Heart Rate and Body Temperature Relationship in Children Admitted to PICU: A Machine Learning Approach

Emilie Lu, Thanh-Dung Le, *Member, IEEE*, Philippe Jouvet, and Rita Noumeir, *Member, IEEE*

Abstract—**Vital signs are crucial clinical measures, with body temperature (BT) and heart rate (HR) being particularly significant. While their association has been studied in adults and children, research in Pediatric Intensive Care Unit (PICU) settings remains limited despite the critical conditions of these patients.** *Objective:* **This study examines the relationship between HR and BT in children aged 0 to 18 admitted to the PICU at CHU Sainte-Justine (CHUSJ) Hospital.** *Methods:* **Machine learning (ML) techniques, including Gradient Boosting Machines (GBM) with Quantile Regression (QR), were applied to capture the relationship between HR, BT, and age, optimizing model performance through hyperparameter tuning.** *Results:* **Analyzing data from 4006 children, we observed a consistent trend of decreasing HR with increasing age and rising HR with higher BT ranges. Linear models often underestimated HR at lower BT ranges and overestimated it at higher ranges, especially in younger age groups. The GBM model demonstrated improved accuracy and supported a user-friendly interface for HR predictions based on BT, age, and HR percentiles. Qualitative observations indicated that linear models underestimated HR at lower BT ranges and overestimated it at higher ones, particularly in younger children. These findings challenge the direct linear association assumed in prior studies.** *Conclusion:* **This study provides new insights into the non-linear dynamics between HR, BT, and age in critically ill children, emphasizing further research to quantify and understand these relationships.** *Significance:* **By refining predictive models and re-evaluating traditional assumptions, this work provides valuable insights for improving clinical decision-making in PICU settings.**

I. INTRODUCTION

PEDIATRIC Intensive Care Unit (PICU) patients' health requires particular attention and ongoing monitoring [1]. Due to their severe conditions or illnesses, these children, particularly those aged 0 to 18 years old, have distinct physiological characteristics requiring special medical evaluation and care [2]. In the PICU of CHUSJ, caregivers often use the Clinical decision support systems (CDSS) in their practice to improve the quality of patient care and outcome [3]. In this high-risk environment, healthcare practitioners rely on vital sign monitoring as fundamental indicators of patients' health across all age groups during their hospitalization [4]. These indicators assist in making clinical decisions and facilitate prompt interventions if necessary. Among these vital signs, heart rate (HR) and body temperature (BT) are crucial parameters providing essential information about a child's health, as their relationship can be influenced by age, agitation, stress, infection, shock, physiological distress, or underlying illness [5], [6]. Consequently, HR and BT offer invaluable insights for guiding physicians in clinical decision-making [7].

Firstly, HR is fundamental in assessing a person's overall cardiac health [8]. This metric varies depending on the child's age, as highlighted by Pediatric Advanced Life Support (PALS) data, where awake patients' reference HR values are age-dependent [6], [9]. Newborns under 28 days of age typically have an HR ranging from 85 to 205 beats per minute (bpm), while infants and toddlers aged 1 to 24 months fall between 100 and 190 bpm. In children from 2 to 10 years old, the normal HR ranges between 60 and 140 bpm. Adolescents between 10 and 18 years old usually maintain an HR averaging between 60 and 100 bpm. Monitoring patients' HR can be an early indicator of various medical conditions [10]. Deviation from the normal HR range can signal underlying health conditions or potential complications. For example, bradycardia, an abnormally low HR, leads to reduced blood flow to organs, leading to dizziness, fatigue, shortness of breath, loss of consciousness, and signs of heart failure [11]. Conversely, tachycardia, i.e., high HR, can indicate a severe medical condition, such as shock [12].

Additionally, BT regulation is vital for assessing an individual's overall health and reflecting the body's heat production and balance, which are crucial for optimal physiological functions. The regulation of BT is managed by the hypothalamus, a central region within the brain that acts as the body's temperature control center. This process, called thermoregulation, aims to adjust BT based on signals perceived by thermoreceptors in blood and skin. The primary goal is maintaining a normal BT around $37 \pm 0,5\,°C$, varying slightly in infants and young children. The hypothalamus plays a crucial role in initiating necessary adjustments in temperature fluctuations, whether a decrease or an increase. Extreme temperatures beyond normal ranges indicate health issues: low temperatures (hypothermia) suggest cold responses or medical problems that can potentially lead to severe complications. Elevated temperatures (hyperthermia), triggered by various external factors, such as infection, environmental temperature, or disruptions in the body's thermoregulation system, often signals illness. Consistent temperature checks are important in early illness detection, enabling timely medical interventions. This practice is essential for taking precautionary measures to avert physiological dysfunction and protect organs' structure and function from damage [13].

A distinct pattern is observed in the pediatric population from 0 to 18 years, where HR tends to decrease with age after the initial month. Notably, there is an observed tendency for HR to rise during the first month following birth, reaching its peak before gradually decreasing [9]. Insights from studies on hospitalized children in primary care and emergency departments consistently show that BT is an independent factor influencing HR, resulting in an average increase of around 10 bpm per degree centigrade [6], [10]. More specifically, the studies [14] and [12] highlighted an age-specific influence of BT on HR. Every $1\,°C$ BT increase in children with acute infections correlates with an HR rise ranging from 9.9 to 14.1 bpm [14]. In contrast, in children aged 0–16 years old attending urgent and emergency care, the influence of a $1\,°C$ increase in HR shows a range of 8.7 to 13.7 bpm, with the observed average being 12.3 bpm [12].

In adults admitted to the emergency department, it has been observed that a $1\,°C$ increase in BT leads to an approximately 7 bpm increase in HR [15]. In contrast, among acute patients within the emergency department, the results indicate that the group of patients with a high BT, i.e., greater than $37.2\,°C$, exhibited the most significant changes in HR, with an increase of 7.4 ± 0.9 bpm per degree [5]. In critically ill adults admitted to the Intensive Care Unit (ICU), it was observed that within the temperature range of $32.0°$ to $42.0\,°C$, a linear correlation existed between BT and HR. Specifically, for every $1\,°C$ increase within this range, the HR showed a linear increase of 8.35 ± 0.50 bpm $(p < 0.0001)$ [16].

