

RESEARCH ARTICLE

Predictive modelling of transport decisions and resources optimisation in pre-hospital setting using machine learning techniques

Hassan Farhat^{1,2,3*}, Ahmed Makhoul^{1,4}, Padarath Gangaram⁵, Kawther El Aifa¹, Ian Howland¹, Fatma Babay Ep Rekik¹, Cyrine Abid⁶, Mohamed Chaker Khenissi¹, Nicholas Castle¹, Loua Al-Shaikh¹, Moncef Khadhraoui⁷, Imed Gargouri⁸, James Laughton¹, Guillaume Alinier^{1,9,10,11}

1 Ambulance Service, Hamad Medical Corporation, Doha, Qatar, **2** Faculty of Sciences, University of Sfax, Sfax, Tunisia, **3** Faculty of Medicine "Ibn El Jazzar", University of Sousse, Sousse, Tunisia, **4** College of Engineering, Qatar University, Doha, Qatar, **5** Faculty of Health Sciences, Durban University of Technology, Durban, South Africa, **6** Laboratory of Screening Cellular and Molecular Process, Centre of Biotechnology of Sfax, University of Sfax, Sfax, Tunisia, **7** Higher Institute of Biotechnology, University of Sfax, Sfax, Tunisia, **8** Faculty of Medicine, University of Sfax, Sfax, Tunisia, **9** University of Hertfordshire, Hatfield, United Kingdom, **10** Weill Cornell Medicine-Qatar, Doha, Qatar, **11** Northumbria University, Newcastle upon Tyne, United Kingdom

* Hfarhat1@hamad.qa



OPEN ACCESS

Citation: Farhat H, Makhoul A, Gangaram P, El Aifa K, Howland I, Babay Ep Rekik F, et al. (2024) Predictive modelling of transport decisions and resources optimisation in pre-hospital setting using machine learning techniques. PLoS ONE 19(5): e0301472. <https://doi.org/10.1371/journal.pone.0301472>

Editor: A. L. MAhfoodh, UNITEN: Universiti Tenaga Nasional, MALAYSIA

Received: October 12, 2023

Accepted: March 11, 2024

Published: May 3, 2024

Copyright: © 2024 Farhat et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data contain potentially identifying or sensitive patient information and owned by a third-party organization (Hamad Medical Corporation) and who has imposed them the Hamad Medical Corporation Research Ethics Committee. All data requests can be submitted to the Hamad Medical Corporation Medical Research Center on MRChelpdesk@hamad.qa.

Abstract

Background

The global evolution of pre-hospital care systems faces dynamic challenges, particularly in multinational settings. Machine learning (ML) techniques enable the exploration of deeply embedded data patterns for improved patient care and resource optimisation. This study's objective was to accurately predict cases that necessitated transportation versus those that did not, using ML techniques, thereby facilitating efficient resource allocation.

Methods

ML algorithms were utilised to predict patient transport decisions in a Middle Eastern national pre-hospital emergency medical care provider. A comprehensive dataset comprising 93,712 emergency calls from the 999-call centre was analysed using R programming language. Demographic and clinical variables were incorporated to enhance predictive accuracy. Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Adaptive Boosting (AdaBoost) algorithms were trained and validated.

Results

All the trained algorithm models, particularly XGBoost (Accuracy = 83.1%), correctly predicted patients' transportation decisions. Further, they indicated statistically significant patterns that could be leveraged for targeted resource deployment. Moreover, the specificity rates were high; 97.96% in RF and 95.39% in XGBoost, minimising the incidence of incorrectly identified "Transported" cases (False Positive).

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Conclusion

The study identified the transformative potential of ML algorithms in enhancing the quality of pre-hospital care in Qatar. The high predictive accuracy of the employed models suggested actionable avenues for day and time-specific resource planning and patient triaging, thereby having potential to contribute to pre-hospital quality, safety, and value improvement. These findings pave the way for more nuanced, data-driven quality improvement interventions with significant implications for future operational strategies.

Introduction

Ambulance services, often called “Emergency Medical Services” (EMS), are undergoing significant transformations globally, reflecting contemporary healthcare systems’ complex, dynamic challenges. In a world where rapid technological advancements often outpace the ability of traditional healthcare infrastructure to adapt, the pressing imperatives include effective and efficient medical intervention strategies. EMS face multifactorial challenges, including increasing patient volumes, finite resources, and the escalating need for speed and precision in life-saving interventions [1]. These challenges are particularly pronounced in multinational societies like Qatar, where a diverse demographic landscape adds layers of complexity to emergency medical care.

However, the utility of advanced techniques, such as machine learning (ML), assumes immense significance. The application of ML, a subfield of artificial intelligence, serves as a catalyst for deep data exploration—unearthing patterns and correlations that often remain invisible in traditional analyses [2]. In an environment where every minute counts, the capacity to predict medical needs and allocate resources judiciously is a dynamic process. ML’s predictive analytics provide an innovative approach to bridging existing gaps in the EMS domain, offering potentially transformative solutions to both old and new challenges [3].

In the literature, the integration of ML in EMS marked a significant shift towards more efficient and responsive healthcare systems. The role of predictive analytics in healthcare, and specifically in EMS. For instance, research demonstrated that ML helped enhance diagnostic processes. Their study showed that ML algorithms could significantly improve the accuracy and speed of medical diagnoses, particularly in complex cases like cancer detection. This advancement in diagnostics was crucial for early intervention and effective treatment planning. Further, other researchers highlighted how ML is used in patient care management by personalising patient care plans, leading to better patient outcomes and more efficient use of healthcare resources. This aspect of ML aligns with the overarching goal of improving healthcare delivery and patient satisfaction. It also helped predict and manage public health crises, such as pandemics. Their research illustrates how ML models can forecast disease spread patterns, aiding in the formulation of effective public health responses. It helped forecast medical emergencies and anticipate specific trends and emergencies, enabling proactive resource allocation. Recent studies have highlighted the potential of ML in transforming EMS operations by enhancing response times and resource allocation in EMS and reducing. Similarly, other studies indicated that ML algorithms could predict high-demand areas, allowing for better ambulance positioning and quicker response time, as well as resource optimisation. Other studies revealed that ML could lead to cost savings in healthcare by reducing unnecessary interventions aligning with the goal of value improvement, indicating that ML can contribute to cost reduction and quality enhancement in healthcare services.

Qatar's multinational population makes it a uniquely challenging environment for the emergency care service [4]. Hamad Medical Corporation Ambulance Service (HMCAS) is the leading provider in this setting and continually evolves its healthcare infrastructure. This makes HMCAS a particularly relevant case study for integrating ML techniques into its operational system. The rich data generated in such a complex environment are fertile grounds for ML algorithms to tease out nuanced insights that can inform more effective resource allocation strategies. With healthcare costs soaring worldwide, there is a growing need for value improvement (VI), a quality improvement (QI) concept that aims not only at cost reduction but also at enhancing the quality of healthcare services [5]. ML fits neatly into this VI culture. ML could contribute to resource optimisation by facilitating advanced predictive analytics of patient transport decisions, ultimately leading to fiscal savings and improved healthcare outcomes. To our knowledge, no previous study approached the patients' transport decisions in a pre-hospital setting using ML.

The study aimed to investigate the role of ML algorithms in predicting patients' transport decisions, which would allow optimal resource allocation in HMCAS in Qatar.

Method

Setting and source of data

A retrospective quantitative analysis with predictive modelling was conducted on 93,712 pre-hospital emergency calls received by HMCAS between January 1st and May 31st, 2023. The electronic patient care record (ePCR) system, overseen by the HMCAS Business Intelligence (BI) division, served as the data source. Compliance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines was maintained throughout the study for machine learning-based analyses [6]. Ethical approval was obtained from the Hamad Medical Corporation Medical Research Centre (Reference: MRC-01-22-264). R-Studio was employed for data processing and analysis.

Participants

The criteria for inclusion encompassed all 999 emergency calls leading to the dispatch of an ambulance, during which a paramedic conducted an on-scene patient evaluation, irrespective of whether the patient was transported to a hospital or not. Exclusion criteria were the cases involving a deceased individual and when the 999 call was made by a healthcare facility.

Statistical analysis

Predictors and outcome variable. The variable "Handover" was considered the outcome variable. It included a binary outcome, which is whether a patient, after receiving emergency treatment by paramedics on scene, was conveyed to a healthcare facility ("Transported" = 1) or not ("Not Transported" = 0).

