ORIGINAL RESEARCH ARTICLE

AI-Enhanced Strategies for COVID-19 Vaccination and Booster Prioritization: A Comprehensive Framework

Sara AlShaya[1], Aji Gopakumar[2*], Rahul Raghupathy[3], Badshah Mukherjee[4], Bachir Abou Elias[5], Sheik Jamal Mohideen[6], Vibhor Mathur[7], Sudheer Kurakula[8]

[1,2,3,5,6,7]Emirates Health Services (EHS), Dubai, UAE
[4]SAS Middle East FZ, Dubai, UAE

DOI: 10.55489/njcm.151220244270

ABSTRACT

Background: Distributing vaccines efficiently during the COVID-19 pandemic presented significant logistical challenges. To address the need for identifying populations at risk of breakthrough infections and those requiring booster shots, Emirates Health Services (EHS) developed a framework utilizing AI-driven digital solutions. Objective: To develop a machine learning (ML) model to identify individuals at risk of breakthrough infections and in need of booster doses, aiming to prioritize booster administration and reduce repeated infections among fully vaccinated individuals in the Northern Emirates of the UAE.

Methods: A monitoring dashboard was developed using the EHS Intelligence (PaCE) platform. The study, conducted in three phases, created models to predict infection risk, COVID-19 severity in ICU patients and breakthrough infection risk, using data from the Wareed EMR system.

Results: The AI models accurately identified high-risk individuals and predicted ICU mortality, achieving AUCs of 75% and 74% for infection risk, 94% and 91% for ICU mortality in training and validation datasets, observed 79% AUC with 85% accuracy for identifying high-risk groups for booster vaccination.

Conclusion: The integration of AI in vaccination prioritization demonstrated its potential to enhance public health initiatives and improve pandemic management in the UAE.

Keywords: AI-augmented EHS Intelligence platform, Machine learning in public health, COVID-19 vaccination, Booster dose, Data analytics, UAE healthcare innovation

ARTICLE INFO

Financial Support: None declared
Conflict of Interest: None declared
Received: 04-06-2024, **Accepted**: 22-10-2024, **Published**: 01-12-2024
***Correspondence:** Dr. Aji Gopakumar (Email: aji.gopakumar@ehs.gov.ae)

How to cite this article: AlShaya S, Gopakumar A, Raghupathy R, Mukherjee B, Abou Elias B, Jamal Mohideen S, Mathur V, Kurakula S. AI-Enhanced Strategies for COVID-19 Vaccination and Booster Prioritization: A Comprehensive Framework. Natl J Community Med 2024;15(12):1013-1028. DOI: 10.55489/njcm.151220244270

www.njcmindia.com | pISSN: 0976-3325 | eISSN: 2229-6816 | Published by Medsci Publications

INTRODUCTION

The COVID -19 vaccinations have been pivotal in managing the pandemic by significantly reducing severe illness, hospitalizations, and deaths. As the virus continues to mutate, booster doses have become increasingly important to sustain immunity, particularly against new variants. Both the initial vaccines and subsequent boosters are vital not only for protecting individuals but also for achieving wider public health goals.[1] It is important to consider the high-risk population who consists of individuals more prone to severe outcomes when exposed to COVID-19. Giving priority to high-risk groups including the older age group, those with comorbidities, and frontline health workers is essential as these population groups are most vulnerable to severe consequences from COVID-19. Vaccinating and boosting these groups can directly decrease hospitalizations and mortality, addressing the most immediate and serious effects of the pandemic. On the other hand, infection risk often raises considerably after a certain period. Data from the National Israeli database highlights the infection risk notably rises 6-months post immunization, indicating that immunity from the vaccine may wane over time, leading to a higher susceptibility to infection.[2,3] This highlights necessity for effective distribution of booster doses to maintain robust immunity and continued protection against COVID-19.[4,5]

Prioritizing COVID-19 booster doses for high-risk groups was increasingly relevant as the virus continues to evolve and immunity wanes over time. Evidence shows that while initial vaccination offers significant protection, the effectiveness of vaccines can diminish, particularly in older adults and those with underlying health conditions. Consequently, booster doses are essential to reinforce immunity, lower the risk of severe/mortality outcomes, and curb the transmission of the virus. By focusing on high-risk populations, the overall effectiveness of vaccination strategies can be enhancing and safeguard those most vulnerable to severe disease. One of the studies has modeled COVID-19 supplementary doses and antiviral medications in 4 middle income countries, incorporating factors such as age, comorbidities, vaccination and infection status. The findings highlight that administering booster doses to at-risk groups is essential, while oral antivirals are expected to have a significant impact across all settings.[6] Emerging evidence suggested that booster shots are effective in mitigating infections, one research discuss about reduction of symptomatic Omicron infections. With the limited availability of bivalent or dual-component booster doses[7], it is essential to identify and prioritize those at highest risk of severe breakthrough infections. Targeting these high-risk populations allows for more efficient distribution of booster doses and enhances protective measures for the most vulnerable individuals.

While there was extensive research on vaccine prioritization during early phase of the COVID -19 vaccine[8-10], application of AI-driven models for identifying risk groups for vaccination and booster prioritization remains rare. There is research focused on creating and validating an AI powered digital assistant designed to reduce hesitancy in vaccine intake, using the Medical research councils (MRC's) framework as a guide.[11,12] AI-powered chatbots and thermal imaging have played a significant role in pandemic management by providing guidance to large groups and helping to enforce public health measures such as social distancing and lockdowns.[13,14] Artificial Intelligence has been increasingly used in research and development of drugs for COVID-19 infection[15], these AI-driven tools and computational strategies have significantly expedited the process of developing vaccines to combat the virus.[16] Studies emphasized the crucial role of artificial intelligence in enhancing clinical trials related to vaccines and drug development through various technological approaches, however there are less studies that shows application of AI models in risk group classification for effective vaccination and booster dose delivery.

Literature shows machine learning models can facilitate biomarker predictions for COVID-19, offering clinicians a clear and interpretable view of how various risk factors influence outcomes.[17] in addition, AI models have the potential to identify high-risk patients early on, analyze its epidemiology and simulate spread of diseases through trained data.[18-19] An AI framework can process vast amounts of data to assess individuals and classify them into risk levels ranging from no risk, minimal-moderate to high risk and the identified risk group can be quarantined earlier, reducing the likelihood of virus transmission.[20] In the early stage, different studies indicated that elderly COVID-19 patients, those exhibiting symptoms like fatigue and expectoration, and individuals with common comorbidities such as hypertension, diabetes, and coronary heart disease, are at a higher risk of progressing to severe disease.[21-23] Later after the emergence of AI in risk identification, machine learning models predicted chronic obstructive pulmonary diseases (COPD), cardiovascular and chronic kidney diseases (CVD and CKD), Type-2 DM, hypertension, malignancy and asthma as key determinants of COVID-19 death rate, with their gender and age being the most significant predictors. It's also highlighted important symptom-comorbidity links, such as Pneumonia with Hypertension and Diabetes, underscoring high-risk groups and aiding in the prioritization of resources.[24] Effective allocation of vaccination and booster doses is essential, as research indicates that the patients requiring a booster shot are greatly influenced by various factors, including the local incidence of the disease, the severity of outcomes, and individual's risk factors for developing moderate to severe illness etc.[25] Therefore, the current study aims to present high-risk populations for COVID-19 infection to ensure effective vaccination coverage across the northern emirates. It also showcases a developed model for prioritizing booster doses to reduce breakthrough infections, which are cases of COVID-19 occurring in individuals who have been fully vaccinated.

Objectives

Emirates Health Services (EHS) is a leading public healthcare system in the UAE, comprising 17 hospitals and over 100 healthcare centers, dedicated to delivering high-quality, patient-centered care. With the pandemic posing unprecedented challenges, especially in predicting infection risks and managing healthcare resources effectively, EHS aimed to leverage AI and machine learning to develop precise, interpretable models. These models would guide vaccination strategies, ensuring that vulnerable groups receive timely and appropriate care, including booster doses, to minimize breakthrough infections and optimize public health outcomes across the region.

In response to the challenge of identifying high-risk groups for COVID-19 infection and prioritizing vaccination, EHS developed highly interpretable machine learning models that predict infection risk based on symptoms and patient characteristics. The primary objective was to create sustainable AI models to enhance healthcare outcomes across the UAE, focusing on effective vaccination strategies and prioritizing booster doses for those most vulnerable to breakthrough infections.

Accordingly, the article focuses on the objectives to establish the power of analytical methods in identifying the high-risk population for COVID-19 infection for effective vaccination coverage across northern emirates and to exhibit the model developed for booster prioritization to reduce breakthrough infection, which refers to the occurrence of COVID-19 in individuals who have been fully vaccinated against it.

METHODOLOGY

This article is methodological, focusing on the implementation of advanced analytical techniques adopting AI models designed to identify high-risk populations for effective COVID-19 booster doses. It emphasizes the procedures and techniques applied in the organization to tackle the booster prioritization, rather than presenting statistical data on vaccination/booster rates and its prioritization factors. Machine learning models are applied for faster and effective decisions within the EHS healthcare entity during the pandemic. By exploring the methodologies employed, the paper aims to provide insights into the success of our framework and its design and implementation, showcasing how it was used to enhance vaccination strategies and response efforts. The research also includes the necessary ethical and administrative approvals (MOHAP/DXB-REC/MJJ/ No. 58/ 2022 and MOHAP/ DXB-REC/ JJJ/ No. 69/ 2022), affirming its adherence to established standards.