Regarding the mentioned findings above, the diverse techniques employed to model the relationship between HR and BT in children and adults reflect the complexity of this physiological association. In studies with children, approaches ranged from percentile curves using advanced statistical distributions [10] to quantile regression with polynomial equations [6], centile charts
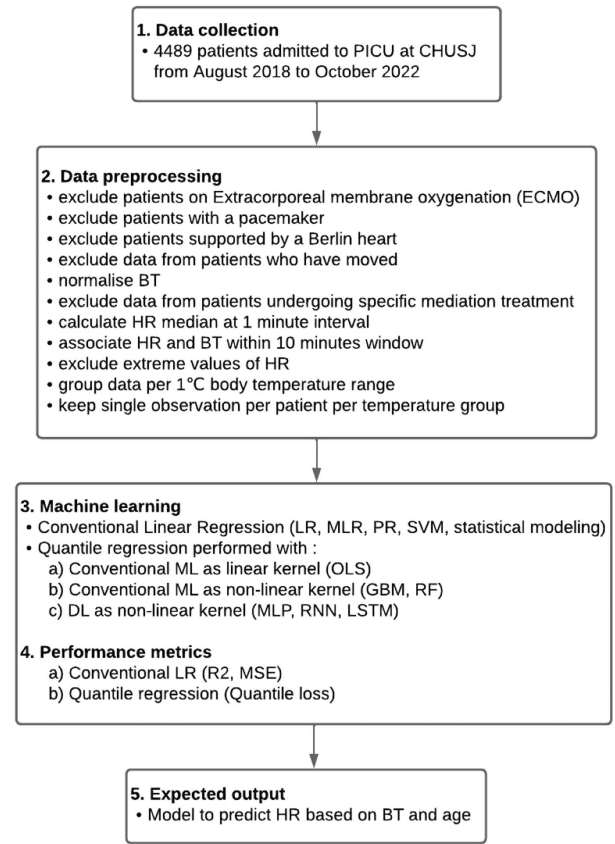

Fig. 1. Workflow to model the relationship between HR and BT.

[14], and multiple linear regression models [12]. Similarly, in adult studies, techniques such as linear regression [15], linear and multiple variable regression analyses [5], and multiple linear regression with various covariates [16] demonstrated a multifaceted understanding of the HR-BT relationship. These methods illustrate the diverse approaches to understanding the HR-BT relationship in different age groups. The findings highlight the need for age-specific considerations, centile charts, and flexible modeling approaches to improve clinical assessments and guide medical interventions. However, the focus on critically ill children remains significantly limited despite the complexity of their medical conditions. Moreover, acknowledging the existing limitations in which traditional methods may not have captured the complexity of the association, the utilization of Machine Learning (ML) and Deep Learning (DL) techniques would enable a more in-depth analysis to accurately capture the intricate patterns between the two vital signs in critically ill children.

II. MATERIALS AND METHODS

This study was conducted following ethical approval from the research ethics board at CHU Sainte-Justine (CHUSJ). The approved protocol number is 2023-5201. Fig. 1 represents an overview of the workflow for this study. There are four main steps, including data collection, data preprocessing, ML

modeling, and evaluation of the expected output. Those steps are discussed in detail as follows:

A. Data Collection

Since 2015, a high-resolution database has been in operation at CHUSJ, collecting clinical data from patients admitted to the PICU at 1-second to 5-second intervals from monitors. Patients were continuously monitored during their stay until their discharge [17]. This project aims to establish a relationship between HR and BT. Patients at CHUSJ PICU were closely monitored for both physiological markers. For BT, continuous monitoring occurred at 30-second intervals, complemented by manual measurements taken every 2-4 hours. The measurement method varied based on the patient's age and included using rectally, esophageal, or axillary temperature probes. HR was monitored using electrodes on the patient's chest, generating an electrocardiogram (ECG). Alternatively, HR could also be obtained using a pulse oximeter attached to the patient's finger, which provides oxygen saturation data and HR measurements.

Children from 0 to 18 years old who were admitted to the PICU between August 2018 and October 2022 with HR and BT records were included in the study. Certain patient groups were excluded from the study in this population to avoid potential confounding factors that interact with BT and HR values. Therefore, removing these data points was essential to achieve accurate modeling. Patients on extracorporeal membrane oxygenation (ECMO), with a pacemaker, or supported by a Berlin heart were excluded. By excluding those patients, our data extraction process yielded information from 4,007 patients admitted between August 2018 and October 2022. From this patient cohort, we extracted patient age at admission, 4 days (96 hours) of HR and BT (value, temperature site, measurement type such as continuous or manual) data along with the corresponding date and time of acquisition, comfort scores and specific medication treatment with the day when the drug treatment was finished.

The average length of stay for the 4,007 pediatric patients included in this study was 7.30 days, with a standard deviation of 29.28 days. We extracted HR and BT data for up to 96 hours (4 days) from each patient's record. For patients with less than 4 days, we used all available HR and BT measurements within the recorded time frame. For patients with more than 4 days of data, only the first 96 hours of data were included in the analysis. This was done to ensure uniformity across the cohort and to focus on the most critical period of the PICU stay, where the majority of significant physiological changes are likely to occur. However, it is possible that including data beyond the 4-day window could provide additional insights, especially for patients with longer stays. Future studies could explore the inclusion of longer observation periods to evaluate its potential impact on the predictive accuracy of our model.

B. Data Preprocessing

Data preprocessing is a fundamental initial step in preparing patient data for the ML algorithm. Effective data preprocessing ensures data quality issues that can impact the validity of any conclusions drawn [18]. To get better reliability of subsequent insights, the process involves the following tasks:

1) Exclude Data From Patients Who Have Moved: Child's movements, including when a patient is restless, crying, or screaming, were excluded from the study. The analysis of the following comfort scores assisted in identifying and excluding data from patients who have moved. In this study, we focused on the relationship between HR and BT during calm, non-moving states because physical activity or agitation can significantly alter this relationship.

1) **Cornell Assessment of Pediatric Delirium (CAPD)** scale was used in PICU for invasively ventilated children who have fluctuations in awareness, attention, and cognition [19]. The evaluation had 8 questions, and each answer was associated with a score between 0 and 4. The total of these scores determined the outcome; delirium would be absent if the scores were less than 9.
2) **Comfort Behavior (Comfort B)** was used to assess pain in ventilated and sedated patients [20]. The score involved observing 6 behaviors (awakening, agitation, ventilation, movement, facial expression, muscle tone) rated on a scale of 1 to 5 to determine the perceived pain level. A total score between 11 to 17 indicated a normal score where the patient had no pain.
3) **FLACC (Face, Legs, Activity, Cry, Consolability)** scale was used to evaluate pain in noninvasive patients under 6 years [19]. Each behavior was scored on a scale from 0 to 2, and patients were included in our study if they had a total score between 0 and 3. This range indicated the absence of pain or mild pain.
4) **revised FLACC (r-FLACC)** was used for patients with intellectual disability [19]. The evaluation was executed the same way as the FLACC scale.
5) **Visual numeric Scale (VNS)** was used to evaluate and monitor pain for patients over 6 years with the ability to communicate [19], [20]. The patient evaluated the pain on a scale of 10, with a score between 0 and 3 indicating the absence of pain or mild pain.
6) **Richmond Agitation-Sedation Scale (RASS)** was a scale used to evaluate agitation and sedation [19]. We only kept data with scores between -5 and $+1$, which signified instances of physical and verbal stimulation.