The continuous predictors were "Age", "Weight", and the hour when the 999 call was received ("Hour_received").

The categorical variables were nationalities categories ("Nationality_CAT"), the entity managing the 999 calls or "CFS_Owner" (A "Call For Service" can be attributed to the Police, Civil Defense, or HMCAS), Program Question AnswerTM (ProQA) "ProtocolName", "Unit_Type", "DispatchType", which is determined according to the ProQA protocol outcome, "Provisional-Diagnoses_CAT", and co-morbidities such as "Seizure", "Asthma", Diabetes Mellitus "DM", Chronic obstructive pulmonary disease ("COPD"), "CurrentlyPregnant", Cerebral Vascular Accident ("CVA"), "Hypertension", Cardio Artery Disease ("CAD"), "Other", and Unknown.

Data pre-processing. Effective data analysis in a medical context hinges upon the quality and coherence of the underlying data [7]. Pre-processing steps were implemented to ensure the dataset's reliability, interpretability, and clinical relevance by renaming, recoding, transforming variables, and handling missing values.

First, the variables' renaming was conducted to make them understandable in a manageable format. This step is foundational, ensuring that subsequent operations are intuitive and less error-prone. This was achieved by mapping column names to more descriptive terms. This included the variables "ProtocolName", "ProtocolCode", "DispatchType", "ProvisionalDiagnosis", and "Seizure". Second, existing values under the variables were recoded to a more understandable form, such as the values under "ProtocolName" that were recoded into [the protocol name and its ProQA™ code between "()"] as defined by the International Academy of Emergency Medical Dispatch [8]. Third, data transformation was performed. For example, values under the variables "ProtocolName" and "DispatchType" were altered based on the corresponding value in the "CFS_Owner" field. Data was restructured to represent the calendar week numbers.

Further, in the transformation, recategorisation of several variables such as "TransportedTo", the hospital facility "PatientTriageArea", "ProvisionalDiagnosis", and "Nationality_CAT" was performed. Fourth, the variables not required after transformation were removed. Fifth, the missing values were assessed by generating a plot to allow the visualisation of the 'missingness' patterns across variables [9]. This was a twofold process: initially, to assess the extent and nature of missing data and second, to verify the effectiveness of the missing values management technique. Missing data were managed using the Multiple Imputation by Chained Equations (MICE) algorithm. MICE is a statistical technique that runs multiple simulations to fill in the missing data, offering a more accurate representation of the data landscape [9]. It uses all available variables to estimate the missing ones, proceeding variable by variable iteratively repeated until the dataset is complete. This process culminates in several datasets that are analysed collectively, enhancing the reliability of the research outcomes. "Location_LAT" and "Location_Long" were excluded from the MICE process as they were only utilised to plot the patient transport/not transported decision map. For continuous variables, the MICE imputation method deployed was the Predictive Mean Matching (PMM), whereas the Classification and Regression Trees (CART) MICE method was utilised for the categorical variables [10]. The imputed values were integrated into the original dataset, fulfilling the replacement of missing values for comprehensive data.

Outliers were assessed by designing boxplots. They evinced no discernible outliers, thus obviating the need for further outlier removal steps. Hence, the initial data was deemed devoid of outliers and was retained for subsequent analysis. Feature selection was conducted by employing two automated methods to ensure the robustness of the chosen variables. The first method was the Random Forest (RF) algorithm. The second method was the Recursive Feature Elimination (RFE) using the "Wrapped Method" [11]. RFE is used to evaluate a combination of features and assign a score based on model accuracy considering the interactions between features [11]. The Wrapped Method helps improve the RFE's performance by automatically selecting the most important features, thereby making the model more accurate and efficient [12].

Further, categories with low counts, such as the non-EMS ProQA™ call-taking protocols, such as "Maritime", "Fire", and "Rescue", under the "ProtocolName" variable, were grouped together to avoid algorithmic errors associated with sparse data during ML analysis. Categorical variables were converted into numerical codes to avoid causing errors when building the models.

It is important to note that despite preliminary feature selection, a substantial correlation existed among variables, as confirmed by Chi-square and Wilcoxon rank-sum tests (S1 File).

Variables selected for inclusion were not determined solely by statistical tests; they were also influenced by the researchers' expertise and what was previously employed in the literature [13]. Other epidemiological analyses that identified significant associations with co-morbidities further informed the selection, including the table in [S1 File](#). Therefore, in addition to "Handover," which served as the outcome variable, the final list of variables retained for supervised predictive modelling included "Unit_Type," "Region," "ProtocolName," "WeekNumber," "Hour_Received," "LocationType," "WeekDay," "Age," "Gender," "Nationality_CAT," "ProvisonalDiagnoses_CAT," "Hypertension," and "DM."

Predictive modelling

Supervised predictive modelling (PM) was adopted for this study.

The pre-processed dataset was segregated into training (80%; $n = 70,783.20$) and test subsets (20%; $n = 17,695.80$). This partitioning is critical for unbiased model evaluation and preventing overfitting [14].

Parallel computing (PC) was utilised to speed up the data analysis process. A PC allows multiple calculations simultaneously by taking advantage of the multiple cores in a processor [15]. This approach benefits complex computations, helping us obtain results more quickly and efficiently.

Hyperparameter tuning was used to optimise the performance of the ML algorithms [16]. This enables the creation of the most effective model for predicting the transport decisions variable with the highest accuracy, increasing the reliability of the results.

Further, four algorithms were employed to train the data: 1) RF utilised a ten-fold cross-validation [14]. This approach randomly divided the training dataset into ten equal-sized subsets. The RF algorithm was trained on nine subsets and tested on the remaining one. This procedure was repeated ten times using a different subset for validation. 2) Support Vector Machine (SVM) was enhanced with a Radial Basis Function (RBF) Kernel. RBF Kernel optimises model performance by measuring the similarity between data points and examining their relative positions in a defined set of characteristics [17]. These characteristics are the pre-determined variables identified as relevant to the model's predictive accuracy [17]. 3) The eXtreme Gradient Boosting (XGBoost) was adopted using its built cross-validation system, which provides a robust and highly accurate algorithmic approach [18]. 4) Adaptive Boosting (AdaBoost) was also utilised. AdaBoost allows boosting the performance of the models by iteratively correcting their errors, offering a straightforward yet effective ensemble method [18].

Each model was evaluated on several metrics such as Accuracy, Sensitivity, Specificity, Matthews Correlation Coefficient (MCC), and Area Under the Receiver Operating Curve (AUC). Prediction's feature importance was assessed. Feature importance allows for identifying the most influential variables in the predictive models. The final predictive models were used to make predictions on the test data. Control charts, specifically p-charts and c-charts, were utilised for statistical quality control to monitor the performance of each model over time [19, 20]. Using the statistical process control (SPC) charts, the prediction models were summarised and visualised by day hour ("Hour_received") and weekday to allow a better understanding of the predictive models' performance at different times of the day and on different weekdays. This is crucial to assess the usefulness of these models in pre-hospital settings.

Results

From June to May 31st, 2023, 93,712 emergency calls were received, resulting in 72,279 patients being transported, 21,301 not being transported, and 132 cases identified as cases of death upon arrival on scene. After data pre-processing and removing the cases that did not