To ensure the framework's relevance and accuracy, insights and experiences from physicians were integrated. These healthcare professionals contributed their practical knowledge and observations, which were essential for contextualizing the data and refining the framework. This combination of data-driven analysis and clinical expertise aimed to create a robust and practical model for improving health service delivery and patient outcomes. The framework was meticulously developed by analyzing retrospective data collected from Emirates Health Services (EHS) facilities, with a primary focus on data obtained from the Wareed platform. This data provided a comprehensive historical view of patient interactions and healthcare outcomes. Initial findings were explored through a pilot study at a single facility. Building on this foundation, an advanced AI-driven machine learning model was created for high-risk infection group in general, evaluating various deep-learning algorithms such as random forest, gradient boosting, Neural networks, decision trees, etc. This model was instrumental in the effective management of COVID-19 predicting higher risk group for infection and ICU related mortality. In the next phase of study, it resulted in a highly interpretable model for prediction of breakthrough infection and booster prioritization. These three models observed with higher accuracy and as significantly enhancing prognosis, prevention, and control within the UAE based on the EHS data. The implementation of these innovative models led to notable improvements in user experience, clinical excellence, and financial outcomes, thereby advancing the overall quality of healthcare services and benefiting the well-being of the healthcare workforce.

Phase-wise Implementation of AI Models: This study conducted in three phases considering the development and integration of three advanced AI models with the EHS intelligence (PaCE) platform to create a dashboard for rapid clinical decision-making. Each model addresses a distinct aspect of COVID-19 management. In Phase 1, the study focused on identifying risk of COVID-19 infection in the various population groups of UAE, following extensive screening and the development of a Risk Scoring Mechanism. In Phase 2, the focus shifted to predicting mortality risk in ICU settings by identifying key severity factors. Finally, in Phase 3, a model was developed to prioritize vaccination and booster administration.

The **first model** focuses on predicting the likelihood of a COVID-19 infection in individuals based on a comprehensive dataset from a nationwide cohort in the United Arab Emirates. The study collected retrospective data from encounters reported by physicians, which was then input into a Patient Under Investigation (PUI) form during COVID-19 screening processes. The key variables used were age, gender, body temperature, comorbidities, and patient symptoms. These variables were fed into machine learning algorithms, ultimately creating a predictive model.

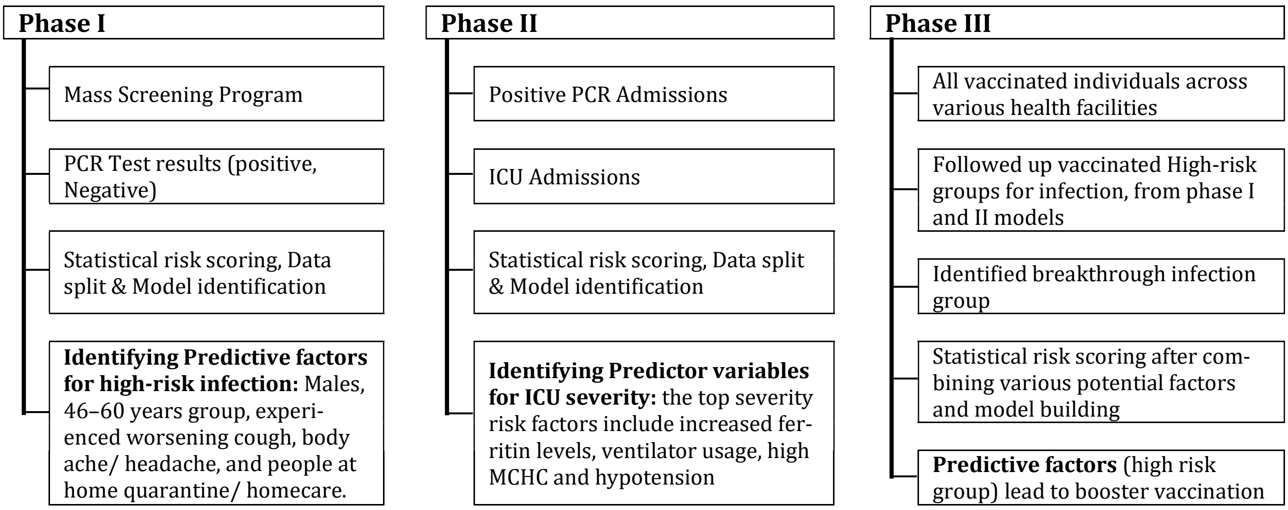
The data was instrumental in developing an infection risk score model, which aggregated patient information to generate a risk score for different population groups that used for classifying level of risk (high, medium, low). The Gradient Boosting Model, chosen for its parsimony and explainability, demonstrated sufficient predictive value, helping the categorization of higher risk populations for quarantine, diagnostic testing and exposure tracking. In other words, the study found Gradient Boosting Model as the optimal model for risk group prediction, which is a machine learning technique used for regression and classification tasks. This model builds a collection of decision trees sequentially, where each tree attempts to correct the errors of the previous ones. These models combine the strengths of multiple weak models to create a more accurate and robust final model, are particularly effective in handling complex data sets like covid infected data and vaccination details with numerous variables and interactions.

The **second AI model** was designed to predict the mortality risk of COVID-19 patients in ICU settings including 71 variables. A sophisticated risk-scoring mechanism was combined with its data attributes to develop this model. The dataset included various patient metrics, and multiple machines learning algorithms, including Gradient Boosting, Random Forest, Ensemble, Decision Tree, Neural Network, and Logistic Regression, were evaluated. The Gradient Boosting Model outperformed the other models, identifying key risk factors for mortality, such as increased ferritin levels, ventilator usage, high MCHC and hypotension. These findings highlighted the model's ability to accurately predict fatal outcomes, providing valuable insights for clinical decision-making in critical care settings.

The **third model** (foster) was developed to prioritize vaccination and booster shots for patients identified as high-risk based on the first two models. By integrating the risk scores generated by the infection and mortality prediction models, the new AI model helped determine which patients should receive vaccination or booster shots first compared to other population groups. This approach ensures that those most vulnerable to severe outcomes from COVID-19 are protected promptly, maximizing the effectiveness of vaccination campaigns.

Overall, the study demonstrated the potential of AI-driven risk scoring mechanisms in enhancing COVID-19 care. The infection risk model provided a robust tool for early identification of high-risk individuals, enabling targeted interventions such as testing and isolation. The mortality risk model in ICU settings offered critical insights into patient management, aiding in the allocation of resources and improving survival rates. Finally, the vaccination prioritization model ensured that the most vulnerable populations received timely protection, optimizing public health efforts. Together, these three models form a comprehensive framework for managing the pandemic, offering scalable and data-driven solutions that could be applied to future infectious disease outbreaks.

Research Framework



The vaccination model integrates diverse data sources from EHS facilities, Private, Pure Health, SEHA, Al Tamouh and External Facilities, including vaccination history, demographic information, and infection rates, which improves its predictive accuracy. This comprehensive approach enables more effective stratification of the population based on risk factors. By identifying high-risk groups, key severity factors and optimizing vaccination coverage strategies, the AI model enables better resource allocation for ensuring booster doses are administered efficiently, reducing healthcare costs and improving patient outcomes. The model supports a highly interpretable dashboard on the EHS Intelligence platform, includes a real-time feedback mechanism to capture user characteristics and leader decisions, which helps in continuously refining the model and its implementation. This ensured that the model remains relevant and effective over time during the entire period of pandemic. Hence the new application enhances the precision of risk identification, improves resource management, and supports evidence-based policy development, offering significant advancements over existing AI models in vaccination prioritization.

Variables and Data collection Process: The model consists of multiple variables including vaccination details such as location of administration, emirates, facilities, sectors, number of doses administered, booster administration details, and patient characteristics such as Age, gender, nationality, comorbidities, etc.

The data collection process was initiated after obtaining approval from REC and EHS. In the phase 1 project, the infection risk model was created using identified data from the Wareed EMR system of EHS facilities. The inclusion criteria targeted adult patients who had a documented encounter with a COVID-19 PCR test within the five months of 2021 pandemic period. Using a random sampling technique, the study analyzed 23,996 encounters, representing 22,865 unique individuals tested for COVID-19. To build the model, a variety of input variables were considered, including patient factors (such as age, gender, health history, recent exposure to COVID-19 positive individuals, and travel history), type of visit (encounter type and admit mode), and a range of disease symptoms (abdominal/stomach pain, abdominal bleeding, body ache, diarrhea, difficulty in breathing, fatigue, headache, muscle pain, history of dry cough, history of hospitalization, nausea, shortness of breath, running nose, sore throat, vomiting, and weakness). Additionally, specific COVID-19 symptoms reported by patients were included.

In phase II, the study population included COVID-19 patients admitted in the ICUs of EHS facility, specifically among the patients who were COVID-19 positive and both gender and nationality. Data from confirmed cases was gathered from January 2021 to May 2021 to encompass the midpoint of the pandemic. During the study period, it was expected to have approximately 800 COVID-19-infected cases per study settings. This population-based data consisted of probable predictors of COVID-19 severity in critical care. The researchers proposed to include all the COVID-19-positive cases reported in the Wareed platform. The study included the main outcome variable as binary, with categories “patients survived/ not” in ICU due to COVID-19 infection. The model considered more than 70 characteristics as explanatory variables. These independent variables include categorical and quantitative types such as patient demographic characteristics, ventilator use, severity status, vaccination status, other clinical parameters and so on. From the literature, important and potential factors were identified and included in the AI model. Other variables were selected as per the nature of ICU facilities and patient experience in the care units of EHS facilities. The missing values for the labs and clinical variables were imputed. The data quality checks showed a minimal discrepancy in patient demographics, comorbidities, medications, laboratories, and clinical attributes, including using a defined and reviewed feature set.

