the feature dependencies, and enhancing model performance by making good selections of features.

Data Pre-processing

The MIMIC dataset was used for feature engineering, concatenation, and cleaning. Among the pre-processing procedures were normalization and handling of missing variables. The main pre-processing steps are summed up as follows:

- **Handle Missing Values:** missing values are handled by context, all correlated events are set to zero. This maintains the structure of the input matrix as well not introducing any bias through imputation. This approach preserves the structure of the input matrix, and also avoids imputation bias.
- **Convert text to lowercase:** Convert the whole text to lower case In this step, all different forms of the same words like "Lung" and "lung" are normalized into a single token which reduces the redundancy and promotes consistency across the dataset.

One-Hot Encoding

Essentially, one-hot encoding is a data preparation technique that converts categorical variables into a numerical format for use by ML algorithms. With this method, you will convert each feature category into a separate column and assign a rate to it, 1 indicates its presence, and 0 indicates its absence.

Data Scaling

Data scaling is the process of altering the range of feature values without altering the actual data, is an essential part of preprocessing. Min-max scaling is a common technique for scaling data that involves transforming data values to lie within a predetermined range, often between 0 and 1. The transformation is defined as in Equation (1):

$$X = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (1)$$

where X is original value, and x_{min} and x_{max} represent feature's lowest and highest values, accordingly. This normalization increases the effectiveness of the suggested model and guarantees that each characteristic has an equal contribution to learning.

Train-Test Split

The dataset must be split into training and test sets once data processing is complete. 80% of dataset is utilized for training, while remaining 20% is used for testing.

Proposed Recurrent Neural Network (RNN) Model

The proposed architecture employs an RNN to learn sequence clinical information, which is capable of learning temporal relationships in patient health records. The RNN has a hidden state, which is updated during each time step in accordance with current input and state of the previous time step, and enables the model to learn time-dependent patterns that are important in risk predictions. The basic functioning of the RNN is provided in Equation (2).

$$h_t = \sigma(W_h h_{t-1} + W_x x_t + t) \qquad (2)$$

where W_h and W_x are weight matrices, b is bias, x_t is the input, h_t is hidden state, and σ is activation function. The Adam optimizer is used to train the model, which has a learning rate of 0.001, a 32-batch size, 20 epochs, 128 hidden units, and a dropout rate of 0.2-0.5. Such a structure enables the related domain to be effectively learned, generalized and improved in predictive accuracy for clinical risk-based monitoring.

Model Evaluation

The model is also tested using standard classification metrics to confirm that it can be reliably evaluated for risk prediction in clinical trials. Accuracy is the percentage of samples that have been appropriately identified of all samples, and is used to estimate the overall accuracy of predictions. A measure of precision shows the proportion of expected positives that are true, or how likely the model is to minimize the number of FP. Recall (Sensitivity) is rate of number of TP, that are recognized, which is crucial in clinical practice to avoid missing at-risk conditions. The model's performance may be properly assessed using F1 (the harmonic mean of acc and rec), which is particularly helpful when working with unbalanced datasets. It is computed by applying the equations (3-6) below:

The number of accurately anticipated positive cases is represented by TP (e.g., correctly identified high-risk patients), while TN denotes correctly predicted negative cases. FP represents the cases when the model errs in predicting positive instances (false alarms), and FN points to the cases when the real positive cases are missed.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad \square\square\square$$

$$Precision = \frac{TP}{TP + FP} \qquad \square\square\square$$

$$Recall = \frac{TP}{TP + FN} \qquad \square\square\square$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad \square\square\square$$

Also, the AUC of the ROC curve is used to evaluate the model's capacity to differentiate between classes at various thresholds.

Results, Discussion & Comparative Analysis

The experiments of this study are carried out in a Dell Inspiron laptop operating on Windows 11, powered by an Intel, Core™ i7-11800H processor. The model training and assessment are implemented in Python in the Google Colab environment.

Table I shows the results of the suggested RNN-based model that was tested on MIMIC-III dataset. The model performs well overall in categorization, as seen by its high accuracy of 92.08. It achieves a prec of 94.8 and a rec of 96.6, meaning it effectively identifies actual risk

cases while minimizing FP and FN. The model's robustness is demonstrated by an F1 score of 95.7%, indicating a reasonable balance between accuracy and recall. Also, an AUC of 0.98 indicates strong discriminative ability. These results validate the efficacy of the suggested risk-based monitoring approach for AI in clinical trials.