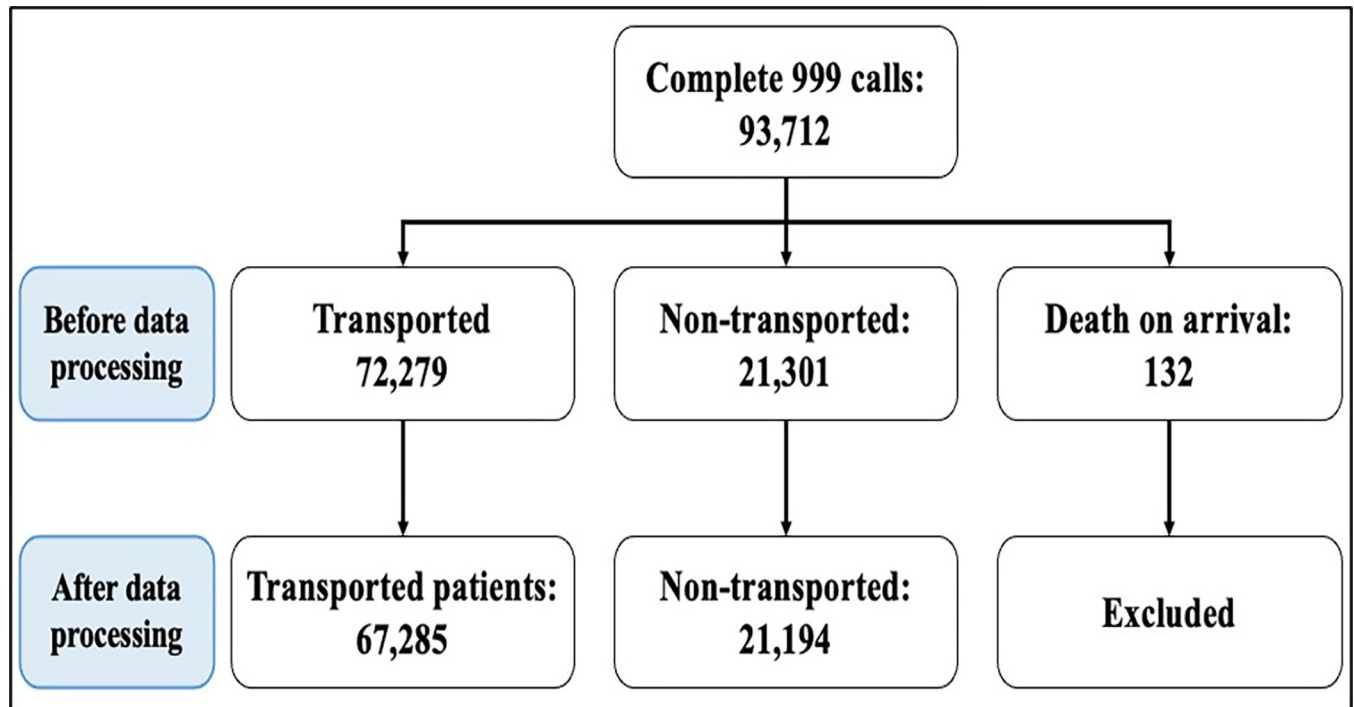


Fig 1. Geographical distribution of transported and not transported patients across Qatar. (This map was created by the first author using the “Leaflet” package in R. The data is available under the Open Database License).

<https://doi.org/10.1371/journal.pone.0301472.g001>

comply with the inclusion criteria, 67,285 cases of transported patients and 21,194 related to patients who were not transported were retained for the predictive modelling (Fig 1).

“S2 File” presents the missing value percentages before and after the imputation.

“S3 File” presented boxplots generated to visualise the outliers within the continuous variables. No discernible outliers were visualised.

Automated feature selection methods were used. The Mean Decrease Accuracy (MDA) and Mean Decrease Gini (MDG) determined by RF and the scree-plot coefficients determined by RFE were plotted respectively in “S3 File”. For the RF method, the MDC indicates how much model accuracy decreases when a feature is randomly permuted. Higher values imply that the feature is important. The MDG measured how each feature contributed to the homogeneity of the nodes and leaves in the resulting Random Forest. The RFE method evaluated the effectiveness of different feature sets in predicting an outcome. Lower Root Mean Square Error and Mean Absolute Error values, along with R^2 values closer to 1, signal better feature sets for accurate predictions; the smaller the associated standard deviations, the more consistently reliable the selected feature is. Although the RFE does not directly provide the variable names in its table and attributes them numbers instead, it provides the most significant variable at the end. The tables in “S4 File” indicated that features such as provisional diagnoses (“ProvisionalDiagnoses_CAT”), priority to hospital (“PriorityToHospital”), ProQA™ call-taking protocol name (“ProtocolName”), and patient triage area (“PatientTriageArea”) were predominant for accurate predictive modelling. Hence, the variables included in the modelling were selected based on the researchers’ clinical experience and operational relevance.

For the predictive modelling, “Table 1” evaluated the four ML algorithms utilised in this study to predict whether patients who call 999 will be transported to a hospital or not. Each of the models’ performance metrics (Accuracy, Sensitivity, Specificity, Kappa, and McNemar’s

Table 1. Predictive models performance metrics.

Random_Forest	XGBoost	AdaBoost	Support_Vector_Machine
Accuracy: 82.90%	Accuracy: 83.10%	Accuracy: 78.25%	Accuracy: 76.65%
95% CI: (82.3%, 83.43%)	95% CI: (82.54%, 83.65%)	95% CI: (77.64%, 78.86%)	95% CI: (76.02%, 77.27%)
No Information Rate (NIR): 76.05%	No Information Rate (NIR): 76.05%	No Information Rate (NIR): 76.05%	No Information Rate (NIR): 76.05%
p-value [Acc* > NIR]: < 0.01	p-value [Acc* > NIR]: < 2.2×10 ⁻¹⁶	p-value [Acc* > NIR]: 2.033×10 ⁻¹²	p-value [Acc* > NIR]: 0.032
Sensitivity: 35.87%	Sensitivity: 44.10%	Sensitivity: 17.39%	Sensitivity: 72.20%
Specificity: 97.69%	Specificity: 95.39%	Specificity: 97.42%	Specificity: 98.55%
Kappa: 0.42	Kappa: 0.46	Kappa: 0.20	Kappa: 0.08
McNemar's Test p-value: < 0.01	McNemar's Test p-value: < 0.01	McNemar's Test p-value: < 0.01	McNemar's Test p-value: < 0.01
Confusion Matrix and Statistics	Confusion Matrix and Statistics	Confusion Matrix and Statistics	Confusion Matrix and Statistics
Prediction 0 1	Prediction 0 1	Prediction 0 1	Prediction 0 1
0 1520 311	0 1869 621	0 737 347	0 306 200
1 2718 13146	1 2369 12836	1 3501 13110	1 3932 13257
Pos Pred Value: 83.01%	Pos Pred Value: 75.06%	Pos Pred Value: 67.99%	Pos Pred Value: 60.48%
Neg Pred Value: 82.87%	Neg Pred Value: 84.42%	Neg Pred Value: 78.93%	Neg Pred Value: 77.13%
Prevalence: 23.95%	Prevalence: 23.95%	Prevalence: 23.95%	Prevalence: 23.95%
Detection Rate: 8.59%	Detection Rate: 10.56%	Detection Rate: 4.17%	Detection Rate: 1.73%
Detection Prevalence: 10.35%	Detection Prevalence: 14.07%	Detection Prevalence: 6.13%	Detection Prevalence: 2.86%
Balanced Accuracy: 66.78%	Balanced Accuracy: 69.74%	Balanced Accuracy: 57.41%	Balanced Accuracy: 52.87%
'Positive' Class: 0	'Positive' Class: 0	"Positive" Class: 0	'Positive' Class: 0
MCC: 47.02%	MCC: 48.47%	MCC: 26.37%	MCC: 14.69%
AUC: 85.05%	AUC: 84.87%	AUC: NULL	AUC: NULL

<https://doi.org/10.1371/journal.pone.0301472.t001>

Test p-value) were determined. Accuracy denotes the proportion of true positives and true negatives among the total observations. The “No Information Rate (NIR) and p-value” indicate the baseline model’s accuracy, which in this case is 76.05%. The p-value suggests that the models are statistically significant in improving prediction compared to a no-information model. Sensitivity and Specificity represent the model’s ability to identify “true positives” and “true negatives” correctly. High Specificity across the models (97.69% in Random Forest to 95.39% in XGBoost) suggests fewer false positives. Kappa offers a chance-corrected measure of agreement between observed and predicted classifications. Higher Kappa suggests better performance. XGBoost had the best Kapp value ($k = 0.46$) and then RF ($k = 0.42$). McNemar’s Test helps us understand whether the ML model was genuinely helpful in making reliable predictions or if it was by chance. The p-value was <0.01 for all models, suggesting they are significantly better at making correct predictions. The MCC represents the quality of binary classifications. The closest to 100%, the better the prediction. XGBoost had the best MCC (48.47%) followed by RF (47.02%). For the AUC (85.05% for XGBoost and 84.87% for RF), the higher the value, the better the model’s performance. Positive Predictive Value, Negative Predictive Value, and Balanced Accuracy are ways to better understand how good a model is at making predictions. True Negatives refer to those cases when patients were “Not Transported” in reality and correctly predicted by the algorithm. False Positives refer to cases when patients were “Not Transported” but were incorrectly predicted as “Transported” by the algorithm. Lastly, False Negatives refer to “Transported” cases but were incorrectly predicted as “Not Transported”.