The further study cohort for vaccination strategies (phase III) included individuals who had received at least one dose of the COVID-19 vaccine, with a focus on all adults /residents of the various locations of UAE, recorded underlying health conditions (diabetes, heart disease, respiratory illnesses). Exclusions were made for individuals with incomplete data or non-residents. The methodology for this project involved an extensive data collection process that integrated multiple data sources from various facilities, including EHS facilities, Private entities, Pure Health, SEHA, Al Tamouh, and External Facilities. The focus was on vaccination administration across a diverse range of sectors, encompassing the Aviation Sector, Education Sector, Economic Activities Sector, Government Sector (including MOHRE, government healthcare workers, and others), Military Sector, General Transportation (such as Buses and Taxis), Police Sector, Public Sectors, Labor Sector, Security Sector, Tourism and Hospitality Sector, and Travel Sectors. This comprehensive approach ensured a wide and representative cohort, enabling a thorough and effective analysis of the vaccination data. However the final AI model developed including data from Wareed, which is in integration with EHS Intelligence Platform, and real time feedback from healthcare staff and local leaders. Though the study methodology began by defining the population to include all vaccinated individuals across various health facilities, this broad approach allowed for a comprehensive overview of vaccination outcomes within diverse healthcare settings. Subsequently, the study focused on a follow-up with vaccinated individuals identified as high-risk groups based on the findings from the Phase I and Phase II models. These models previously identified individuals with a higher probability of COVID-19 infection and mortality, respectively, by analyzing factors such as age, comorbidities, and initial infection risk. The next step involved identifying breakthrough infection groups, which included individuals who became infected with COVID-19 despite receiving the full recommended two doses of the vaccine. By closely examining these cases, the study aimed to uncover patterns, combinations of risk factors and commonalities among those who experienced breakthrough infections. This analysis led to the identification of significant factors contributing to breakthrough infections among the high-risk group, which in turn, informed the recommendation for booster vaccinations. Factors such as the emergence of new variants, waning immunity over time, and individual health conditions were scrutinized to understand their impact on the efficacy of the initial vaccination. This thorough evaluation allowed for a nuanced understanding of the need for booster doses, thereby optimizing the allocation of booster vaccinations to those most likely to benefit from them.

Initially in phase III, the project aimed to include the entire population of vaccinated individuals across various locations of UAE, but later, model development focused solely on data from the Wareed EMR system which includes detailed and complete data on clinical parameters of the individuals visited different facilities and centers of EHS. This shift might exclude certain groups or individuals and lead to selection bias. The study's mention of missing values in comorbidities being imputed may lead to potential biases in the results. Despite these challenges, the project aimed to refine its AI-driven model continuously, ensuring accurate identification of high-risk individuals and effective vaccination strategies. Overall, while the study attempts to integrate diverse data sources and maintain a comprehensive scope, all the related potential biases were carefully considered and addressed in the analysis and interpretation of the findings to ensure accurate and reliable conclusions.

Analytical Methods: To effectively implement vaccination strategies during the pandemic, a rigorous scientific approach was necessary. Previous studies had shown that there was no universal standard for predictive models capable of prioritizing vaccination groups at different stages, including initial doses and boosters. This highlighted the need for developing a robust data analytical method using advanced machine learning technologies. Such an approach was crucial in creating an AI model that could accurately identify patient characteristics associated with high-risk groups for COVID-19 infection, serving as a critical first step in optimizing vaccination efforts.

The analysis began with the estimation and exploratory data analysis of health indicators, focusing on patient characteristics and healthcare parameters. Multiple machine learning algorithms, including decision tree-based models, Random Forest, and neural networks etc, were developed to predict high-risk population groups for immediate vaccine administration. The analysis further extended to the estimation of health indicators like infection rate, vaccination rate and severity/mortality risk rates. As part of model evaluation, the model's discriminatory power was assessed using the Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) with model performance comparisons made across these metrics. The overall infection and vaccination results, including key insights from the champion model, were included in the analytical results, highlighting trends in vaccination and booster administration with major predictors of booster prioritization.

Missing data was a significant challenge during the investigation and follow-up, it was effectively managed through improved tracking systems. The dashboard developed includes visual presentation of Covid-19 vaccination administration trends, broken down by location, year, UAE emirates, facility/entity, patient gender, nationality, and months, among other factors. It also includes predictions of health indicators generated by various machine learning models, along with a table comparing model performance. The results of the champion model are highlighted, showcasing the misclassification rate across training, testing, and validation datasets, ROC curve, AUC, confusion matrix, accuracy level, variable importance highlighting key predictors for booster vaccination, and partial dependence (PD) plots for deeper insights. In nutshell, this AI-driven dashboard on the EHS intelligence platform was built to provide district-level insights into COVID-19 vaccination and booster allocation data, enhancing transparency and visibility for health sector stakeholders. It supports the monitoring of high-risk groups and key performance indicators through advanced SAS-based analytics and visualizations, ensuring that evidence-based practices are shared and implemented across EHS facilities.in addition, a real-time feedback mechanism was implemented to gather input from staff, physicians, and leaders, that used for the Model Improvement.

Stages of Stakeholder Engagement: The project aimed to establish sustainable AI-driven models to predict key health indicators and deliver clinical solutions that benefit the entire country. To ensure the project's success and demonstrate our commitment to social responsibility, we actively involved all relevant stakeholders in each phase of implementation.

Healthcare providers played a central role by continuously updating, uploading, and verifying patient records through the Wareed online platform. This ensured that patient data remained accurate and up to date, which was critical for identifying those at highest risk for vaccination prioritization. The project involved extensive coordination, with statisticians and data scientists from The Data and Statistics Department (DSD) conducting regular meetings with key stakeholders to share progress and ensure that evidence-based practices were implemented across EHS facilities. DSD in collaboration with public health professionals, clinical doctors, and hospital administrators, coordinated for the project to identify groups most "at risk" for COVID-19 vaccination and boosters based on a variety of health indicators. The department took the initiative to organize and conduct regular training sessions for staff, positioning themselves as the key oversight body to ensure the reliability of vaccination data. This was essential for facilitating effective decision-making within the health system. Given the vast amount of data EHS handles, a technology-driven approach was essential. This solution allowed for precise analysis, with vaccination statistics being displayed on an advanced, visually intuitive dashboard. The dashboard provided a clear representation of trends, patterns, and predictive rates, making the data accessible and easily interpretable for quick decision-making. The SAS-based EHS intelligence platform enabled users to monitor current statistics and projected outcomes via dashboard, supporting timely interventions in healthcare delivery.

EHS's excellence in cost efficiency was reflected in the program that developed a prediction model to compact healthcare expenses and effective resource utilization in all the EHS facilities. An innovative AI approach was formed to foresee the risk group for vaccination priorities which significantly enhance patient outcome and reduced losses in healthcare facilities. These projects helped the EHS entity to set targets and develop strategies for policy recommendations on vaccination administration and to improve timely access for the general public to healthcare settings. The developed AI models empowered the EHS providers to plan the necessary resource planning, and care strategy for improved patient outcomes20-21.

RESULTS

The current paper aimed at showcasing an effective data analytical framework for quick decision making and highlights our technological achievements in vaccination prioritization. The result part summarizes the AI model's performance, including presentation of vaccination details across UAE emirates and key factors like dose type and patient characteristics. The machine learning models developed in the EHS healthcare system have significantly improved patient outcomes by identifying high-risk populations for vaccination, optimizing coverage across northern emirates, and targeting strategies for vulnerable groups. These models have also prioritized individuals for booster doses to mitigate breakthrough infections and assessed the timing for boosters based on increased infection rates following the second dose.