Based on scores extracted from the database, we assigned movement or not based on the score. Afterward, we associated the scores with the nearest date-time values of BT value since it was monitored at a lower frequency (hours) than HR (seconds). If any of the scores indicated the presence of movement, we excluded the data associated with those movement instances to work with data from calm and non-moving patients.

2) Normalize Body Temperature: We normalized the temperature data, specifically for measurements obtained from the axillary site, to ensure consistency and accuracy in our analysis. The interpretation of axillary temperatures was underestimated by approximately $0.5\,^{\circ}\mathrm{C}$ compared to other methods. Temperature documentation in the PICU was done by recording the actual temperature value displayed on the thermometer. The site from which the temperature was taken (oral, axillary, or

rectal) was also noted. Among these methods, rectal temperature measurement was usually considered the most precise [21].

3) Exclude Data From Patients Undergoing Specific Medication Treatment: Patients in the PICU who received Dexmedetomidine and medication affecting HR were only included once their medication treatment was completed. Specifically, medications used to treat heart conditions that slow down HR included antiarrhythmics, beta-blockers, calcium channel blockers, digoxin, and ivabradine [22], [23]. Conversely, medications that increased HR included Dobutamine, Dopamine, Epinephrine, Milrinone, Norepinephrine (or Noradrenaline), and Salbutamol. To exclude the medication administration period, we excluded data when the monitored timestamp of vital signs coincided with the medication administration.

4) Calculate Heart Rate Median at 1-Minute Interval: In the PICU, HR values were recorded initially every second for all patients. Given the high frequency of measurements, we simplified the data volume by doing data aggregation. More specifically, calculating the median HR at one-minute intervals made the original high-frequency HR data more manageable and insightful.

5) Associate HR and BT Within 10 Minutes Window: Establishing the association between HR and BT was important as it provided valuable insights into the patient's physiological response. Since HR and BT were monitored at different PICU intervals (seconds vs. hours), we associated these two variables by calculating the median of HR measurements within a 10-minute window, more precisely, within ± 5 minutes of each BT timestamp.

6) Exclude Extreme Values of Heart Rate: Extreme HR measurements—values lower than 30 or higher than 240 bpm—were excluded from the dataset. We hypothesized that these numbers could result from unusual clinical diseases or data entry errors, which could have introduced inaccuracies in the analysis. By removing these outliers, we wanted to ensure our dataset was more resilient and reliable. This would have allowed us to base future analysis and modeling efforts on a more accurate depiction of physiological parameters within the target population.

Regarding extreme HR values, we excluded observations with temperatures below 30 °C or above 43 °C and HR values below 30 or above 240 beats per minute, as these readings often indicate data entry errors or rare clinical conditions that could skew the analysis. HR values below 30 or above 240 bpm, as these readings often indicate data entry errors or pre-death (HR below 30/min) or tachy arrythmia (HR above 240/min) conditions that are not linked to core temperature but states that require specific interventions. Our approach aligns with previous studies, such as [10], which also excluded similar extreme values to maintain data quality and accuracy in their analyses. We recognize that excluding these data points limits the model's applicability to less critical cases, and future research should focus on finding advanced techniques to handle motion and outliers to develop more generalized conclusions applicable to a broader patient population.

7) Group Data Per 1 °C Body Temperature Range: We segmented the data into temperature ranges of 1 degree Celsius

(°C) each to conduct specific temperature value analyses. This temperature-based segmentation created groups ranging from 33 °C to 33.9 °C up to 40 °C to 40.9 °C. By doing so, we performed detailed investigations into how patient's physiological responses vary across different temperature ranges, providing insights that could be valuable for clinical and research purposes.

8) Keep Single Observation Per Patient: The decision to preserve only one observation per patient per temperature interval was implemented to address the substantial variability in data volume per patient and ensure our dataset's integrity and balance. By reducing the data to a single observation per patient per interval, we aimed to give each patient an equal contribution to the final model, avoiding any skew caused by patients with disproportionately more data points. Our analysis selected the observation closest to the group's median value for each patient with multiple HR-BT observations within the same BT group (e.g., 36.0 to 36.9 °C). This approach aimed to minimize bias from extreme values or outliers while selecting a representative HR-BT pair for that temperature range. For patients with multiple BT values across different temperature groups (e.g., 36.0 to 36.9 °C, 37.0 to 37.9 °C, etc.), we preserved one HR-BT observation per BT group to ensure the relationship between HR and BT was captured across the range of temperatures experienced. This process resulted in approximately one HR-BT observation per patient per BT group, reducing redundancy and maintaining a manageable dataset. However, relying on the population-level median as a reference may introduce bias by favoring group-based patterns over individual-specific nuances, potentially overlooking the unique HR-BT relationship for each patient. A more tailored approach, such as using the patient's median HR within the BT range as a reference, could better reflect individual variability. While our method effectively balanced representation and dataset efficiency, future work should explore more advanced strategies, such as incorporating all available data points or employing patient-specific reference values, to enhance the model's robustness and generalizability.

After preprocessing, our dataset comprised 4462 pairs of HR-BT values represented in Table I. The values were organized based on BT values, where we used temperature grouping with 1-degree increments. This decision allows for a more granular representation of temperature variations and makes it easier for the ML algorithm to identify patterns in the data [18]. Moreover, the table provides a breakdown of patient counts according to the patient's age, offering additional insights into the age distribution within the dataset. The age groups include newborn (0-28 days), infant (29 days - 1 year), toddler (1-2 years), child (2-12 years) and teenager (12-18 years). This meticulous grouping strategy corresponds to the Food and Drug Administration (FDA) standard age categorization, with a more detailed grouping specifically for patients aged 1 to 2 years [24]. It also facilitates a more nuanced analysis, considering temperature variations and age-specific patterns in pediatric patients.