“Fig 2” shows a comparative evaluation of four ML algorithms was performed. RF demonstrated a robust sensitivity in identifying “Transported” cases with a high True Positive count (13,146). Still, it struggled to correctly predict the “Not Transported” cases since relatively low True Negative cases ($n = 1,520$) and high False Positive cases ($n = 2,718$) were predicted.

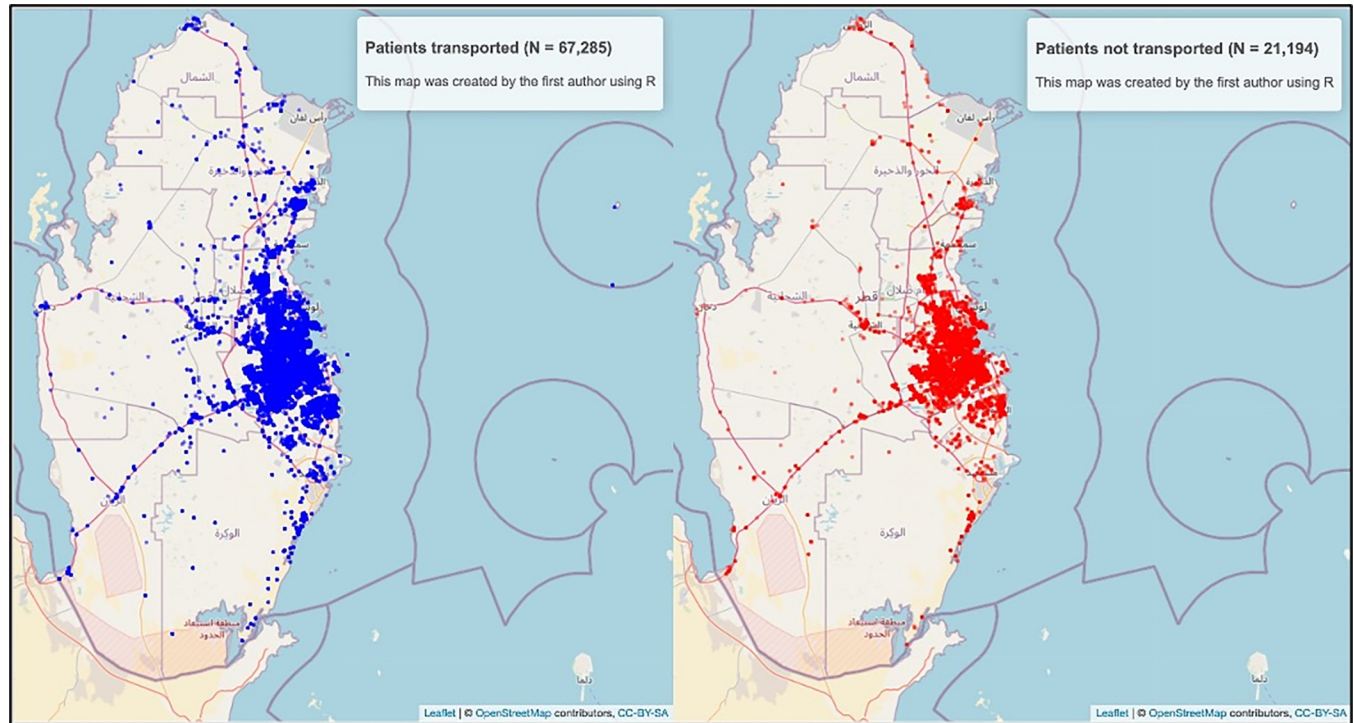


Fig 2. Confusion matrix plots of the four algorithms.

<https://doi.org/10.1371/journal.pone.0301472.g002>

XGBoost performed similarly, although with a slightly improved predicted True Negative rate ($n = 1,869$) but a considerable number of False Positives ($n = 2,369$) and False Negatives ($n = 621$). SVM, with an extremely high True Positive rate (1,3257), evidently failed in identifying “Not Transported” cases, giving an extremely low True Negative ($n = 306$) and high False Positive ($n = 3,932$) cases. AdaBoost showed a more balanced performance than SVM, with a commendable True Positive rate ($n = 13,110$) and a better True Negative rate ($n = 737$), although it also predicted a significant number of False Positives ($n = 3,501$) and False Negatives ($n = 347$).

Additionally, “Fig 3” was created to identify the most important variables in the predictive models. The call-tacking “ProtocolName” and “ProvisonalDiagnoses_CAT” were the most significant features across all variables.

In “Figs 4 and 5”, and “S5 File”, the prediction data showcases the frequency of 999 calls categorised into (Yes = Transported) and (No = Not Transported) across different days of the week and hours of the day. For RF and XGBoost, the best performing models, Sunday’s 999 call rates remained relatively stable between both models, with only a slight decrease in XGBoost, whilst certain days like Monday showed more variability. The late-night and early-morning hours consistently have fewer 999 calls across all days in both datasets. The frequency of transported patients tends to be higher during the day than at night for almost all days of the week (Except on Fridays). Peaks in the transported patients can be seen during late morning to early afternoon hours, notably for Mondays and Wednesdays. For the “Not Transported” patients, the frequency generally remains lower compared to “Transported” patients but still shows some hourly variations.

The following examples are illustrated to explain how to use the data in “S5 File” for the prediction of 999 calls using the XGBoost algorithm, also considering the other predictors in the model:

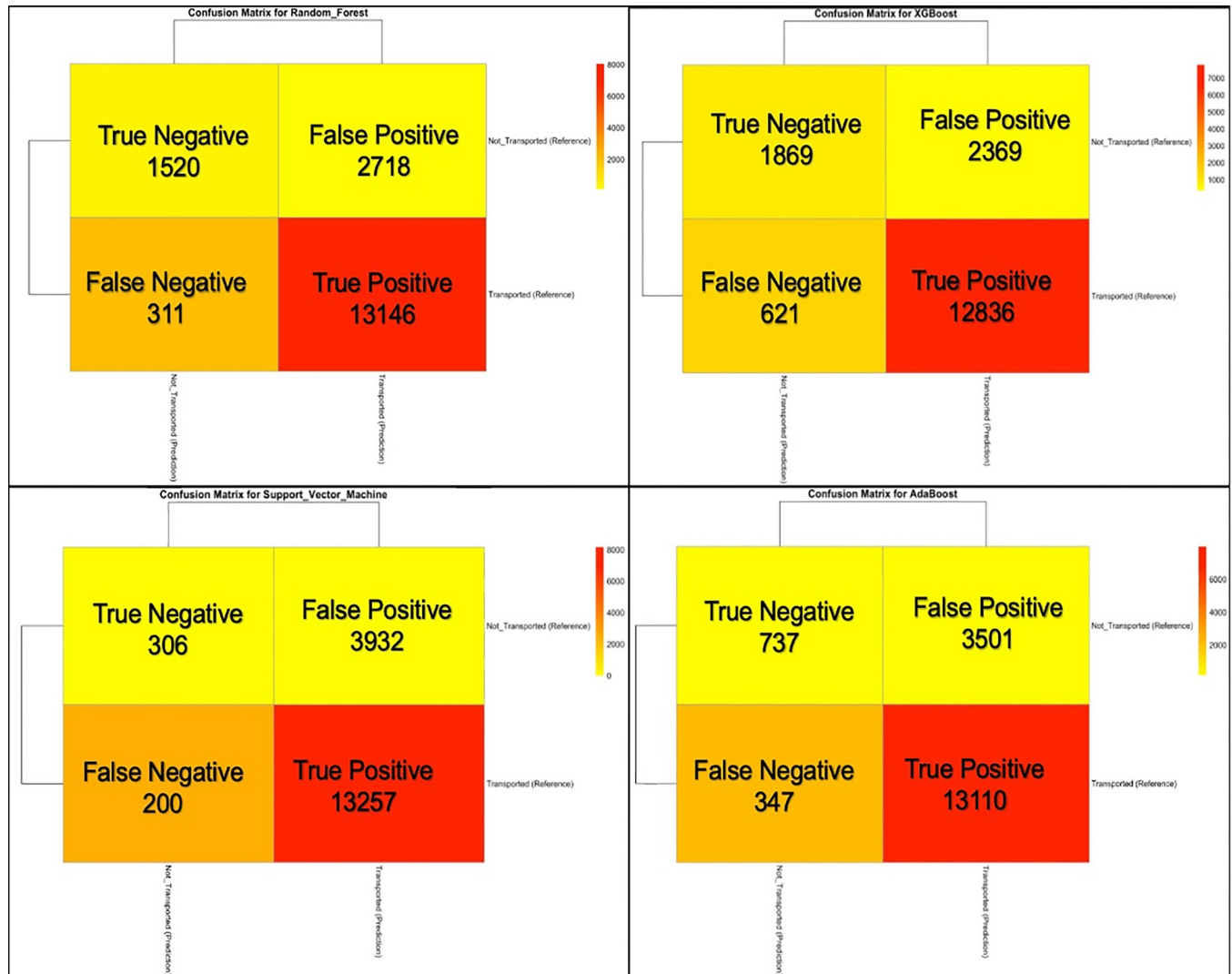


Fig 3. Feature importance of the four machine learning models.