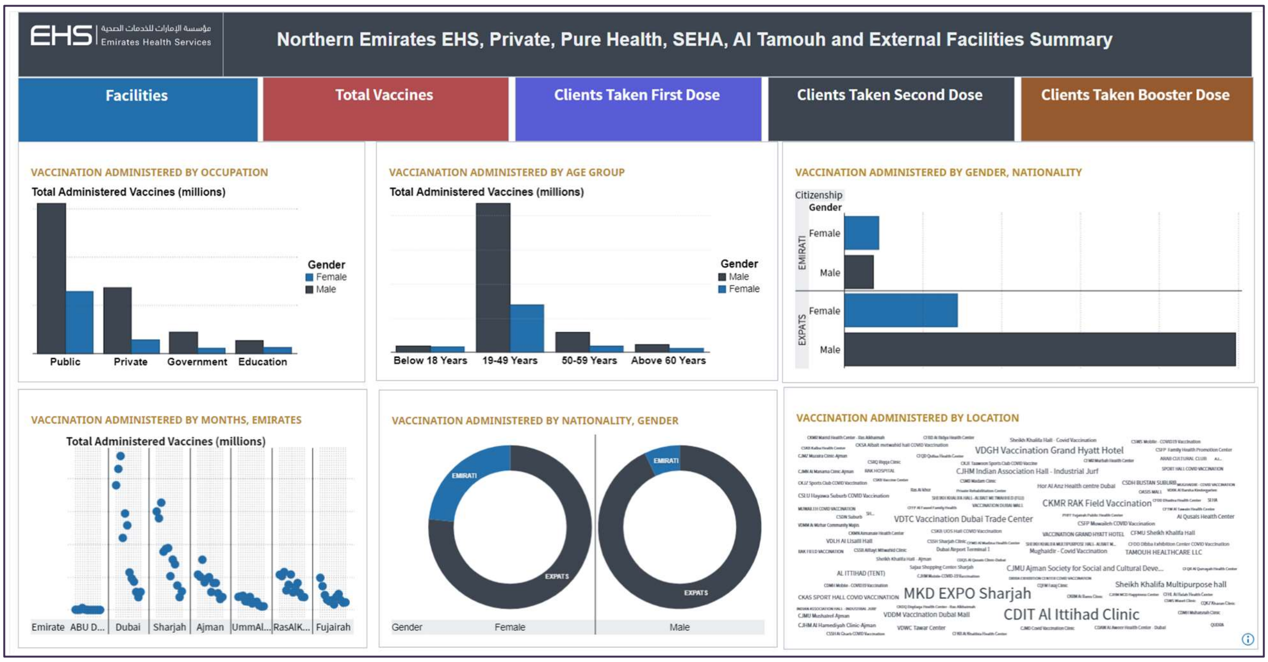


Figure 1: Covid-19 vaccination Profile

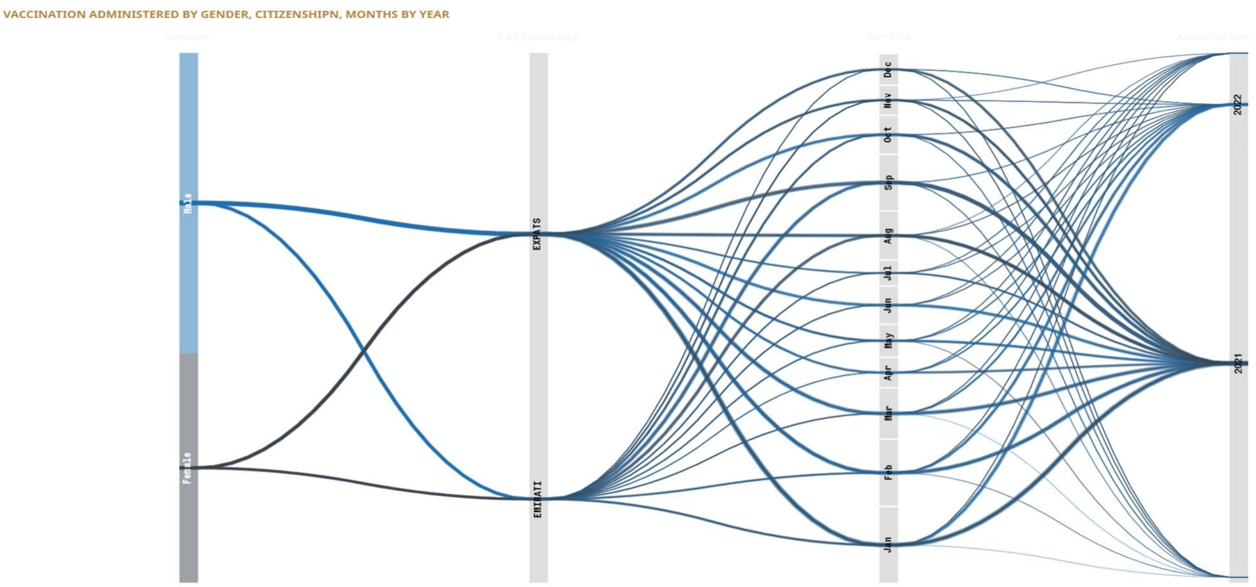


Figure 2: Covid-19 vaccination administration Trend – by Gender, Nationality and Months/year

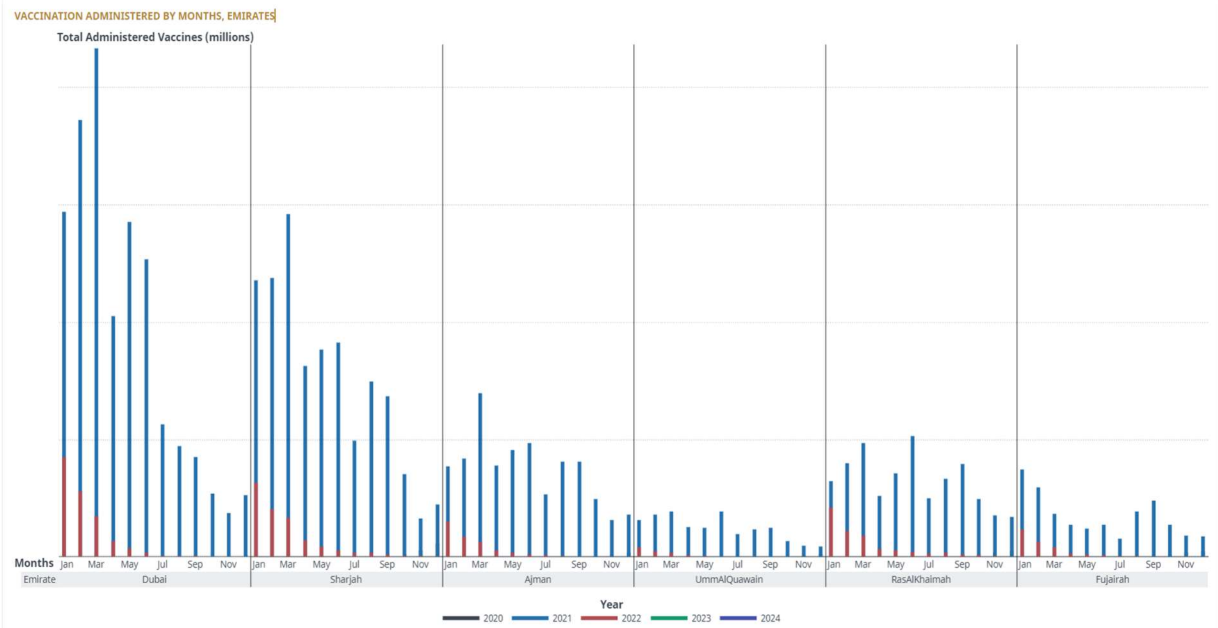


Figure 3: Covid-19 vaccination administration Trend – by Emirates and Months/Year

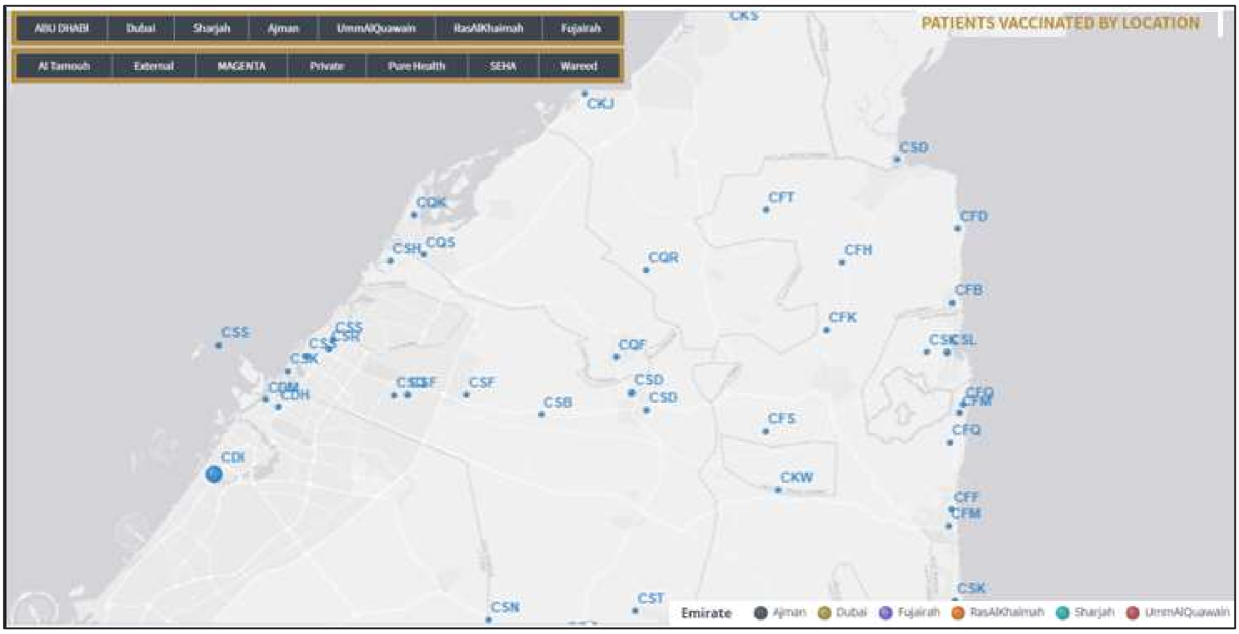


Figure 4: Covid-19 vaccination Administration Trend – By Location/UAE

During the pandemic, dashboard results were distributed to the leaders and healthcare professionals of the organization, which includes vaccination details by patient demographics (fig. 1), Covid-19 vaccination administration Trend by different factors (fig. 2-4), first dose/2nd dose and booster dose details (fig. 5). The results on model performance (table 1-4, Fig. 6-11) were also shared with healthcare providers including physicians and administrative staff of EHS facilities via AI driven dashboard.

Advanced AI models were compared based on deep-learning-based algorithms (decision tree-based, Random Forest, Neural Network, and Gradient Boosting models) for providing care excellence during the pandemic. A risk score model was developed, and investigators/clinical doctors worked together on the development of this machine model. The population with various characteristics has a different likelihood ratio for risk.