TABLE 1: PROPOSED RESULTS OF CLINICAL TRIAL ON MIMIC-III DATASET

Measures	**RNN**
Accuracy	92.08
Precision	94.8
Recall	96.6
F1 Score	95.7
AUC	0.98

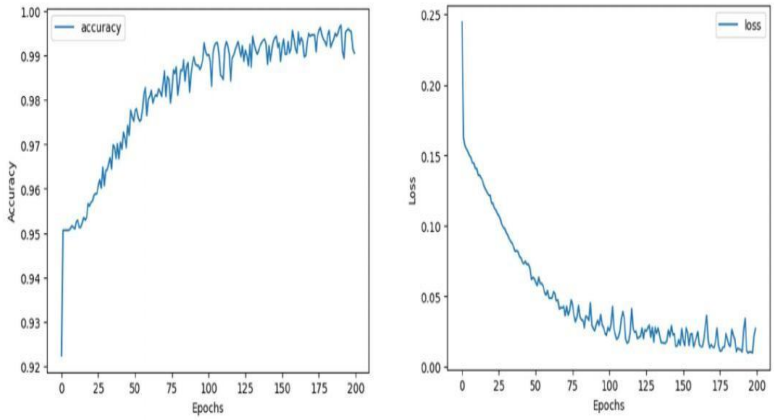

Fig 4: Training Accuracy and Loss Curves of the Proposed RNN Model

The suggested RNN model's learning behavior during training epochs on the MIMIC-III dataset is depicted in Figure 4. Accuracy in left plot demonstrates a gradual increase with an accuracy which implies a successful model learning and convergence. The right plot shows a steady diminution of loss, close to near zero, which is a confirmation of minimized prediction error. All these trends indicate stable training, good convergence and the ability of the model to produce high predictive performance with minimal overfitting.

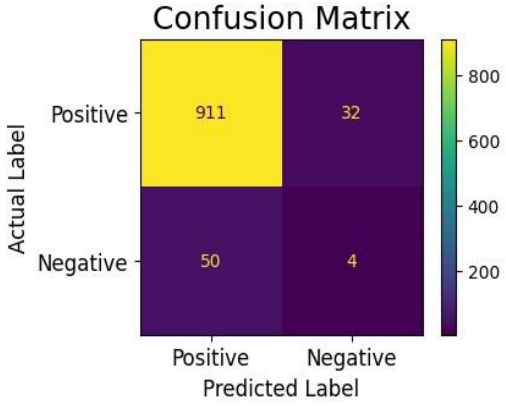

Fig 5: Confusion Matrix of the Proposed RNN Model

The confusion matrix of suggested RNN model evaluated on the dataset is displayed in Figure 5. The model effectively classifies with 911 positive cases (True Positives) and 4 negative cases (True Negatives), which is good classification performance. But 32 false

positives and 50 false negatives are also noted, implying that there are small misclassifications, especially on the negative cases.

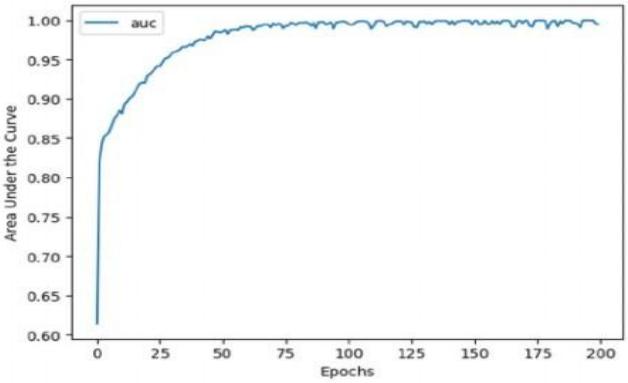

Fig 6: AUC Curve of the Proposed RNN Model

Figure 6 illustrates how the suggested RNN model's Area Under the Curve (AUC) varies across training epochs using the MIMIC-III dataset. The AUC value grows quite fast and the value reaches about 0.62 and then approaches almost 1.0, which means that there is a great enhancement in ability of model to discriminate between positive and negative classes. This supports fact that model is very efficient in the classification of risks, thus it can be used to ensure accurate AI-driven clinical trial monitoring.

Comparative Analysis

Table II shows a comparison of proposed RNN-based model with other methods that are currently in use to predict clinical trials: MLP, XGBoost, and SVM. The proposed RNN performs better than all the existing models, with the highest accuracy (92.08%), precision (94.8%), recall (96.6%), F1-score (95.7%), and AUC (0.98). Conversely, the classic models like MLP and SVM show moderate performance, whereas XGBoost demonstrates relatively low performance in some of the measures. The high performance of the RNN shows that it is effective in the representation of temporal dependencies in clinical data and thus is more appropriate for accurate and reliable risk-based monitoring in clinical trials.