<https://doi.org/10.1371/journal.pone.0301472.g003>

- i. **Monday at Hour 0:** On Mondays at midnight, we can expect approximately 50 are predicted to be transported, while 15 will likely not be transported.
- ii. **Wednesday at Hour 11:** On Wednesdays at 11 AM, we can expect approximately 385 to be transported, while 61 will likely not be transported.
- iii. **Friday at Hour 18:** On Fridays at 6 PM, we can expect approximately about 50 are predicted to be transported, while 12 will likely not be transported.
- iv. **Sunday at Hour 22:** On Sundays at 10 PM, we can expect approximately 62 are predicted to be transported, while 14 will likely not be transported.

Discussion

In the feature selection process, variables such as “PatientTriageArea” and “PriorityToHospital” emerged as significant predictors for constructing a robust model. However, these

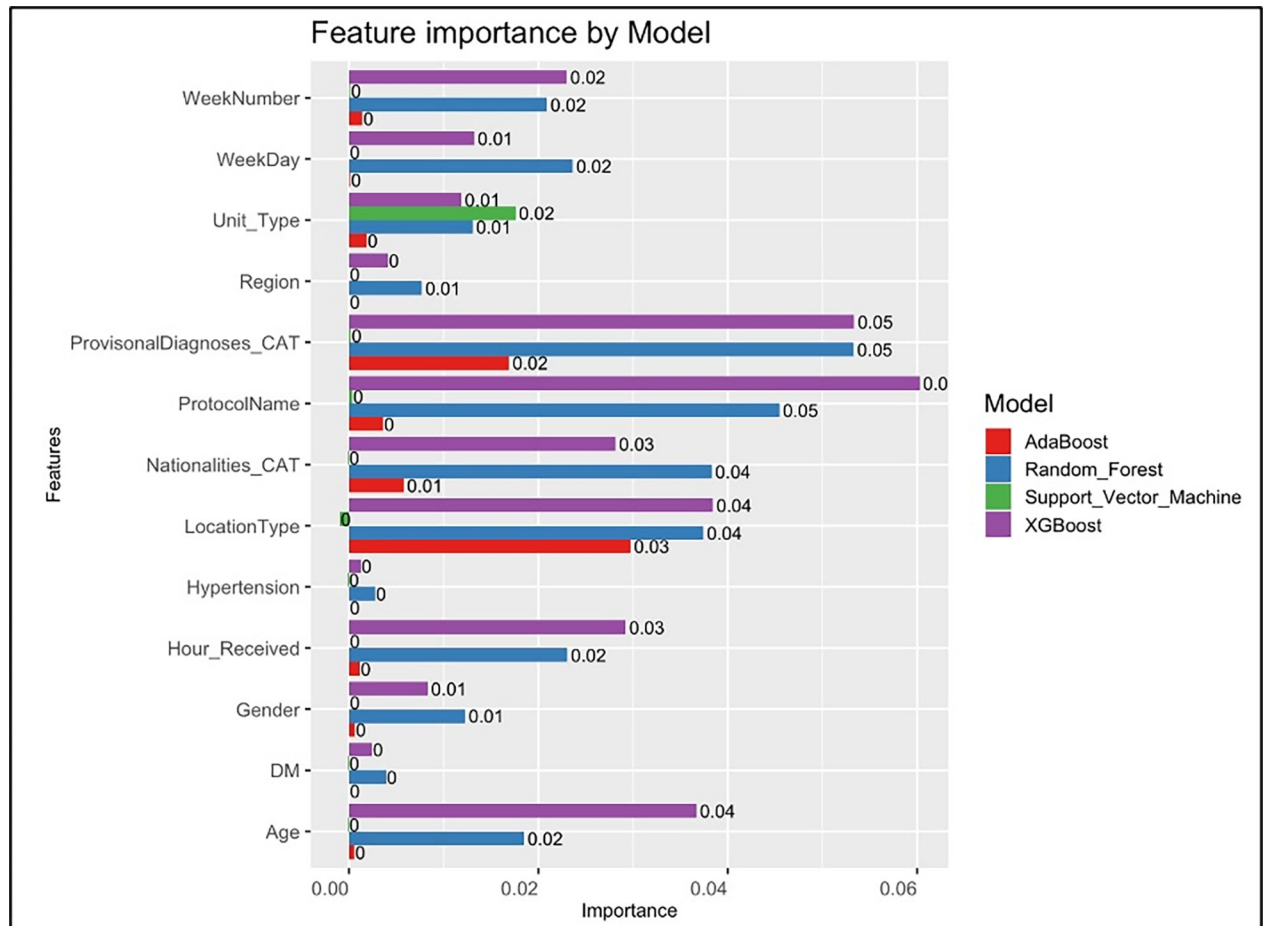


Fig 4. Statistical process control of the prediction data of the hourly patients not transported per model.

<https://doi.org/10.1371/journal.pone.0301472.g004>

variables were not incorporated into the final models for several reasons. Firstly, the area where a patient is triaged and the transport priority are frequently subject to change. Specifically, the patient triage area is often not determined until the ambulance is near the hospital. This is because patients' medical conditions are dynamic, evolving during the transport journey, thereby affecting the final decision regarding identifying their destination area and the priority level, when the patient has already approved to be transported. Secondly, alternative feature selection methods, such as the Chi-square test in "S1 File", revealed that most variables correlated highly with the outcome variable. This suggests that they, too, could substantially contribute to constructing a predictive model. Therefore, while automated feature selection techniques provide valuable insights, it is crucial also to consider other statistical and non-statistical measures to ensure a comprehensive understanding of how different variables could influence the model's performance. This multifaceted approach enriches the robustness of the model, accommodating wider clinically and operationally significant factors like the call-taking protocol name. Moreover, prior studies have emphasised that feature selection serves as a guide towards pertinent variables, and features can be selected based on both their mutual information and correlation coefficients [11, 21, 22]. Nevertheless, the expertise of the researchers remains integral in selecting variables for inclusion in the predictive model. Ultimately, the model's performance metrics serve as the definitive evaluation of its efficacy and