In phase I, a random sample of 23996 encounters accounting for 22865 people across UAE was considered for the development of the machine learning models to predict the COVID infection risk. 3818 encounters came out to COVID-19 positive from the given sample. Data was split into training (70%) and Validation (30%) datasets. Patients with fever were 2.8 times more likely to be infected than patients without fever. In the gender analysis the male patients are 2 times more likely to be infected with COVID- 19 when compared to female patients. Similarly, patients in the age group 46-60 were 1.5 times more likely to be infected compared to patients in other age groups.

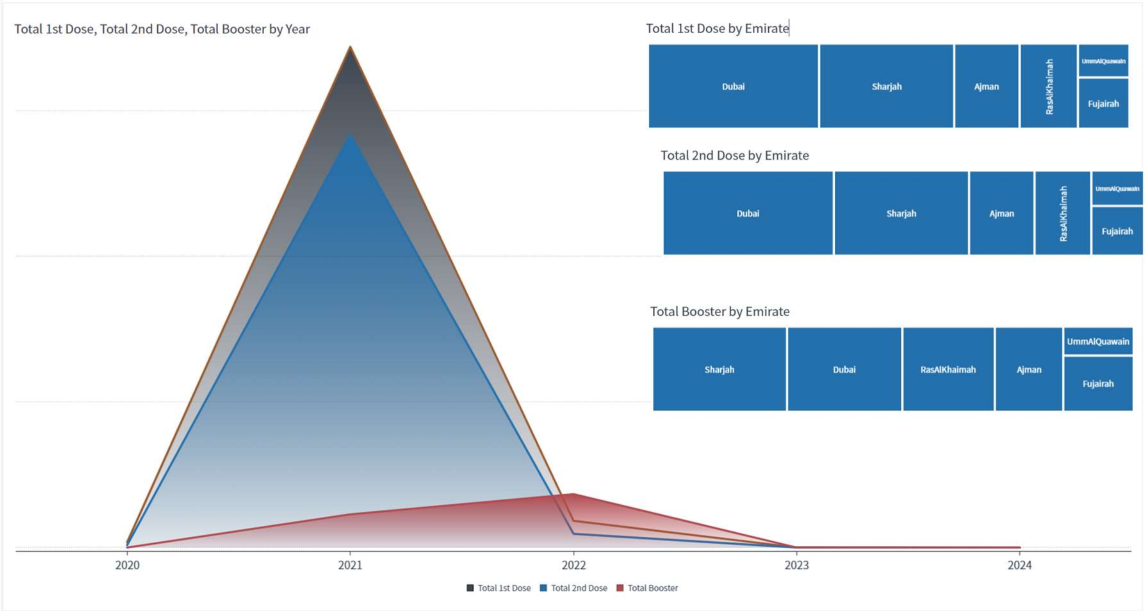


Figure 5: Covid-19 Booster Dose vaccination Trend – By Year and Emirates

Table 1: Details regarding Champion Model

The indicator predicted by the model	Best ML Model	AUC: Training	AUC: Validation
Covid Infection Risk among population	Gradient Boosting	75%	74%
Covid related ICU mortality	Random forest	94%	91%

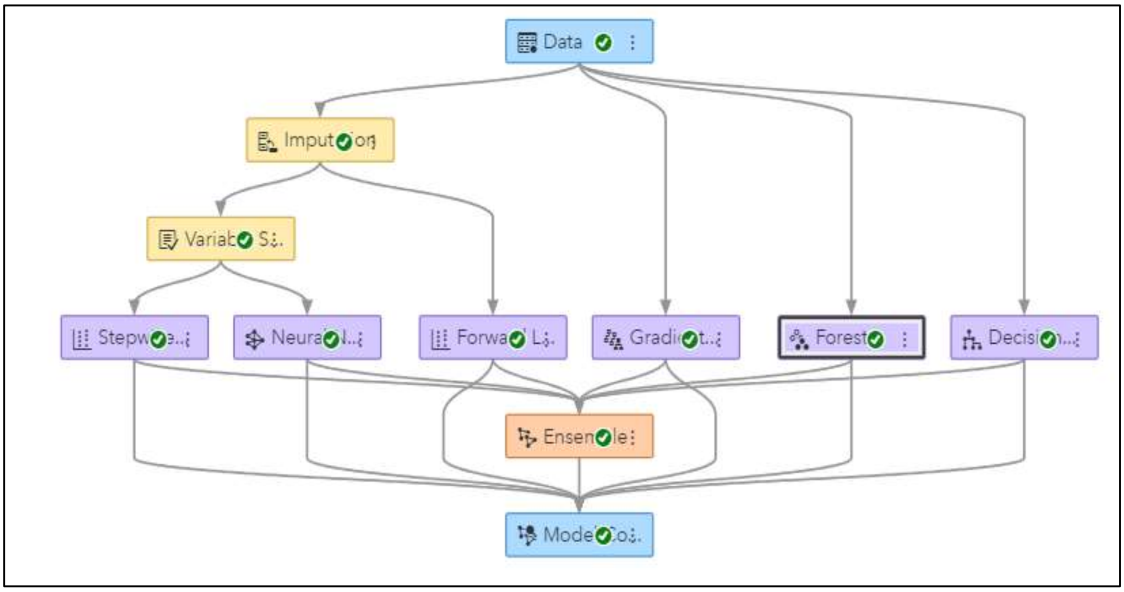


Figure 6: Pipeline

While predicting the COVID related ICU mortality (phase II), the prediction using the Forest Algorithm produced high level accuracy, 91.8% of the real-time data was correctly classified by the model. The training and validation accuracy was 94% and 91%. The model effectively identified the risk score, and concluded that higher Ferritin levels, Ventilator usage, MCHC as the top clinical variables, and Hypotension as the top comorbid condition for predicting COVID related mortality in ICU. The relative importance of these variables was determined by using a one-level decision tree. The model also identified some of the key factors including comorbidities and lab results that were very specific to the EHS population. This empowered the EHS providers to plan on the necessary resource mobilization, lab test planning, and care strategy for improved outcomes. results of phase I-II are presented in the below table (table 1).

In phase III, the goal was to develop an AI model that could effectively prioritize patients for vaccination or booster shots based on their risk of severe COVID-19 outcomes. The table 2 and pipeline (Fig. 6) presents a comparison of several machine learning models using various performance metrics to evaluate their effectiveness in this task.

Table 2: Comparison of Model Performance

Model Applied	Root Average Squared Error (RMSE)	(AUC)	Accuracy	F1 Score	Gini Coefficient	Lift	False Positive Rate	Misclassification Rate	Gain	KS (Youden)
Gradient Boosting	0.337	0.790	0.846	0.247	0.580	2.523	0.016	0.154	2.118	0.414
Forest	0.337	0.790	0.846	0.230	0.580	2.385	0.014	0.154	2.085	0.418
Ensemble	0.337	0.789	0.847	0.238	0.577	2.420	0.013	0.153	2.086	0.413
Neural Network	0.339	0.783	0.845	0.235	0.566	2.393	0.016	0.155	2.024	0.400
Forward Logistic Regression	0.339	0.783	0.845	0.238	0.565	2.377	0.016	0.155	2.024	0.400
Stepwise Logistic Regression	0.339	0.783	0.845	0.238	0.565	2.377	0.016	0.155	2.024	0.400
Decision Tree	0.341	0.769	0.845	0.235	0.537	2.416	0.016	0.155	2.026	0.383

AUC – **Area Under ROC**

The Forest model and Gradient Boosting model were identified as the champion models, key metrics of the Gradient Boosting model show Root Mean Squared Error (RMSE) as 0.337, an Area Under the ROC Curve (AUC) of 0.790, and an accuracy of 0.846. Its F1 score of 0.247 is the highest among the models, indicating a balanced precision and recall. It also has the highest lift (2.523) and gain (2.118), suggesting it is particularly effective at identifying individuals most at risk. The Gradient Boosting model achieves these results with a slightly higher false positive rate (0.016) compared to the Forest and Ensemble models.

The Ensemble model also perform well, with similar RMSE, AUC, and accuracy scores. The Ensemble model has the highest accuracy (0.847) and the lowest false positive rate (0.013), making it a strong contender for practical applications where minimizing false positives is crucial. However, its F1 score and lift are slightly lower than those of the Gradient Boosting model.

Neural Network and Logistic Regression models have slightly higher RMSE values (0.339), indicating marginally less precision in prediction. Their AUC values are lower (0.783) than those of the Gradient Boosting and Forest models, and they also have a higher false positive rate. Despite these differences, their lift and gain values are relatively competitive, indicating they can still identify high-risk individuals effectively.

The Decision Tree model has the lowest AUC (0.769) and Gini coefficient (0.537), reflecting a reduced ability to discriminate between high and low-risk individuals. However, it maintains a competitive lift and accuracy, making it a viable option for certain applications.

In summary, the Forest model was identified as the most effective model in identifying and prioritizing high-risk individuals for vaccination or booster shots, with a balance of high precision, recall, and a low misclassification rate.

Champion model results: The champion model for this project is Forest. The model was chosen based on the KS (Youden) for the Test dataset (0.42). 84.59% of the Test dataset was correctly classified using the Forest model. Out of 54 potential variables, the five most important factors identified for booster prioritization are Emirates of vaccination administration, patients' occupation, Age, Day to Second Dose and presence of Diabetes/CKD. The results of the champion model are detailed below.

The plot (Fig. 7) shows how the misclassification rate changes as the number of trees in the forest increases. The training error typically decreases as the number of trees increases, but the error for the validation dataset gives an indication of how well the model generalizes. For this model, the minimum error for the validation dataset is 0.1547 and occurs for 75 trees, so the validation error is still decreasing at the last tree.