Our main goal was to accurately capture the true relationship between HR, BT, and age in stable states of pediatric patients admitted to the PICU. Data down-selection during preprocessing was necessary to ensure the dataset's quality, as data quality directly influences the ML model's ability to identify meaningful

TABLE I
NUMBER OF OBSERVATIONS AND PERCENTAGES FOR EACH BODY TEMPERATURE CATEGORY BY AGE GROUP FOLLOWING DATA PREPROCESSING

BT range (°C)	Newborn	Infant	Toddler	Child	Teenager	Median HR (bpm)	Total	Percentage
33-33.9	2	2	0	1	1	105	6	0.14%
34-34.9	8	3	0	4	2	118	17	0.39%
35-35.9	10	15	3	18	13	113	59	1.34%
36-36.9	110	335	139	640	391	114	1615	36.77%
37-37.9	136	442	201	752	420	119	1951	44.42%
38-38.9	48	168	53	243	117	131	629	14.32%
39-39.9	5	28	13	78	28	133	152	3.46%
40-40.9	0	3	2	21	7	132	33	0.75%
Subtotal	**319**	**996**	**411**	**1757**	**979**		**4462**	**100%**

patterns. Therefore, we removed instances that could introduce confounding factors, such as patient motion or extreme HR values, as these factors can introduce noise into the model, potentially leading to prediction inaccuracies. Previous studies have predominantly focused on the linear relationship between HR and BT; however, our observations suggest that this relationship is not clearly linear, with linear models underestimating or overestimating HR in specific BT and age ranges, as evidenced by the improved accuracy of our GBM model. Additionally, the number of data points per patient varied, which could lead to an imbalance in the model. To address this, we ensured that each patient contributed equally to the final model. We selected observations that were the most representative of stable states to achieve a balanced dataset. While this approach helped us focus on capturing the fundamental dynamics between HR, BT, and age, we acknowledge that it may also result in the loss of some information. Originally, our dataset included 10,135 HR-BT pair values, which were reduced to 4,462 observations for 4007 patients after applying our down-selection strategy. While this represents 56% data loss, this approach avoids overrepresenting patients with disproportionately large data points, ensuring that no single patient dominates the dataset. However, we recognize that this method may lead to a loss of variability, as only a subset of the available data is retained. While this reduces noise, it may limit the ability to capture finer details in individual patient variations, which we highlight as a limitation in our study. In future studies, preserving as much data as possible and employing advanced techniques to handle motion artifacts and outlier values will be crucial. This will help develop models with more generalized conclusions that apply to a broader range of clinical conditions.

C. Machine Learning

Following the preprocessing steps, the objective is to capture the intricate relationship between HR and BT in critically ill pediatric patients. We will validate this finding with our data by employing conventional linear regression models based on previous studies highlighting linear relationships. For this study, four popular algorithms were used for modeling the linear relationship [18]:

1) Conventional Linear Regression:

a) **Linear Regression (LR)** is used to determine the best linear function that fits a given set of input-output pairs [18]. This approach was also used by Kirschen, Singer, Thode Jr, & Singer, 2020 and Jensen & Brabrand, 2015 where they respectively observed that a 1 °C rise in BT would cause a 7 bpm and a 7.4 ± 0.9 bpm increase in HR among adults admitted to the emergency department [5], [15]. The limitation of this technique is only one independent variable can be used to predict the dependent variable, making this approach less accurate for prediction. [18].

b) **Multiple Linear Regression (MLR)** is also used to build a linear model to predict the dependent variable related to one or many independent variables, which makes this technique a better prediction model than the linear regression [15]. This technique was used by Heal et al., 2022, Jensen & Brabrand, 2015 and Broman, Vincent, Ronco, Hansson, & Bell, 2021 where they all concluded in a linear change between HR, BT and age [5], [12], [16].

c) **Polynomial Regression (PR)** with first degree, is a linear regression equation. The first-degree polynomial, often called simple LR, considers terms up to the first power, resulting in a model where the response variable is a linear function of the predictors. [18]

d) **Support Vector Machine (SVM)** can be used for both classification and regression problems. In this work, we employed the linear kernel, and the objective was to fit as many data points as possible to the hyperplane line inside the demarcation lines, known as the decision boundaries. This approach's strength is its excellent prediction accuracy and robustness against outliers. However, it does not attempt to minimize the errors between the true and predicted values, in contrast to other regression algorithms [18].

e) **Statistical modeling** approach by Davies & Maconochie, 2009 enabled them to predict HR at different quantiles, leading to the conclusion that "BT serves as an independent determinant of HR, resulting in an approximate increase of 10 bpm per degree centigrade" [6]. In our study, we used their established equation to determine its coefficients, then evaluated its effectiveness in predicting HR with our data [6]. Their statistical modeling is shown below:

$$\text{Expected HR value} = \text{BT} \cdot a + \text{Age} \cdot b + (\text{Age}^2) \cdot c + \text{constant} \quad (1)$$

where BT is the body temperature in °C and Age is the age in month [6].

2) Quantile Regression: Furthermore, we incorporated Quantile regression (QR) into our analysis for a more comprehensive exploration, combined with various traditional ML algorithms and DL models. This method allowed us to go beyond linear assumptions and enhanced our ability to uncover linear and non-linear patterns in our dataset. We also included interaction terms BT and age to enhance the accuracy of the predictive model for HR and provide a more comprehensive understanding of the dynamic.

QR is a statistical method used to estimate the response variable at various quantiles (e.g., 0.10, 0.25, 0.50, 0.75, 0.90). Its primary purpose is to analyze the impact of predictor variables (independent variable) on the output variable (dependent variable) across these quantiles. Unlike traditional regression methods, such as Least squares regression, which focus on estimating the mean, QR can offer insights into how the predictors influence various parts of the response distribution, making it valuable in situations where the relationship is not uniform across all quantiles and suitable for capturing a broader range of information about the response variable [25]. This method is typically utilized when the assumptions of linear regression are unmet. One of the advantages of QR lies in its robustness, as it is less sensitive to outliers compared to the least squares method. This robustness enhances the reliability of the technique in the presence of atypical data points. Additionally, QR offers flexibility by enabling the capture of the entire distribution of the outcome variables, providing a more comprehensive view of the relationship between variables. Moreover, it addresses the issues related to heteroscedasticity, where the variance of the dependent variable varies differently depending on the levels of the independent variable. This implies that the variance of residuals is not constant across all levels of the response variable [25]. Another benefit of using QR is that we may customize the kernel, which allows us to improve our capacity to identify and comprehend subtleties in the relationships between variables. Because of its adaptability, we continue to analyze three different kernels: conventional ML as a linear kernel, conventional ML as a non-linear kernel, and DL as a non-linear kernel.

a) Conventional ML as a linear kernel

    i) **Ordinary Least Squares (OLS)** is utilized to model and estimate the coefficients of a linear relationship between variables, particularly when the assumed relationship is linear. OLS aims to comprehend the correlation between variables and make predictions while minimizing the sum of all squares, also called the residual sum of squares (RSS), which is the difference between observed and predicted values. This method proves efficient in providing accurate parameter estimates when the assumption of linearity holds. It also offers interpretability where the coefficients present clear insights into the relationship between the variables [26].