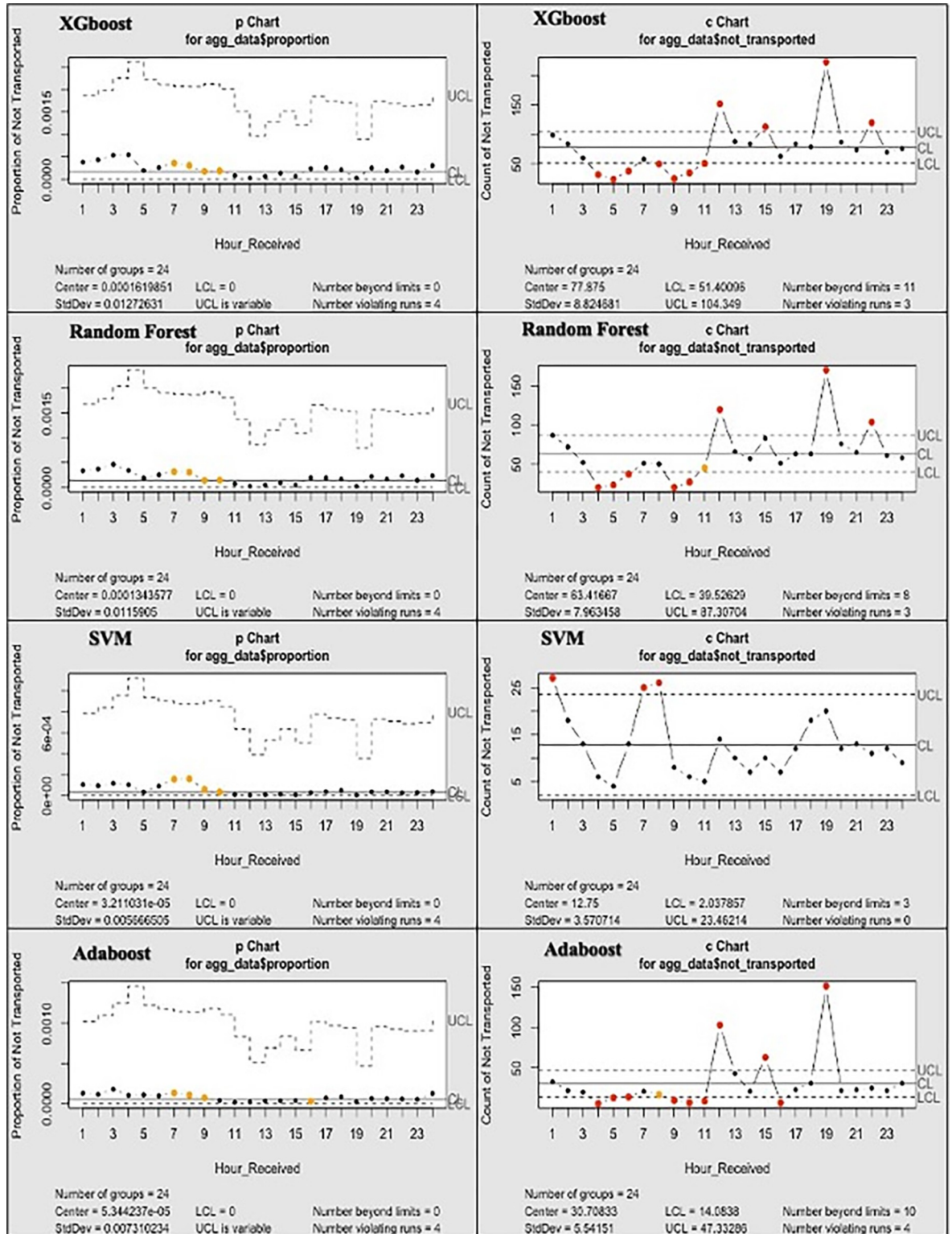


Fig 5. Hourly and weekly number of patients predicted to be transported to the hospital for all models.

<https://doi.org/10.1371/journal.pone.0301472.g005>

reliability. While feature selection algorithms can inform variable importance, the researchers' nuanced understanding and domain knowledge are indispensable for robust and clinically relevant modelling.

Effective decision-making is often encumbered by the lack of robust, predictive data analytics in pre-hospital care [23]. Integrating ML techniques into healthcare has foreshadowed a new era of efficiency, predictability, and QI, particularly in pre-hospital settings [3, 24]. This study demonstrated the robustness of ML algorithms with high accuracy rates in predicting whether an emergency call would result in the patient being "Transported" or not. These findings have multiple implications for HMCAS, offering avenues for systemic QI interventions that were previously untapped.

Firstly, the predictive prowess of these algorithms can serve as a valuable tool for resource optimisation. When medical resources are finite, but demands are often unpredictable and urgent, allocating resources judiciously might be essential. In our case, RF and XGBoost models can help identify high-risk time periods likely to experience peaks in emergencies. By doing so, HMCAS can pre-position ambulances, personnel, and all other necessary equipment, thereby reducing response times and potentially increasing the effectiveness of the care provided. Many recent studies explored the potential for predictive analytics to improve pre-hospital medical teams' performance [3, 25].

Secondly, the high specificity rates of these models can drastically reduce the incidence of false positives, ensuring that resources are allocated only where truly needed. In the context of patient transport decisions, false positives are erroneously predicting that a patient needs to be transported when they do not, which can lead to unnecessary use of pre-hospital resources. This strains healthcare systems and could divert critical resources from patients requiring urgent care. Therefore, even models with high specificity must be cautiously interpreted and continually validated to minimise these risks [26, 27].

The feature importance abilities of ML models (Fig 3) enabled the precise identification of key predictive variables, such as the ProQA™ protocol name and provisional diagnoses in our case. Based on these specific variables, this enhances the pre-hospital system's proficiency in identifying patients at risk of not being transported. Such targeted predictive analysis can substantially ease financial and logistical pressures on pre-hospital care systems, for instance, by fine-tuning dispatch protocols to recognise better patients with low-acuity complaints who may request not to be transported. In such instances, alternative resources like "non-transport units" equipped with less costly medical equipment could be allocated rather than dispatching emergency response units, which are over-equipped and crewed by highly qualified staff like critical care paramedics as they are designed for more critical patients [28, 29]. This approach may not only lead to fuel, workforce, and vehicle maintenance savings. It could also mitigate the risks associated with high-priority driving, with increased risk of road traffic accidents and under-triage associated with not transporting a patient [30–33]. It is crucial to highlight that numerous recent studies have accentuated the significant risks of failing to transport patients who have called for pre-hospital emergency assistance, thereby depriving them of advanced in-hospital care [34, 35].

Thirdly, ML algorithms are fundamentally adaptive and can continually evolve when updated with new data, offering dynamic models that can adjust to societal behaviour, technological progress or healthcare guidelines [36, 37]. This adaptability is crucial for continuous QI as it allows pre-hospital EMS to be agile and responsive to new challenges or shifts in pre-hospital care needs [38]. For example, the ML models could be updated to consider the emergence of new health crises such as pandemics, seasonal trends in certain medical conditions, or changes in urban layouts that can affect transport times. Incorporating these findings into operational protocols can potentially revolutionise the pre-hospital care paradigm. For

instance, ambulances and medical teams could be pre-positioned strategically during identified peak hours, thus decreasing response times and enhancing patient outcomes [39, 40]. Additionally, staff rosters could be adapted to align with these temporal patterns, ensuring sufficient medical personnel are available during high-demand periods. The availability of advanced telemedicine options could also serve as alternative care pathways for non-transport cases, thus averting unnecessary pressure on hospitals.

Finally, it is important to note that while ML holds great promise for enhancing EMS capabilities, its effective implementation necessitates a multi-disciplinary approach that includes healthcare professionals, data scientists, and policymakers. It is not only about algorithmic sophistication but also about ethical considerations, data privacy, and regulatory compliance. Further, published studies cautioned against over-reliance on algorithms, highlighting ethical concerns such as data misuse, patient confidentiality, and the potential for algorithmic bias [41–43]. The incorporation of ML into pre-hospital EMS practices represents more than just a technological advance; it signifies a paradigm shift with the potential to fundamentally transform the benchmarks of emergency pre-hospital care [44, 45]. Considering these complexities, integrating continuous QI principles with ML becomes instrumental when implementing predictive pre-hospital care models. These QI methodologies, such as those derived from Deming and Kaizen principles, are helpful since they focus on iterative testing and constant monitoring, ensuring that the ML algorithms (S6 File), when deployed, remain accurate, ethical, and aligned with healthcare objectives, enabling real-time assessment and their fine-tuning [46, 47]. Concurrently, the culture of “continuous improvement” can perfectly merge with the adaptive nature of ML, encouraging a culture of regular feedback and instantaneous adjustments, creating a self-updating system that can mitigate risks and dynamically evolve to meet the ever-changing demands on pre-hospital care.

Limitations

While the models demonstrated commendable accuracy rates, it is noteworthy to mention that their sensitivity rates, particularly in the case of SVM, dipped to an inadequate 7.2% highlighting existing limitations in identifying “true positives”. This suggests that the models are potentially more inclined towards avoiding false positives at the expense of missing out on genuine emergency cases [48]. In this study, while all four algorithms excelled in sensitivity by correctly predicting a high number of “Transported” cases, they showed varying degrees of specificity and error control, highlighting the need for further fine-tuning to improve overall performance. Such a gap in model performance highlights the pressing need for future research to refine these algorithms to achieve a more balanced interplay between sensitivity and specificity with new data. Moreover, missing values in the dataset could contribute to this issue. Filling in these gaps is more than a data-cleaning operation, it is also a fundamental step towards improving the robustness of predictive modelling. Missing data, if not addressed, could introduce bias and compromise the validity of the findings, thus mitigating the utility of ML in this critical context. Given this, we recommend that HMCAS reviews policies and procedures, specifically to address non-systematic missing data entry, which will pave the way for more robust data analysis and predictive modelling in the future.