The Receiver Operating Characteristic (ROC) curve (Fig. 8) was used in the study to assess the performance of selected classification models. It plotted the true positive rate (sensitivity) against the false positive rate (1-specificity) at various threshold settings.

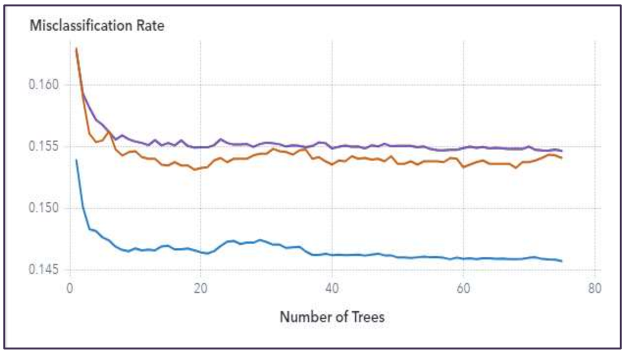


Figure 7: Misclassification rate in Train/Test/Validation datasets

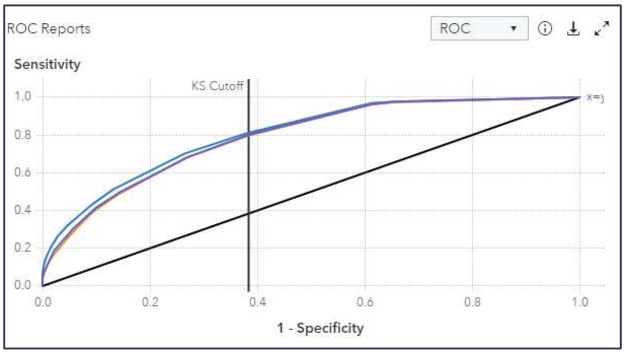


Figure 8: ROC curve and AUC of the champion model

Each point on the curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. and then AUC was computed to summarize the model's discriminatory power. The obtained AUC value of 0.8 and above indicates excellent discrimination power of the model. In this study, a higher AUC value demonstrates that our model effectively differentiates between individuals who are at high risk of infection required booster dose and those who are not, thereby supporting targeted vaccination and booster dose strategies. These tools helped to determine how well the AI model can distinguish between high-risk and low-risk individuals for COVID-19 infection and subsequent booster doses.

This ROC curve rapidly approaches the upper-left corner of the graph, where the difference between sensitivity and 1-specificity is the greatest, indicates the selected model is more accurate. This ROC curve plots sensitivity (the true positive rate) against 1-specificity (the false positive rate), which are both measures of classification based on the confusion matrix. These measures are calculated at various cutoff values.

Figure 9: Confusion Metrix

Table 3: Champion Model's performance at different thresholds

Threshold	True Positives (%)	True Negatives (%)	False Positives (%)	False Negatives (%)
0.15	79.75	61.64	38.36	20.25
0.20	70.36	73.41	26.59	29.64
0.50	13.12	98.73	1.27	86.88

To help identify the best cutoff to use when scoring data, the KS Cutoff reference line is drawn at the value of 1-specificity where the greatest difference between sensitivity and 1-specificity is observed for the validate partition. The KS Cutoff line is drawn at the cutoff value 0.15, where the 1-specificity value is 0.384 and the sensitivity value is 0.797. As the cutoff values range from 0 to 1, inclusive, in increments of 0.05, the predicted classification is the event when infection flag is greater than or equal to the cutoff value. otherwise, it was considered as a non-event.

The above event classification report is a visual representation of the confusion matrix at various cutoff values for each dataset (Fig. 9). The classification cutoffs used in the plot are the default (0.5) and the KS value for the TRAIN dataset (0.2), VALIDATE dataset (0.15), TEST dataset (0.15). For this data, for the bar corresponding to the event level of Infection "1 (yes)", the segment of the bar colored as blue (Correct) corresponds to true positives.

The above table (Table 3) illustrates the model's performance across different thresholds. At a threshold of 0.15, the model achieves high sensitivity with 79.75% true positives but also a high false positive rate of 38.36%, indicating it identifies many infections at the risk of over-predicting. At 0.20, the model strikes a better balance, with 70.36% true positives and 73.41% true negatives, reducing false positives to 26.59% but increasing false negatives to 29.64%. At a threshold of 0.50, the model becomes very conservative, identifying only 13.12% of true positives and nearly all non-infections correctly (98.73%), but with a high false negative rate of 86.88%. This shows a shift from high sensitivity and false positives at lower thresholds to high specificity and missed infections at higher thresholds. The choice of threshold thus involved it is a trade-off between sensitivity and specificity based on the desired balance of identifying true infections versus avoiding false positives. Here to balance between correctly identifying true cases and minimizing false positives, the 0.20 threshold is the most appropriate choice.

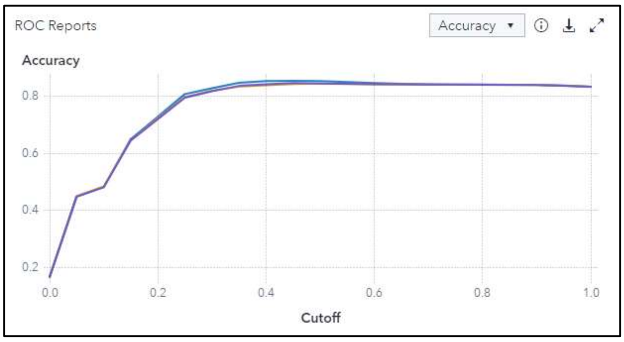


Figure 10: Accuracy Level of the champion model

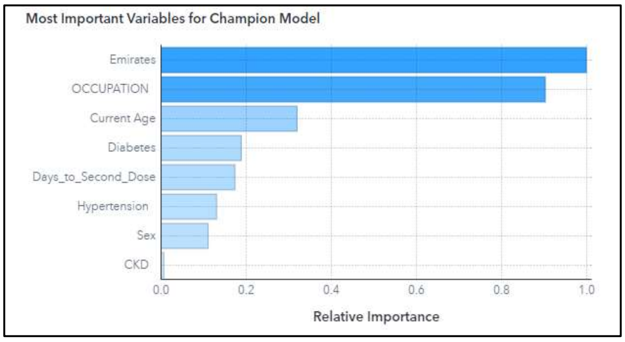


Figure 11: Relative Importance of the Factors

Table 4. Other Metrics

Threshold	Precision (%)	Recall (%)	F1 Score
0.15	29.1	79.8	42.4
0.20	34.5	70.4	46.0
0.50	67.2	13.1	21.8

For the selected model, the accuracy in the TEST partition (Fig. 10) at the cutoff of 0.5 is 0.8459, in the TRAIN partition at the cutoff of 0.5 is 0.8542 and the accuracy in the VALIDATE partition at the cutoff of 0.5 is 0.8453. Accuracy is the proportion of observations that are correctly classified as either an event (infection risk – yes) or non-event (infection risk – No) for booster intake, which was calculated at various cutoff values, as (true positives + true negatives) / (total observations).

This representation (Table 4) allows for a more nuanced understanding of the model's performance in terms of its ability to correctly identify positive cases, its coverage, and the balance between precision and recall.

Variable importance/Major predictors of booster vaccination: Variable importance (8 / 54 variables) is calculated using a surrogate model, presented in Fig. 11, a one-level decision tree for each input where the target is the predicted class or value. Inputs with a positive importance value are determined to be important. The most important inputs across the champion and challenger models for this project appear as follows.

PD plots for deeper insights

The partial dependence (PD) plots were reviewed (Fig. 12), and it provides deeper insights into the model's behavior, revealing the impact of specific features on the predicted risk of severe COVID-19 outcomes for effective vaccination. The "Emirate" feature shows that residents of Fujairah have the highest predicted risk (0.29), suggesting a priority for booster vaccinations in this region, while Dubai shows the lowest (0.03). Occupation-wise, government healthcare workers exhibit the highest predicted risk (0.42), potentially due to higher exposure, whereas laborers show the lowest (0.12). Age-wise, individuals around 21 years old face a slightly elevated risk (0.19), while comorbidities like diabetes (0.39), hypertension (0.36), and chronic kidney disease (0.27) significantly increase vulnerability. Additionally, females (0.19) are predicted to have a marginally higher risk compared to males (0.15). Lastly, a shorter gap between vaccine doses correlates with a slightly higher risk. These insights, facilitated by the AI augmented EHS intelligence platform, assist stakeholders in implementing targeted COVID-19 precautions and vaccination strategies for at-risk populations.

Overall, the statistics distribution was done through the dashboards created on the EHS intelligence platform, which supports EHS stakeholders in knowledge sharing and decision-making. This platform serves as a central hub for accessing, analyzing, and sharing current healthcare information. By highlighting high-risk groups, it helps healthcare providers implement effective COVID-19 precautions and vaccination strategies.

DISCUSSION

The current study focused on epidemic trends and the orchestration of safe vaccination campaigns. EHS's early adoption of AI-driven strategies exemplifies the potential of technology in navigating the complexities of public health crises. In the broader context, machine learning has been extensively employed to predict various COVID-19-related outcomes in EHS. The present project includes classifying patients with increased vulnerability who are more susceptible to severe complications and thus more likely to require hospitalization or experience severity/mortality[28-29], as documented in numerous studies. However, our research highlights a notable gap in the application of machine learning specifically for the prioritization of vaccination efforts. While existing models effectively predict high-risk cases, less attention has been paid to optimizing vaccine distribution strategies. The strategic prioritization of vaccinations plays a pivotal role in identifying which segments of the population should be immunized initially, taking into account aspects such as age demographics, job-related exposure risks, comorbidities, and general infection susceptibility. Our findings suggest that AI can significantly enhance decision-making processes in vaccination strategies by identifying those who would benefit most from early immunization.