TABLE 2: COMPARISON OF PROPOSED MODELS WITH EXISTING METHODS FOR CLINICAL TRIALS

Measures	**RNN**	**MLP**[18]	**XGBoost**[19]	**SVM**[20]
Accuracy	92.08	75	73.01	76
Precision	94.8	-	-	75
Recall	96.6	0.65	-	78
F1 Score	95.7	-	31.21	-
AUC	0.98	0.72	63.92	86

The proposed model is an RNN-based framework aimed at monitoring risk in clinical trials using MIMIC-III data. The model can capture temporal dependencies and sequential trends in patient data and accurately predict risk conditions. Compared to conventional models, the proposed RNN achieves higher acc, prec, rec, F1, and AUC. Its major strengths are its ability to detect risks early, better management of time-series clinical data, fewer false predictions, and decision support for monitoring trial participants. Also, the model is better suited for real-time monitoring of clinical trials and patient safety because it is more generalizable and stable.

Limitations of this Study

This study has made good performance, it has a number of limitations. First, model is tested on MIMIC-III dataset, which may not generalize well to other clinical trials and populations. Second, the suggested RNN model, being time-dependent, might not represent long-term dependencies as effectively as more sophisticated architectures such as LSTM or transformer models. Third, the research largely uses structured and somewhat processed data, with minimal use of unstructured clinical notes and multimodal data sources. Moreover, the model is not explainable, and thus, predictions are hard to interpret in critical healthcare situations. Lastly, no real-time deployment and validation in actual clinical trial settings have been investigated, which is necessary for practical adoption.

Conclusion & Future Work

The adoption of AI in healthcare has enhanced the effectiveness and consistency of clinical trial monitoring in the past few years. This paper introduces a risk-based risk monitoring framework that is based on AI and implemented with an RNN on MIMIC-III dataset. The suggested model works well, obtaining a high F1-score, AUC, recall, accuracy, and precision, and is helpful for identifying temporal connections in clinical data. The experimental findings and comparative research prove that, compared to conventional ML models, RNN model is more efficient and a trustworthy solution for early risk detection and enhanced patient safety in clinical trials. The strategy also ensures strong pre-processing, strong feature processing and convergence of the model all of which work towards better predictive performance. In the future, the given framework can be extended with even more advanced DL models, i.e., LSTM, GRU, or transformer-based models, which can further improve the results. In addition, it is possible to combine real-time streaming data and multi-modal inputs (medical imaging and genomic data) in order to increase accuracy and generalizability of predictions. To increase model interpretability and confidence in clinical decision-making, explainable AI techniques such as SHAP and LIME can also be explored. Furthermore, its applicability in large-scale, real-time risk-based monitoring systems would be more appropriate by testing the model using a variety of real-world clinical trial data and applying it in a real-world healthcare environment.

References

[1] R. Liu, H. Liu, L. Li, Z. Wang, and Y. Li, "Predicting in-hospital mortality for MIMIC-III patients: A nomogram combined with SOFA score," *Medicine (Baltimore).*, vol. 101, no. 42, p. e31251, Oct. 2022, doi: 10.1097/MD.0000000000031251.

[2] R. Snehamrutha, "Patient Engagement Strategies in Community Pharmacies and their Effect on Vaccination Uptake and Medication Synchronizations," *ESP J. Eng. Technol. Adv.*, vol. 3, no. 3, pp. 163–173, 2023, doi: 10.56472/25832646/JETA-V3I7P120.

[3] J. A. Kachhia, "Healthcare Predictive Analytics based on Machine Techniques for Identifying Cardiovascular Risks Screening," *Int. J. Curr. Eng. Technol.*, vol. 13, no. 6, pp. 635–642, 2023, doi: 10.14741/ijcet/v.13.6.17.

[4] Z. He, A. Erdengasileng, X. Luo, A. Xing, N. Charness, and J. Bian, "How the clinical research community responded to the COVID-19 pandemic: An analysis of the COVID19 clinical studies in ClinicalTrials.gov," *JAMIA Open*, 2021, doi: 10.1093/jamiaopen/ooab032.

[5] K. Larson *et al.*, "COVID-19 interventional trials: Analysis of data sharing intentions during a time of pandemic," *Contemp. Clin. Trials*, vol. 115, p. 106709, Apr. 2022, doi: 10.1016/j.cct.2022.106709.

[6] A. Warrier, “Real-Time Healthcare Event Processing: Stream Analytics for Clinical Decision Support,” *Int. J. Emerg. Res. Eng. Technol.*, vol. 1, no. 4, December, pp. 47–54, 2020, doi: 10.63282/3050-922X.IJERET-V1I4P106.