b) Conventional ML as a non-linear kernel

    i) **Gradient Boosting Machine (GBM)** creates a predictive model in the form of a decision tree, where each tree in the ensemble corrects the errors of its predecessor, leading to an overall model with enhanced accuracy and robustness. The primary objective of GBM is to generate a predictive model capable of capturing complex relationships in the data. This versatile technique can be applied to both regression and classification problems. One of its key strengths lies in its iterative improvement process, where subsequent trees focus on addressing the residuals left by previous ones. Its iterative and sequential construction allows it to capture and model complex, non-linear interactions between variables. This flexibility makes GBM a valuable tool for tasks where the underlying relationships may not follow a simple linear structure. Some notable advantages of GBM include high predictive accuracy, the capability to handle both linear and non-linear relationships effectively, and robustness against overfitting. The model's ability to learn from its mistakes and continuously refine its predictions makes it a valuable tool for tasks requiring precision and adaptability in the face of complex data structures [27], [28], [29].

    ii) **Random Forest (RF)** constructs many decision trees, each built on a subset of the data during training. Consequently, the output represents these trees' mean or average prediction for regression problems. Each tree focuses on a specific subset, and the final prediction aggregates these individual tree predictions. The algorithm offers several benefits: flexibility to handle regression and classification problems, robustness against overfitting compared to decision trees, ability to highlight feature importance for better predictive performance, effectiveness in capturing complex relationships and non-linear patterns, and reduced sensitivity to outliers [30].

c) Neural network as a non-linear kernel

The decision to use these advanced methods was driven by the need to address non-linear associations and outliers within the dataset, which are challenging to model using traditional approaches. Quantile Regression was specifically chosen to provide a robust analysis of the data distribution across different percentiles, highlighting variations in HR that mean-based methods could obscure. In neural network models, we design a small-scale model with only three hidden layers and a small number of neurons. Additionally, we employed techniques like early stopping to mitigate overfitting and ensure that the models generalized well despite the limited data points.

    i) **Recurrent Neural Network (RNN)** is designed to discover patterns in sequential data, making them suitable for time series analysis and natural language processing. Their fundamental advantage lies in their ability to use their collection of previous inputs to make real-time decisions by doing the same task on each sequence element. This model's incapacity to provide background

knowledge over an extended period is one of its limitations [31].

ii) **Long Short-Term Memory (LSTM)**, an advanced form of RNN, addresses the issue of learning long-term dependencies. Its unique architecture incorporates four interacting layers containing chains of repeated modules. This model's memory cells enable it to store and retrieve information selectively over time. This makes LSTMs highly effective in tasks requiring understanding context over extended periods [31], [32].

iii) **Multilayer Perceptron (MLP)** is characterized by three layers (input, hidden, and output layers) of interconnected nodes (neurons), where each node in one layer is linked to every node in the next layer with associated weights. Also known as a supplement of feed forwards NN, the data move in one direction from the input, hidden to the output layer. The multiple hidden layers represent the computational engine of MLP. It facilitates feature learning, allowing the network to acquire hierarchical representations of data and capture intricate patterns and features due to their nonlinear characteristics. Moreover, The MLP's adaptability and effectiveness make it a popular choice for various ML applications, including regression and classification [33].

D. Performance Metrics

The following performance metrics were used to assess the effectiveness of ML and DL models. We used R-squared and Mean square errors to evaluate conventional linear regression. While QR models were evaluated using quantile loss. These metrics provided a comprehensive evaluation of the models' performance, offering insights into their ability to predict and capture the underlying patterns within the data. Each metric contributed to a nuanced understanding of the model's accuracy for the regression task [34].

- R-squared (R2), or the coefficient of determination, represents "the proportion of variance in the dependent variable (output) that is predictable from the independent variables (input)" [35]. The interpretation of R2 provides a quantitative measure of the quality fit. Potentially ranging from 0 to 1, an R2 value near 1 indicates that the model explains all of the dependent variable's variability, signifying a robust fit. In contrast, an R2 value nearing 0 indicates that the model fails to explain the variability in the dependent variable, implying a less effective fit [36]. The R2 equation is defined below:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \tilde{Y})^2} \tag{2}$$

where n is the number of observations, Y_i is the true value, $\hat{Y}_i$ is the predicted value and $\tilde{Y}$ is the mean of Y values. More precisely, the numerator is the sum of the squared residual (SSR), representing the difference between the actual and predicted values. The denominator is the total variations of the sum of

squares, representing the sum of the distance between the data and the mean all squared [36].

- Mean Square Error (MSE) measures the average squared difference between the true and the predicted values. Being only a positive value, the interpretation of the MSE value is as follows: a lower MSE signifies that the predicted values exhibit closer proximity to the actual values, reflecting an enhanced overall fit. This metric provides a standardized measure for evaluating regression models' precision and accuracy, facilitating straightforward comparisons between different models [37], [38]. The MSE equation is defined below :

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{3}$$

where n is the number of observations, Y_i is the true value and $\hat{Y}_i$ is the predicted value. [37]

- Quantile Loss (QL), also known as Pinball Loss, is a metric used in QR to calculate the performance of conditional quantiles, providing insight into how well a model captures the distribution of the data [39]. Unlike other ML algorithms, R2 is not a good metric to evaluate performance since it is not based on distribution assumptions [40]. A reminder that QR is a statistical method that aims to model the response variable's conditional distribution through different quantiles, allowing us to have information on the entire distribution, not just the mean. To interpret QL results, a low QL value means a superior performance from the model, suggesting that the predicted values closely align with the actual values at the designated quantile [41]. The QL equation is shown below:

$$QL_\tau(y, \hat{y}) = \max(\tau(y - \hat{y}), (\tau - 1)(y - \hat{y})) \tag{4}$$

where τ is the quantile level, y is the true value, $\hat{y}$ is the predicted value [41]. To compare the performance of quantile regression for each model, it is necessary to compute the total quantile loss, achieved by calculating the average quantile loss across all quantiles. This enables the performance analysis of all quantiles within a model [42].

The ultimate goal is to identify patterns and associations and develop a robust predictive model. The model aims to assist healthcare professionals in anticipating changes in HR based on variations in BT, age, and quantile of interest. Through the integration of advanced ML methodologies, this study seeks to propel the field of pediatric critical care forward, providing clinicians with actionable insights to tailor interventions and enhance overall patient care in the PICU.