Figure 12: PD plot by various Factors

The models developed in our study incorporated a range of factors, including demographic variables, health status, and exposure risks, to predict the likelihood of severe COVID-19 outcomes and, subsequently, to prioritize vaccination for the most vulnerable populations. Thus, EHS emerged as a pioneering entity in the region by integrating AI techniques into the management of the COVID-19 pandemic.

Nations across the globe have committed significant resources in an effort to control the spread of this deadly virus. In this context, AI has proven to be an essential tool, as many researchers utilize it to create mathematical models that analyze the pandemic's dynamics through shared data, ultimately supporting the well-being of society.[25] One of the studies enhances COVID-19 outcome prediction by using advanced machine learning classifiers on datasets of confirmed cases, recoveries, and fatalities.[26] The primary dataset for their study was sourced from Kaggle, a trusted platform for real-world datasets. Researchers gathered and pre-processed raw data related to COVID-19, including demographic details, admission and discharge dates, and patient outcomes such as recoveries and fatalities. This dataset, which combines both textual and numerical data, was used to develop and train COVID-19 prediction model, focusing only on the attributes relevant to analysis. Among the models they tested, the Gradient Boosting classifier achieved the highest accuracy with 90% for both confirmed cases and recoveries, and 92% for fatalities. This improvement supports better patient triage and reduces the burden on healthcare systems, offering valuable tools for managing current and future epidemics.

Another research focused on developing an intelligent web application for automatic COVID-19 detection using various machine learning techniques.[27] After preprocessing the dataset, including handling missing data and applying SMOTE for balancing, different classifiers and deep learning models were trained and evaluated by the researchers. The hybrid CNN-LSTM model, combined with the SMOTE approach, achieving 96.34% accuracy and a 0.98 F1 score using data from the Israeli Ministry of Health. This model was then deployed to a website, allowing users to receive instant COVID-19 predictions based on their symptoms.

Another study applied advanced AI techniques to enhance the accuracy of COVID-19 detection and prediction, demonstrating the growing role of artificial intelligence in managing and understanding the pandemic.[30] This research identified COVID-19 positive patients using machine learning and ANN models on anonymized full blood count data, without relying on symptoms or history. Data was taken from the Hospital Israelita Albert Einstein showed that Random Forest and ANN models achieved high accuracy, with AUC values of 94-95% for hospital patients and 80-86% for community cases. A simple linear combination of four blood counts achieved an AUC of 85% for community patients. This method detected characteristic immune response changes, offering a potential improvement for early COVID-19 screening where PCR tests are limited.

Effective diagnosis of emerging infectious diseases is critical for patient management and controlling outbreaks. This study aimed to validate a machine learning approach for predicting COVID-19 based on clinical features in emergency departments (ED).[31] Data from a US ED and an external ED were used to train and test models with gradient boosting, random forest, and extra trees classifiers. The random forest model achieved the highest performance with an AUC of 0.785, compared to 0.774 for gradient boosting and 0.720 for extra trees. The study confirms that ML can effectively predict COVID-19 and shows promise for future use in diagnosing emerging infectious diseases, particularly in settings with limited diagnostic tools.

Accurate prediction of mortality risk in hospitalized COVID-19 patients is critical for improving medical management and optimizing healthcare resources. One of the research projects investigates the application of explainable artificial intelligence methods to predict mortality risk based on admission data, aiming to enhance both the precision and interpretability of predictions.[32] Data from 824 patients admitted to Azienda Ospedaliera were analyzed using 19 machine learning models. The JRIP model, which provides clear and understandable explanations, performed best, achieving high accuracy in both cross-validation and a separate validation set from the third COVID-19 wave. The results demonstrate that explainable AI models can effectively assess patient risk and support optimized healthcare management, even across different phases and variants of the pandemic. As it was necessitating advanced tools for effective clinical decision-making and resource management, a study enhanced healthcare outcomes to predict ICU mortality in COVID-19 patients by applying machine learning classifiers to demographic, hematological, and biochemical data from ICU admissions during the second and third waves of the pandemic.[33] Using eight classifiers from the caret package in R, the Random Forest model achieved the highest performance with an AUC-ROC of 0.82, while k-nearest neighbors performed the worst (AUC-ROC: 0.59). The model demonstrated the highest sensitivity (0.7). Key predictors of mortality identified by the Random Forest model included serum urea, age, hemoglobin, C-reactive protein, platelets, and lymphocyte count. A review article provides a comprehensive summary of common medical history, clinical examinations, radiology, and laboratory data related to SARS-CoV-2.[34] It synthesizes literature from PubMed, Science Direct, and EMBASE on COVID-19 diagnosis and proposes algorithms for accurate and timely diagnosis of the disease. The review highlights the importance of developing precise and prompt diagnostic algorithms as the ideal approach for managing COVID-19.

AI can also revolutionize how we predict vaccine efficacy, tailor vaccination strategies to individual patient profiles, and identify potential adverse reactions before they become widespread issues. But the field of vaccination research where AI applications are still relatively scarce. One study identified the most effective COVID-19 vaccine for individuals while minimizing severe adverse reactions and assessing mortality risk, addressing the rare application of AI in vaccination studies.[35] By analyzing a dataset on vaccine adverse reactions, various classification models were evaluated, including Random Forest, Decision Tree, Light Gradient Boosting, and Extreme Gradient Boosting, with class balancing techniques like SMOTE and SMOTETOMEK to enhance performance. Key predictors included pre-existing conditions, age, and gender. The Random Forest model achieved the highest accuracy, ranging from 75% to 87%, with improved performance using class balancing methods. The study highlights the potential of AI in vaccination research and highlights the importance of considering medical history to optimize vaccine outcomes and safety. our study also illustrates how a healthcare organization in the UAE effectively implemented an AI-driven strategy to identify and prioritize vulnerable populations for COVID-19 vaccination administration and supplementary doses. One of the systematic reviews highlights the importance of refining mathematical models for vaccine prioritization, especially for age-specific strategies, to enhance future pandemic preparedness.[36] By analyzing various models, it highlights how different assumptions impact outcomes and informs better public health decisions for future infectious disease outbreaks.

With global vaccine supply being limited and the need for efficient distribution growing, developing machine learning models for vaccine prioritization is vital. Our paper advocates for progress in this field, demonstrating how such models can aid public health officials and policymakers in the effective allocation of scarce vaccine doses, ensuring that individuals at the highest risk of severe disease or transmission receive timely access to vaccination.

Project Impact

The project significantly impacted the healthcare sector by addressing the challenge of identifying the population groups at elevated susceptibility to COVID-19 infection for prompt immunization and booster administration. The AI model developed through the project integrated both traditional and advanced machine learning techniques, leading to the creation of a highly accurate model for stratifying vulnerable populations. This model not only improved vaccination strategies but also highlighted increased breakthrough infection rates months after the second dose, facilitating timely interventions. The selection of optimal AI model resulted in the most favorable patient outcomes during the pandemic in the country. The refinement of the model enabled the identification of patients with elevated vulnerability to COVID-19 complications and facilitated the evaluation of its associated ICU severity rates, which demonstrating the effectiveness of AI in improving pandemic response and healthcare management.

When comparing the current approach with previous methodologies, it can be seen that traditional approaches often relied on broad demographic categories or static risk factors to identify high-risk groups, which could lead to overgeneralization and missed at-risk individuals. These methods typically lacked real time data integration and feedback mechanisms, resulting in slower response times and less adaptive strategies. such earlier approaches do not fully integrated diverse data sources, leading to incomplete analyses and less effective risk stratification. AI-Driven Approach use real time data and integrates multiple variables (age, gender, nationality, vaccination history, infection rates, etc.) to dynamically and precisely identify high-risk individuals. This allows for more accurate and timely identification of those who need booster doses, ensures efficient use of resources and reduces healthcare costs.

Conventional models/applications often had limited visualization capabilities, making it difficult for healthcare staff to interpret data and make informed decisions. It potentially leads to wastage or shortages in critical areas. Whereas the AI-Driven dashboard on the EHS Intelligence platform offers a district-level view of vaccination data, enhancing transparency and provides clear actionable insights, helping healthcare staff to develop targeted further sustainable strategies. previous approaches have struggled to scale quickly or adapt to new data and changing circumstances. In comparison with this, the AI model is designed to be scalable and flexible, capable of adapting to new data inputs and evolving health scenarios. This makes it a robust tool for managing vaccination strategies in dynamic environments. Overall, the models resulted in significant policy enhancements and improved departmental strategies across EHS facilities. By quantifying resource demand and enhancing vaccination services through AI-based technology, it enabled precise risk stratification and identification of vulnerable groups. This solution updated health statistics and facilitated timely healthcare delivery. Consequently, COVID-19 instructional manuals and health policies were revised, introducing precautionary measures and preventive steps across all EHS facilities in the UAE. The project established EHS as a regional healthcare leader, utilizing advanced AI models for informed decision-making and practical clinical interventions.