Conclusion

In conclusion, this study serves as a seminal exploration into utilising ML algorithms for predicting transport decisions, optimising resource allocation and improving the quality of pre-hospital emergency care. It presented compelling evidence for the efficacy of predictive modelling in differentiating “Transported” from “Not Transported” patients, thereby aiding in

identifying their transport decision-making process in advance. While the study aligns with previous research in underlining the criticality of timely interventions and the need for nuanced, data-driven approaches, it also accentuates areas needing further refinement—namely, the improvement of model sensitivity and the imperative of handling missing data for more robust predictions. In doing so, the study sets the stage for future investigations that could refine these predictive algorithms and further enhance HMCAS' responsiveness and effectiveness. The findings offer valuable contributions to personalised, efficient, and pre-hospital life-saving medical interventions as we move towards an increasingly data-centric health-care paradigm.

Supporting information

S1 File. Summary statistics and test results of the variables for patients transported and not transported.

(PDF)

S2 File. Summary of the missing values before and after the imputation process.

(PDF)

S3 File. Boxplots for continuous variables data visualisation.

(PDF)

S4 File. Tables of the automated feature selection methods.

(PDF)

S5 File. Table of the four models' predictions (Yes = Transported; No = Not Transported).

(PDF)

S6 File. The R-code that was utilised for the predictive modelling.

(PDF)

Acknowledgments

We wish to acknowledge all the executive and management teams in the HMCAS BI Team.

Open access funding was provided by Qatar National Library.

Author Contributions

Data curation: Hassan Farhat.

Formal analysis: Hassan Farhat.

Methodology: Hassan Farhat.

Project administration: Moncef Khadhraoui, James Laughton, Guillaume Alinier.

Resources: Mohamed Chaker Khenissi, Nicholas Castle, Loua Al-Shaikh, Imed Gargouri.

Software: Hassan Farhat.

Supervision: Moncef Khadhraoui, Imed Gargouri, James Laughton.

Validation: Ahmed Makhoulouf, Cyrine Abid, Guillaume Alinier.

Visualization: Hassan Farhat.

Writing – original draft: Hassan Farhat.

Writing – review & editing: Padarath Gangaram, Kawther El Aifa, Ian Howland, Fatma Babay Ep Rekik, James Laughton, Guillaume Alinier.

References

1. Alhabdan N., Alhusain F., Alharbi A., Alsadhan M., Hakami M., and Masuadi E., 'Exploring emergency department visits: factors influencing individuals' decisions, knowledge of triage systems and waiting times, and experiences during visits to a tertiary hospital in Saudi Arabia', *Int. J. Emerg. Med.*, vol. 12, no. 1, p. 35, Nov. 2019, <https://doi.org/10.1186/s12245-019-0254-7> PMID: 31752662
2. M. A. Ahmad, C. Eckert, and A. Teredesai, 'Interpretable Machine Learning in Healthcare', in *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, in BCB '18. New York, NY, USA: Association for Computing Machinery, Aug. 2018, pp. 559–560. <https://doi.org/10.1145/3233547.3233667>
3. Olave-Rojas D. and Nickel S., 'Modeling a pre-hospital emergency medical service using hybrid simulation and a machine learning approach', *Simul. Model. Pract. Theory*, vol. 109, p. 102302, May 2021, <https://doi.org/10.1016/j.simpat.2021.102302>
4. Al Muftah H., 'Demographic Policies and Human Capital Challenges', in *Policy-Making in a Transformative State: The Case of Qatar*, Tok M. E., Alkhater L. R. M., and Pal L. A., Eds., London: Palgrave Macmillan UK, 2016, pp. 271–294. https://doi.org/10.1057/978-1-137-46639-6_10
5. Teisberg E., Wallace S., and O'Hara S., 'Defining and Implementing Value-Based Health Care: A Strategic Framework', *Acad. Med.*, vol. 95, no. 5, pp. 682–685, May 2020, <https://doi.org/10.1097/ACM.0000000000003122> PMID: 31833857
6. Collins G. S., Reitsma J. B., Altman D. G., and Moons K. G. M., 'Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD)', *Circulation*, vol. 131, no. 2, pp. 211–219, Jan. 2015, <https://doi.org/10.1161/CIRCULATIONAHA.114.014508> PMID: 25561516
7. Adnan K., Akbar R., Khor S. W., and Ali A. B. A., 'Role and Challenges of Unstructured Big Data in Healthcare', in *Data Management, Analytics and Innovation*, Sharma N., Chakrabarti A., and Bsalas V. E., Eds., in *Advances in Intelligent Systems and Computing*. Singapore: Springer, 2020, pp. 301–323. https://doi.org/10.1007/978-981-32-9949-8_22
8. Clwason J. J., Dernocoeur K. B., and Murray C., *Principles of Emergency Medical Dispatch, 5th edition*, 5th edition. USA: The International Academy of EMD, 2014.
9. Sievert C., *Interactive Web-Based Data Visualization with R, plotly, and shiny*. CRC Press, 2020.
10. van Buuren S., *Flexible Imputation of Missing Data*, Second Edition. CRC Press, 2018.
11. Kuhn M. and Johnson K., *Feature Engineering and Selection: A Practical Approach for Predictive Models*. CRC Press, 2019.
12. Theerthagiri P., 'Predictive analysis of cardiovascular disease using gradient boosting based learning and recursive feature elimination technique', *Intell. Syst. Appl.*, vol. 16, p. 200121, Nov. 2022, <https://doi.org/10.1016/j.iswa.2022.200121>
13. Horvat T. and Job J., 'The use of machine learning in sport outcome prediction: A review', *WIREs Data Min. Knowl. Discov.*, vol. 10, no. 5, p. e1380, 2020, <https://doi.org/10.1002/widm.1380>
14. Jena O. P., Bhushan B., and Kose U., *Machine Learning and Deep Learning in Medical Data Analytics and Healthcare Applications*. CRC Press, 2022.
15. Sierra-Sosa D. et al., 'Scalable Healthcare Assessment for Diabetic Patients Using Deep Learning on Multiple GPUs', *IEEE Trans. Ind. Inform.*, vol. 15, no. 10, pp. 5682–5689, Oct. 2019, <https://doi.org/10.1109/TII.2019.2919168>
16. Topçuoğlu B. D., Lapp Z., Sovacool K. L., Snitkin E., Wiens J., and Schloss P. D., 'mikropml: User-Friendly R Package for Supervised Machine Learning Pipelines', *J. Open Source Softw.*, vol. 6, no. 61, p. 3073, 2021, <https://doi.org/10.21105/joss.03073> PMID: 34414351
17. Nguyen T. T., Nguyen D. D., Nguyen S. D., Prakash I., Tran P. V., and Pham B. T., 'Forecasting Construction Price Index using Artificial Intelligence Models: Support Vector Machines and Radial Basis Function Neural Network', *J. Sci. Transp. Technol.*, pp. 9–19, Dec. 2022, <https://doi.org/10.58845/jstt.utt.2022.en.2.4.9-19>
18. Aggarwal A. et al., 'COVID-19 Risk Prediction for Diabetic Patients Using Fuzzy Inference System and Machine Learning Approaches', *J. Healthc. Eng.*, vol. 2022, p. e4096950, Apr. 2022, <https://doi.org/10.1155/2022/4096950> PMID: 35368915
19. Farhat H. et al., 'Quality improvement tools to manage emergency callbacks from patients with diabetes in a prehospital setting', *BMJ Open Qual.*, vol. 12, no. 1, p. e002007, Jan. 2023, <https://doi.org/10.1136/bmjoc-2022-002007> PMID: 36599502