III. RESULTS AND DISCUSSION

To ensure a reliable evaluation of our models, the `train_test_split` function in Scikit-learn [43] is a widely used technique to partition a dataset into training and testing subsets for machine learning tasks. It ensures a randomized split based on specified proportions (80% for training and 20% for testing) and shuffles the data by default to avoid biases caused by the original order. In combination with early stopping techniques, this strategy helps to monitor performance on a
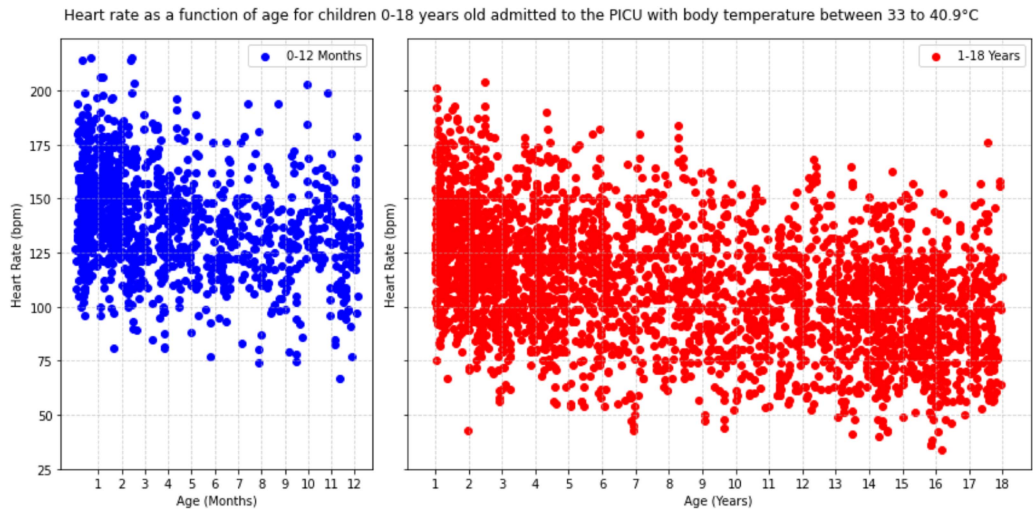
Fig. 2. Median heart rate as a function of age for children from 0-18 years old admitted to the PICU with body temperature between 33 to 40.9 °C.

TABLE II
HYPERPARAMETERS FOR CONVENTIONAL ML MODELS

Hyperparameters	GBM	RF
# Estimators	300	1000
Max depth	3	N/A
Learning rate	0.01	N/A
Min samples leaf	9	1
Min samples split	9	N/A
Random state	N/A	3

TABLE III
HYPERPARAMETERS FOR NEURAL NETWORK-BASED MODELS

Hyperparameters	RNN	LSTM	MLP
Hidden layers	3	3	3
Epochs	200	200	200
Batch size	32	32	32
Units (#neurons)	64	64	512
Optimizer	Adam	Adam	Adam
Learning rate	0.001	0.01	0.001

TABLE IV
MODEL PERFORMANCE FROM TRADITIONAL LINEAR MACHINE LEARNING
TECHNIQUES TO PREDICT HEART RATE FROM CHILDREN 0-18 YEARS OLD
WITH BODY TEMPERATURE BETWEEN 33 TO 40.9 °C

Conventional linear regression	R2	MSE
Linear regression	0.3145	621.1870
Multiple Linear Regression	0.3563	583.2989
SVM (kernel: linear)	0.3152	620.5312
Statistical model (quantile=0.5)	0.3576	581.2691

validation set, effectively preventing overfitting and improving the model's generalization capability.

For hyperparameter tuning, we employed a grid search technique, systematically testing different combinations of critical parameters, as shown in Tables II and III, to identify the optimal settings that maximize prediction accuracy. This method allowed us to fine-tune each model to achieve the best performance based on the specific characteristics of our dataset. We used early stopping criteria for deep learning models to avoid overfitting by halting training when the validation loss ceased to improve, thereby ensuring that the model did not learn noise from the training data.

Firstly, our initial findings from the research involve the generation of a graph showing HR concerning age for a temperature range presented in Fig. 2. Upon analysis, we observe a trend where HR decreases with the patient's age. This observation aligns with the findings presented in the article by the study [9]. The graphical representation provides valuable insights into the relationship between age and HR within distinct temperature

ranges, highlighting a potential correlation that needs further investigation. This initial exploration sets the foundation for a more in-depth analysis of age-related variations in HR across different BT intervals.

Expanding on these insights, we applied ML techniques to uncover subtle patterns and dependencies within our dataset. Each algorithm underwent meticulous hyperparameter tuning to optimize the model performance. Hyperparameters were set before the training process. Meticulous tuning suggests a careful and detailed examination of these hyperparameters to achieve the best possible performance from the model. It involves iteratively adjusting individual hyperparameters, training the model, and evaluating its performance until an optimal configuration is found. The goal is to enhance the model's effectiveness and accuracy. This meticulous approach to hyperparameter tuning is essential for achieving optimal results and maximizing the model's potential.

Since prior research has consistently underscored the linear association between HR and BT [5], [6], [10], [12], [14], [15], [16], we initially used conventional linear regression models. We started with LR to predict HR based on age for each BT range. Subsequently, we explored MLR and PR degree 1 to incorporate BT and age as independent variables. Finally, we experimented with SVM with a linear kernel and the statistical model. Examining the outcomes of our analyses, the performances of these traditional linear regression methods in capturing the linear relationship for BT ranging from 37 to 37.9 °C are summarized in Table IV. The R2 values for the linear models

ranged from 0.3145 to 0.3576. These values indicate that the models explain approximately 31.5% to 35.8% of the variability in HR based on BT. The relatively low R2 values suggest that linear models are limited in capturing the complex relationship between HR and BT, which may be influenced by non-linear interactions or additional physiological variables not considered in these models. Also, the MSE values range from 581.2691 to 621.1870, indicating the average squared difference between the predicted and actual HR values. Higher MSE values reflect more significant inaccuracies in the model's predictions. In this case, the linear models exhibit relatively high MSEs, suggesting that their predictive accuracy is limited, likely due to the over-simplified assumption of a linear relationship between HR and BT. Our study uses advanced ML techniques to identify subtle variations in HR that might not be detected using conventional linear models, thereby providing a more accurate understanding of the physiological interactions in critically ill children. Our work sets the foundation for more sophisticated analysis in this domain, challenging established views and encouraging further exploration into the complex relationship between HR and BT. From the comparative study, we can conclude that there are no obvious linear relations between HR and BT because of significantly small R2 values and a significant value of MSE loss. These values suggest that the linear models struggled to account for a substantial portion of the variability in HR based on BT and age. Based on these findings, there may be a potential need for more advanced methodologies to effectively capture the intricacies of the relationship between HR and BT.