[7] A. Darwiesh, A. H. El-Baz, A. Z. Abualkishik, and M. Elhoseny, “Artificial Intelligence Model for Risk Management in Healthcare Institutions: Towards Sustainable Development,” *Sustainability*, vol. 15, no. 1, p. 420, Dec. 2022, doi: 10.3390/su15010420.

[8] M. D. McCradden, S. Joshi, J. A. Anderson, and A. J. London, “A normative framework for artificial intelligence as a sociotechnical system in healthcare,” *Patterns*, vol. 4, no. 11, p. 100864, Nov. 2023, doi: 10.1016/j.patter.2023.100864.

[9] A. Warrier, “COVID - 19 Contact Tracing : Privacy - Preserving Integration Architectures for Public Health Surveillance,” *Int. J. AI, BigData, Comput. Manag. Stud.*, vol. 2, no. 1, pp. 88–96, 2021, doi: 10.63282/3050-9416.IJAIBDCMS-V2I1P109.

[10] S. Pachiappan, S. P. Samundi, and A. Gnanasambandam, “Artificial Intelligence in Clinical Trials- Future Prospectives,” *Bioequivalence Bioavailability. Int. J.*, vol. 7, no. 1, pp. 1–8, jun. 2023, doi: 10.23880/beba-16000196.

[11] A. Muley, P. Muzumdar, G. Kurian, and G. P. Basyal, “Risk of AI in Healthcare: A Comprehensive Literature Review and Study Framework,” *Asian J. Med. Heal.*, vol. 21, no. 10, pp. 276–291, Aug. 2023, doi: 10.9734/ajmah/2023/v21i10903.

[12] Y. Yuan *et al.*, “A Multi-center Clinical Trial for Camera-based Infant Sleep and Awake Detection in Neonatal Intensive Care Unit,” in *2023 IEEE International Conference on E-health Networking, Application & Services (Healthcom)*, IEEE, Dec. 2023, pp. 319–322. doi: 10.1109/Healthcom56612.2023.10472347.

[13] X. Wang, D. Gupta, M. Killian, and Z. He, “Benchmarking Transformer-Based Models for Identifying Social Determinants of Health in Clinical Notes,” in *2023 IEEE 11th International Conference on Healthcare Informatics (ICHI)*, IEEE, Jun. 2023, pp. 570–574. doi: 10.1109/ICHI57859.2023.00102.

[14] F. Di Martino, F. Delmastro, and C. Dolciotti, “Malnutrition Risk Assessment in Frail Older Adults using m-Health and Machine Learning,” in *ICC 2021 - IEEE International Conference on Communications*, IEEE, Jun. 2021, pp. 1–6. doi: 10.1109/ICC42927.2021.9500471.

[15] X. Liu, L. A. Finelli, G. L. Hersch, and I. Khalil, “Attention-Based LSTM Network for COVID-19 Clinical Trial Parsing,” in *2020 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2020, pp. 3761–3766. doi: 10.1109/BigData50022.2020.9378451.

[16] E. R. Hutchison, Y. Zhang, S. Nampally, J. Weatherall, F. Khan, and K. Shameer, “Uncovering Machine Learning-Ready Data from Public Clinical Trial Resources: A case-study on normalization across Aggregate Content of ClinicalTrials.gov,” in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, Dec. 2020, pp. 2965–2967. doi: 10.1109/BIBM49941.2020.9313362.

[17] L. Tong, J. Luo, R. Cisler, and M. Cantor, “Machine Learning-Based Modeling of Big Clinical Trials Data for Adverse Outcome Prediction: A Case Study of Death Events,” in *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, IEEE, Jul. 2019, pp. 269–274. doi: 10.1109/COMPSAC.2019.10218.

[18] J. Gligorijevic *et al.*, “Optimizing clinical trials recruitment via deep learning,” *J. Am. Med. Informatics Assoc.*, vol. 26, no. 11, pp. 1195–1202, Nov. 2019, doi: 10.1093/jamia/ocz064.

[19] E. Kavalci and A. Hartshorn, “Improving clinical trial design using interpretable machine learning-based prediction of early trial termination,” *Sci. Rep.*, vol. 13, no. 1, p. 121, Jan. 2023, doi: 10.1038/s41598-023-27416-7.

[20] X. Jiang, W. Dai, and Y. Cai, “Comparison of machine learning algorithms to SAPS II in predicting in-hospital mortality of fractures of the pelvis and acetabulum: analyzes based on MIMIC-III database,” *All Life*, vol. 15, no. 1, pp. 1000–1012, Dec. 2022, doi: 10.1080/26895293.2022.2125448.