Conclusion

A key challenge in the public health sector was accurately identifying vulnerable groups for COVID-19 infection and optimizing vaccine distribution among the varied populations in the UAE. Public health systems must efficiently identify patients in need of urgent care while also streamlining the vaccination process. To address this challenge, advanced machine learning algorithms were created, integrating a range of symptoms and patient profiles to evaluate the risk of COVID-19 infection. The project outcomes highlight the transformative impact of AI on public health strategies, particularly in managing the spread of infectious diseases like COVID-19. The study presents how a UAE healthcare organization successfully developed and deployed an AI-driven approach to identify and prioritize vulnerable populations for COVID-19 vaccination and Booster dose management. The study observed that the champions models significantly improve decision-making processes within healthcare systems, resulting in enhanced patient care, optimized resource allocation, and cost savings. The efforts on the integration of AI into priority setting for vaccination and booster shots exemplifies its potential to revolutionize public health initiatives and strengthen pandemic management. This innovative approach demonstrates AI's crucial role in advancing global health responses to future pandemics.

References

1. Haas EJ, McLaughlin JM, Khan F, et al. Infections, hospitalizations, and deaths averted via a nationwide vaccination campaign using the Pfizer-BioNTech BNT162b2 mRNA COVID-19 vaccine in Israel: a retrospective surveillance study. *Lancet Infect Dis.* 2021;**22**:357–366.
2. Goldberg Y, Mandel M, Bar-On YM, et al. Waning immunity after the BNT162b2 vaccine in Israel. *N Engl J Med.* 2021;

385:e85.

3. Tang P, Hasan MR, Chemaitelly H, et al. BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 delta variant in Qatar. *Nat Med.* 2021;27:2136–2143
4. Mahase E. COVID-19: booster vaccine gives "significant increased protection" in over 50s. *BMJ.* 2021;**375**
5. Korosec, C.S., Dick, D.W., Moyles, I.R. et al. SARS-CoV-2 booster vaccine dose significantly extends humoral immune response half-life beyond the primary series. Sci Rep 2024;14:8426. DOI: https://doi.org/10.1038/s41598-024-58811-3
6. Bilgin GM, Lokuge K, Munira SL, Glass K. Assessing the potential impact of COVID-19 booster doses and oral antivirals: A mathematical modelling study of selected middle-income countries in the Indo-Pacific. Vaccine X. 2023 Sep 9;15: 100386.
7. Irving SA, Sundaram ME. Prioritisation of COVID-19 boosters in the omicron era. Lancet. 2022 Oct 15;400(10360):1282-1283.
8. Chapman, L. A. C. et al. Risk factor targeting for vaccine prioritization during the COVID-19 pandemic. *Sci. Rep.* **12**(1), 3055 (2022).
9. Matrajt, L., Eaton, J., Leung, T. & Brown, E. R. Vaccine optimization for COVID-19: Who to vaccinate first? *Sci. Adv.* **7**(6), eabf1374 (2021).
10. Bubar, K. M. et al. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* **371**(6532), 916–921 (2021).
11. Li Y, Lee K-C, Bressington D, Liao Q, He M, Law K-K, Leung AYM, Molassiotis A, Li M. A Theory and Evidence-Based Artificial Intelligence-Driven Motivational Digital Assistant to Decrease Vaccine Hesitancy: Intervention Development and Validation. Vaccines. 2024; 12(7):708.
12. Larson HJ, Lin L. Generative artificial intelligence can have a role in combating vaccine hesitancy. BMJ. 2024 Jan 16; 384:q69.
13. Ma X, Wang Y, Gao T, He Q, He Y, Yue R, et al. Challenges and strategies to research ethics in conducting COVID-19 research. *J Evid Based Med.* (2020) 13:173–177.
14. Naudé W. Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. *AI Soc.* (2020) 35:761–65.
15. Yassine HM, Shah Z. How could artificial intelligence aid in the fight against coronavirus? *Expert Rev Anti Infect Ther.* (2020) 18:493–7.
16. Sharma A, Virmani T, Pathak V, Sharma A, Pathak K, Kumar G, Pathak D. Artificial Intelligence-Based Data-Driven Strategy to Accelerate Research, Development, and Clinical Trials of COVID Vaccine. Biomed Res Int. 2022 Jul 6;2022:7205241.
17. Yagin FH, Cicek İB, Alkhateeb A, Yagin B, Colak C, Azzeh M, Akbulut S. Explainable artificial intelligence model for identifying COVID-19 gene biomarkers. Comput Biol Med. 2023;154:106619.
18. Al-Qaness MAA, Saba AI, Elsheikh AH, Elaziz MA, Ibrahim RA, Lu S, et al. Efficient artificial intelligence forecasting models for COVID-19 outbreak in Russia and Brazil. *Process Saf Environ Prot.* (2021) 149:399–409.
19. Fang C, Bai S, Chen Q, Zhou Y, Xia L, Qin L, et al. Deep learning for predicting COVID-19 malignant progression. *Med Image Anal.* (2021) 72:102096.
20. Srinivasa Rao ASR, Vazquez JA. Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone-based survey when cities and towns are under quarantine. Infect Control Hosp Epidemiol. 2020 Jul;41(7): 826-830.
21. Du Y, Tu L, Zhu P, Mu M, Wang R, Yang P, et al. Clinical Features of 85 Fatal Cases of COVID-19 from Wuhan. A Retrospective Observational Study. Am J Respir Crit Care Med. 2020 Jun 1;201(11):1372-1379.
22. Daoust JF. Elderly people and responses to COVID-19 in 27 countries. PLoS One. 2020;15(7):e0235590.
23. Li J, Chen Z, Nie Y, Ma Y, Guo Q, Dai X. Identification of symptoms prognostic of COVID-19 severity: multivariate data analysis of a case series in Henan province. J Med Internet Res. 2020 Jun 30;22(6):e19636.
24. Aktar S, Talukder A, Ahamad MM, Kamal AHM, Khan JR, Protikuzzaman M, Hossain N, Azad AKM, Quinn JMW, Summers MA, Liaw T, Eapen V, Moni MA. Machine Learning Approaches to Identify Patient Comorbidities and Symptoms That Increased Risk of Mortality in COVID-19. Diagnostics (Basel). 2021 Jul 31;11(8):1383.
25. Lasya KL, Lahari DR. Akarsha AL, Prakash KB, Tran D. Analysis and Prediction of COVID-19 datasets using Machine Learning Algorithms. *ICEEICT 2022*; 01-03. doi: 10.1109/ICEEICT 53079.2022.9768598.
26. Raja, M. A., Ullah, I., Babar, M., & Aziz, T. (2023). Prediction of COVID-19 using machine learning techniques. The Asian Bulletin of Big Data Management, 3(1), 221–234.
27. Sanzida S, Azmiara A, Chand SM, Muktadir M, Riasat K. Automatic COVID-19 prediction using explainable machine learning techniques. Intl J Cognitive Computing in Engineering, 2023:4;36-46. Doi: 10.1016/j.ijcce.2023.01.003.
28. Sara AlShaya, Maryam Sayed Jaffar, Sheik Abdullah Jamal Mohideen, Aji Gopakumar, Vibhor Mathur, Maimoona Saeed. A Machine Learning Model to Predict COVID-19 Infection Risk in the United Arab Emirates. J Commun Dis. 2023;55(4):71-79.
29. Sara Alshaya, Maryam Sayed Jaffar, Aji Gopakumar, Sheik Abdullah Jamal Mohideen, Vibhor Mathur, Sudheer Kurakula, Badshah Mukherjee. Artificial Intelligence (AI) Model to predict the risk of COVID-19 ICU Severity: A Pandemic Success Story. J Commun Dis. 2024;56(2):6-14.
30. Abhirup B, Surajit R, Bart V, Joanne K, Michail M, Simonne W, Mark B, Louise SM. Use of Machine Learning and Artificial Intelligence to predict SARS-CoV-2 infection from Full Blood Counts in a population, International Immunopharmacology 2020;86:106705. Doi: 10.1016/j.intimp.2020.106705.
31. Tay J, Yen YH, Rivera K, Chou EH, Wang CH, Chou FY, et al. Development and External Validation of Clinical Features-based Machine Learning Models for Predicting COVID-19 in the Emergency Department. West J Emerg Med. 2024 Jan;25(1): 67-78. doi: 10.5811/westjem.60243.
32. Bottrighi, A., Pennisi, M., Roveta, A. *et al.* A machine learning approach for predicting high risk hospitalized patients with COVID-19 SARS-Cov-2. *BMC Med Inform Decis Mak* **22**, 340 (2022). https://doi.org/10.1186/s12911-022-02076-1
33. Sakagianni A, Koufopoulou C, Verykios V, Loupelis E, Kalles D, Feretzakis G. Prediction of COVID-19 Mortality in the Intensive Care Unit Using Machine Learning. Stud Health Technol Inform. 2023 May 18;302:536-540. doi: 10.3233/SHTI230200. PMID: 37203743.
34. Mouliou DS, Pantazopoulos I, Gourgoulianis, KI. COVID-19 Smart Diagnosis in the Emergency Department: all-in in Practice. *Expert Review of Respiratory Medicine* 2022;*16*(3):263–272. https://doi.org/10.1080/17476348.2022.2049760
35. Monadhel H, Abbas AR, Mohammed AJ. COVID-19 Vaccine: Predicting Vaccine Types and Assessing Mortality Risk Through Ensemble Learning Agorithms. F1000Res. 2023 Sep 25;12:1200. doi: 10.12688/f1000research.140395.1.
36. Gonzalez-Parra G, Mahmud MS, Kadelka C. Learning from the COVID-19 pandemic: a systematic review of mathematical vaccine prioritization models. medRxiv [Preprint]. 2024 Mar 7:2024.03.04.24303726. doi: 10.1101/2024.03.04.24303726. Update in: Infect Dis Model. 2024 May 15;9(4):1057-1080. doi: 10.1016/j.idm.2024.05.005.