Our experimental findings suggest that linear models might not be able to adequately capture the complexity of the relationship observed in critically ill patients. This assertion is supported by the studies [44], [45], highlighting the non-linear correlation between HR and BT. Specifically, in Momo et al. (2023) article [44], the authors critique the work of Heal et al. (2022) [12], who utilized linear regression models to demonstrate a significant increase in HR of 12.3 bpm for every 1 °C rise in BT among emergency department patients. Additionally, Momo et al. (2023) underscore the significance of findings from other studies employing non-linear models. These studies reveal a distinct, curve-like increase in HR within the temperature range of 37 °C to 38 °C, highlighting the intricate relationship between HR and BT [44].

The authors respond to the previously mentioned critiques in the article by Heal et al. (2023). They acknowledge that employing linear models may not be the most accurate approach and suggest using polynomial regression. They also note that contrary to earlier research suggesting a general 1 °C increase correlating with a 10 bpm rise, their findings indicate a 12.3 bpm increase, emphasizing potential variations within specific temperature ranges. This distinction gains particular significance when considering the application of such predictive tools to individual patients in a clinical setting [45].

Since traditional linear models do not adequately capture the complex and potentially non-linear relationship, we conducted a comprehensive evaluation using QR performed with ML and DL as a linear and non-linear kernel. Through this, we aim to enhance the depth and precision of our analysis. The advantage
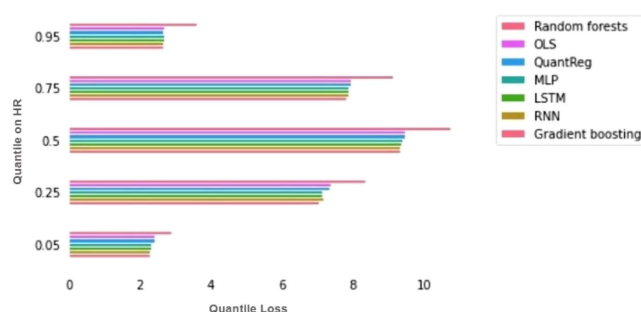

Fig. 3. Comparison of quantile loss per quantile for each technique for a body temperature range between 33 to 40.9 °C from critically Ill patients 0-18 years old from 1 experiment.

TABLE V
MEAN TOTAL QUANTILE LOSS AND SD FROM 5 EXPERIMENTS FOR DIFFERENT METHODS PERFORMED WITH QR WITH A TEMPERATURE RANGE BETWEEN 33 TO 40.9 °C FROM CRITICALLY ILL PATIENTS 0-18 YEARS OLD

Model	Mean total quantile loss ± SD
GBM	6.5069 ± 5.2507e-05
RNN	6.5633 ± 0.0078
LSTM	6.5990 ± 0.0144
MLP	6.6148 ± 0.0236
OLS	6.6867 ± 0.0
QR	6.6903 ± 0.0
RF	7.6584 ± 0.0

of QR lies in its ability to handle data with varying distributions and to provide a more nuanced understanding of the conditional distribution of the response variable across different quantiles [25]. This is particularly relevant when the relationship between variables may not be constant across the entire distribution. To evaluate the models with QR, it is unsuitable to use R2, as mentioned in the literature [40]. Instead, it is essential to investigate the quantile loss (QL) for each quantile as the performance metric for QR. The results presented in Fig. 3 reveal that the lowest QL value is seen at each quantile in QR performed with GBM, showing that the underlying relationship is effectively captured.