20. Singh Ajit Pal, 'QUALITY IMPROVEMENT USING STATISTICAL PROCESS CONTROL TOOLS IN GLASS BOTTLES MANUFACTURING COMPANY', *Int. J. Qual. Res.*, vol. Volume17, no. Issue1, p. 20, 2013.
21. Kou G., Yang P., Peng Y., Xiao F., Chen Y., and Alsaadi F. E., 'Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods', *Appl. Soft Comput.*, vol. 86, p. 105836, Jan. 2020, <https://doi.org/10.1016/j.asoc.2019.105836>
22. Zhou H., Wang X., and Zhu R., 'Feature selection based on mutual information with correlation coefficient', *Appl. Intell.*, vol. 52, no. 5, pp. 5457–5474, Mar. 2022, <https://doi.org/10.1007/s10489-021-02524-x>
23. Darnall N. and Edwards D. Jr., 'Predicting the cost of environmental management system adoption: the role of capabilities, resources and ownership structure', *Strateg. Manag. J.*, vol. 27, no. 4, pp. 301–320, 2006, <https://doi.org/10.1002/smj.518>
24. Rathore N., Jain P. K., and Parida M., 'A Sustainable Model for Emergency Medical Services in Developing Countries: A Novel Approach Using Partial Outsourcing and Machine Learning', *Risk Manag. Healthc. Policy*, vol. 15, pp. 193–218, Dec. 2022, <https://doi.org/10.2147/RMHP.S338186> PMID: 35173497
25. Yu M., Kollias D., Wingate J., Siriwardena N., and Kollias S., 'Machine Learning for Predictive Modelling of Ambulance Calls', *Electronics*, vol. 10, no. 4, Art. no. 4, Jan. 2021, <https://doi.org/10.3390/electronics10040482>
26. Yan G. et al., 'Covert COVID-19 and false-positive dengue serology in Singapore', *Lancet Infect. Dis.*, vol. 20, no. 5, p. 536, May 2020, [https://doi.org/10.1016/S1473-3099\(20\)30158-4](https://doi.org/10.1016/S1473-3099(20)30158-4) PMID: 32145189
27. Sayghe A. et al., 'Survey of machine learning methods for detecting false data injection attacks in power systems', *IET Smart Grid*, vol. 3, no. 5, pp. 581–595, 2020, <https://doi.org/10.1049/iet-stg.2020.0015>
28. Farhat H. et al., 'Epidemiological Determinants of Patient Non-Conveyance to the Hospital in an Emergency Medical Service Environment', *Int. J. Environ. Res. Public Health*, vol. 20, no. 14, Art. no. 14, Jan. 2023, <https://doi.org/10.3390/ijerph20146404> PMID: 37510636
29. Gangaram P., Menacho A. M., and Alinier G., 'Crisis resource management in relation to empowering people to speak up in emergency medical service clinical practice settings', *J. Paramed. Pract.*, vol. 9, no. 2, pp. 60–65, Feb. 2017, <https://doi.org/10.12968/jpar.2017.9.2.60>
30. G. Alinier et al., 'Initial results from a driving safety survey distributed to ambulance paramedics in Qatar', *J. Local Glob. Health Sci.*, vol. 2015, no. Proceedings of the 24th World International Traffic Medicine Association Congress, Qatar 2015, p. 59, Nov. 2015, <https://doi.org/10.5339/jlghs.2015.itma.59>
31. Waalwijk J. F. et al., 'Evaluating the effect of driving distance to the nearest higher level trauma centre on undertriage: a cohort study', *Emerg. Med. J.*, vol. 39, no. 6, pp. 457–462, Jun. 2022, <https://doi.org/10.1136/emered-2021-211635> PMID: 34593562
32. Forsgärde E.-S., Elmqvist C., Fridlund B., Svensson A., Andersson R., and Rööst M., 'Patients' aged ≥ 65 years dispositions during ambulance assignments, including factors associated with non-conveyance to hospital: a longitudinal and comparative study', *BMJ Open*, vol. 10, no. 11, p. e038885, Nov. 2020, <https://doi.org/10.1136/bmjopen-2020-038885> PMID: 33243795
33. Ebben R. H. A. et al., 'A patient-safety and professional perspective on non-conveyance in ambulance care: a systematic review', *Scand. J. Trauma Resusc. Emerg. Med.*, vol. 25, no. 1, p. 71, Jul. 2017, <https://doi.org/10.1186/s13049-017-0409-6> PMID: 28716132
34. Laukkanen L., Lahtinen S., Raatiniemi L., Ehrola A., Kaakinen T., and Liisanantti J., 'Emergency department admission and mortality of the non-transported emergency medical service patients: a cohort study from Northern Finland', *Emerg. Med. J.*, vol. 39, no. 6, pp. 443–450, Jun. 2022, <https://doi.org/10.1136/emered-2020-209914> PMID: 33879493
35. Carrigan S., Goldstein J., Carter A., Asada Y., and Travers A., 'The Prevalence and Characteristics of Non-Transports in a Provincial Emergency Medical Services System: A Population-Based Study', *J. Emerg. Med.*, vol. 62, no. 4, pp. 534–544, Apr. 2022, <https://doi.org/10.1016/j.jemermed.2021.12.009> PMID: 35131130
36. Futoma J., Simons M., Panch T., Doshi-Velez F., and Celi L. A., 'The myth of generalisability in clinical research and machine learning in health care', *Lancet Digit. Health*, vol. 2, no. 9, pp. e489–e492, Sep. 2020, [https://doi.org/10.1016/S2589-7500\(20\)30186-2](https://doi.org/10.1016/S2589-7500(20)30186-2) PMID: 32864600
37. Riley R. D. et al., 'External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges', *BMJ*, vol. 353, p. i3140, Jun. 2016, <https://doi.org/10.1136/bmj.i3140> PMID: 27334381
38. Wilson A., Saeed H., Pringle C., Eleftheriou I., Bromiley P. A., and Brass A., 'Artificial intelligence projects in healthcare: 10 practical tips for success in a clinical environment', *BMJ Health Care Inform.*, vol. 28, no. 1, p. e100323, Jul. 2021, <https://doi.org/10.1136/bmjhci-2021-100323> PMID: 34326160

39. Alinier G., Wilson P., Reimann T., and Morris B., 'Influential factors on urban and rural response times for emergency ambulances in Qatar', *Mediterr. J. Emerg. Med.*, no. 26, Art. no. 26, 2018.
40. Aljaziri Abdunnassir FH and Alinier, 'Access to Emergency Healthcare', in *Healthcare Access*, UK: IntechOpen, 2022, p. 171.
41. Richens J. G., Lee C. M., and Johri S., 'Author Correction: Improving the accuracy of medical diagnosis with causal machine learning', *Nat. Commun.*, vol. 12, no. 1, Art. no. 1, Mar. 2021, <https://doi.org/10.1038/s41467-021-21494-9> PMID: 33790263
42. Chen R. J., Lu M. Y., Chen T. Y., Williamson D. F. K., and Mahmood F., 'Synthetic data in machine learning for medicine and healthcare', *Nat. Biomed. Eng.*, vol. 5, no. 6, Art. no. 6, Jun. 2021, <https://doi.org/10.1038/s41551-021-00751-8> PMID: 34131324
43. Mhasawade V., Zhao Y., and Chunara R., 'Machine learning and algorithmic fairness in public and population health', *Nat. Mach. Intell.*, vol. 3, no. 8, Art. no. 8, Aug. 2021, <https://doi.org/10.1038/s42256-021-00373-4>
44. B. Blobel and M. Giacomini, *pHealth 2019: Proceedings of the 16th International Conference on Wearable Micro and Nano Technologies for Personalized Health 10–12 June 2019*, Genoa, Italy. IOS Press, 2019.
45. Al-Sarem M., Boulila W., Al-Harby M., Qadir J., and Alsaeedi A., 'Deep Learning-Based Rumor Detection on Microblogging Platforms: A Systematic Review', *IEEE Access*, vol. 7, pp. 152788–152812, 2019, <https://doi.org/10.1109/ACCESS.2019.2947855>
46. Rüttimeann B. G., *Transactional Lean: Preparing for the Digitalization Era: A Systematic Approach to Industrialize Office Processes*. Cham: Springer International Publishing, 2019. <https://doi.org/10.1007/978-3-030-22860-6>
47. Willis J., *Deming's Journey to Profound Knowledge: How Deming Helped Win a War, Altered the Face of Industry, and Holds the Key to Our Future*. IT Revolution, 2023.
48. Kaur H., Pannu H. S., and Malhi A. K., 'A Systematic Review on Imbalanced Data Challenges in Machine Learning: Applications and Solutions', *ACM Comput. Surv.*, vol. 52, no. 4, p. 79:1–79:36, Aug. 2019, <https://doi.org/10.1145/3343440>