To assess the performance of each model with QR, the performance metric is the total quantile loss, representing the overall performance across all quantiles within a model. This metric is obtained by calculating the average QL across all quantiles. The data presented in Table V detail the average total quantile loss and the standard deviation (SD) for each predictive model derived from the outcomes of five distinct experiments. The best model is once again QR performed with GBM. The model demonstrates superior efficacy by achieving the lowest total quantile loss value. The exceptional performance of this model suggests its effectiveness in handling the complexities in the data, contributing to enhanced predictive capabilities. This advanced method outperforms conventional approaches with linear kernels, such as OLS, RF, and standalone QR. Specifically, the higher total quantile loss result for RF implies potential limitations in its effectiveness within this context. These results underscore the significance of employing advanced ML and DL
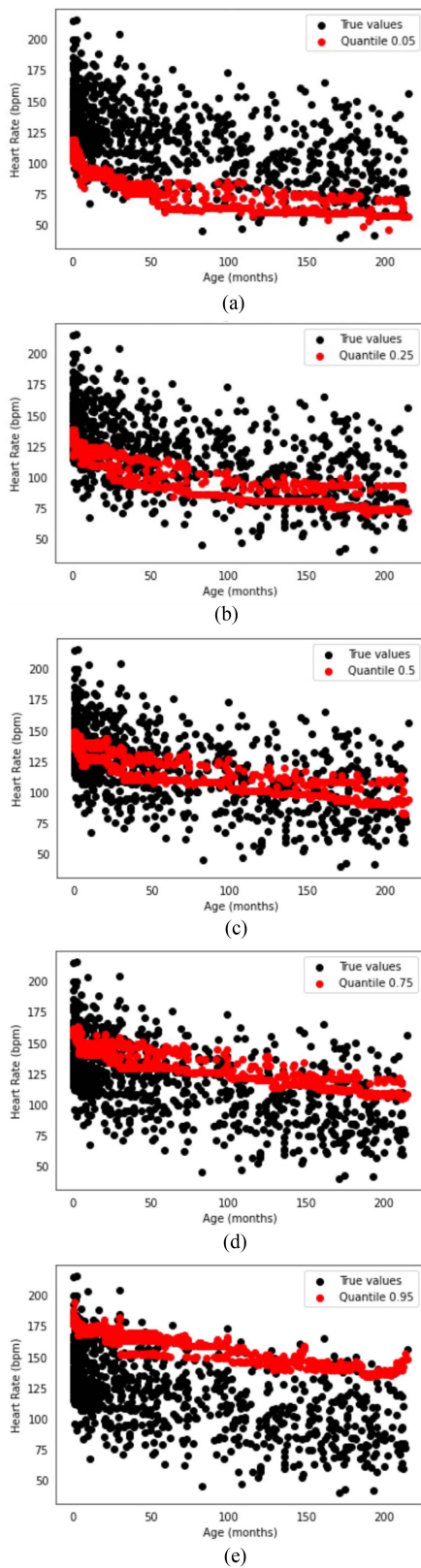
Fig. 4. QR performed with GBM model shows a non-linear relationship between HR, BT, and age for children 0-18 years old admitted to PICU with a temperature range of 34 to 40.9 °C. True values are shown in black, while predictions are shown in red. The analysis is performed at different quantiles, where each row represents a specific quantile, from 0.05 to 0.95. Subfigure (a) illustrates model predictions at quantile 0.05, (b) at quantile 0.25, (c) at quantile 0.50, (d) at quantile 0.75, and (e) at quantile 0.95.
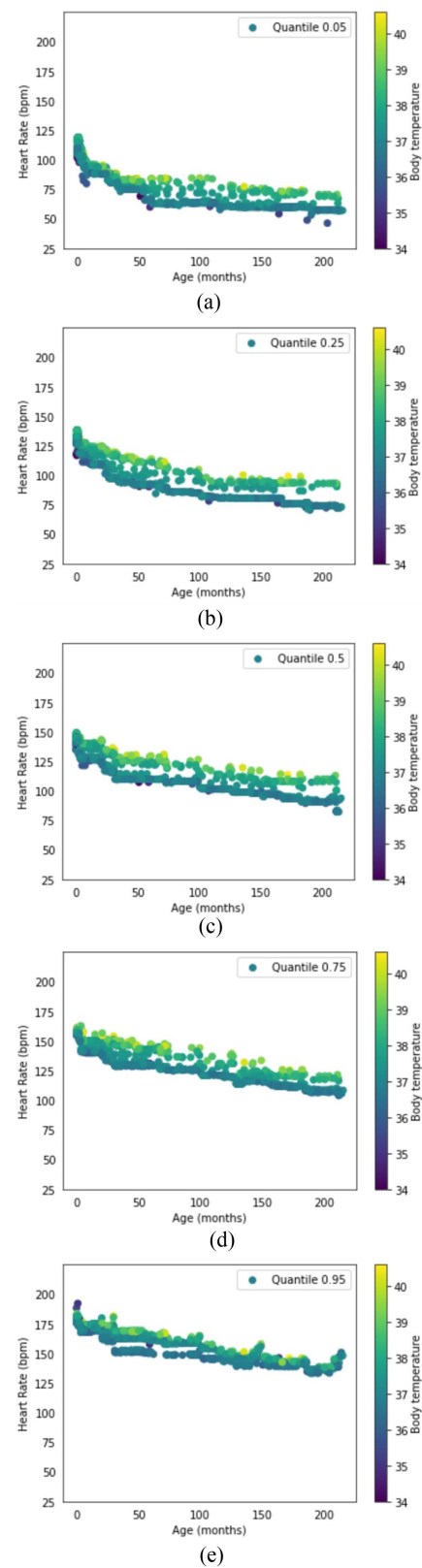
Fig. 5. Exploration of the HR model predictions influenced by age along the x-axis and BT through variation of 34 to 40.9 °C in the color scale for children 0-18 years old admitted to PICU. The analysis is performed at different quantiles, where each row represents a specific quantile, from 0.05 to 0.95. Subfigure (a) illustrates model predictions at quantile 0.05, (b) at quantile 0.25, (c) at quantile 0.50, (d) at quantile 0.75, and (e) at quantile 0.95.
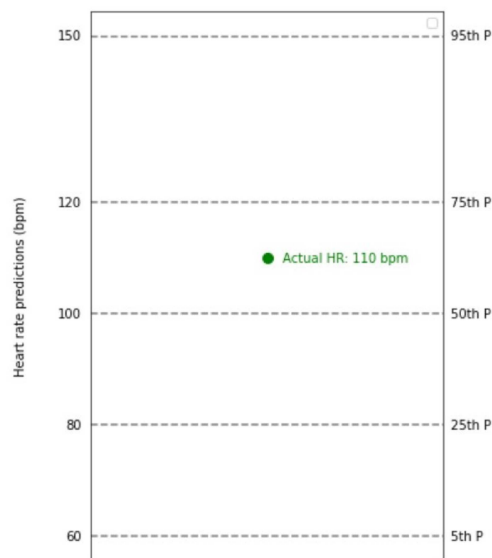
Fig. 6. HR Predictions at different percentiles (5th, 25th, 50th, 75th, 95th) for a given patient's age and BT will be integrated into the CDSS system (a) User interface featuring three key inputs - current HR, current BT, and patient's age - for HR predictions (b) Example of case scenario with a normal HR is indicated by a green dot, which falls within the normal range defined by the 5th and 95th percentiles (c) Another example of a case scenario shows an abnormal HR, marked by a red dot located outside the normal range.

techniques for capturing nuanced patterns in critical care scenarios, providing valuable insights into the relationships between physiological parameters.

Besides the RF approach, the other approaches achieved similar results to those achieved by GBM with QR (considered the best). For example, results achieved by OLS are 6.6867 ± 0.0 vs. $6.5069 \pm 5.2e\text{-}05$ (from GBM), with only 2.8% down performance (in mean total quantile loss measure) GBM offers advantages over OLS, making it particularly suitable for clinical applications. First, GBM can capture complex, non-linear relationships between features, often crucial in clinical data where interactions among variables—such as age, body temperature, and heart rate—tend to be highly non-linear. Conversely, OLS assumes a linear relationship, which may oversimplify these interactions and miss subtle but clinically significant patterns that could influence decision-making. Additionally, GBM provides robustness and flexibility by iteratively boosting predictions to minimize errors from prior models, making it more resilient to various data distributions and outliers [46]. This iterative adjustment allows GBM to handle the variability inherent in clinical data more effectively than OLS, which can be sensitive to

outliers and may require extensive preprocessing to yield reliable results.

Fig. 4 presents the predictive accomplishment of the best-performing model—QR using GBM kernel—highlighting its capacity to capture the complex non-linear relationships at varying quantiles (0.05, 0.25, 0.50, 0.75, 0.95). From top to bottom, the figure sequentially exhibits the model's proficiency in forecasting HR at multiple quantiles, from the 5th to the 95th percentile. This projection is based on patient age, ranging from infancy to nearly 18 years (0-200 months), and BT between $34\,^{\circ}\text{C}$ to $40.9\,^{\circ}\text{C}$. The x-axis denotes the patient's age in months, and the y-axis records HR in beats per minute (bpm). The sub-figure presents the actual HR data as black dots, with the model's predicted values shown in red, facilitating a straightforward comparison of the observed and predicted HR. This allows for a direct appraisal of the model's precision in replicating the true HR values.

Furthermore, in response to the nature of the wide BT range (from $34\,^{\circ}\text{C}$ to $40.9\,^{\circ}\text{C}$), Fig. 5 is introduced to deliver enhanced detail concerning the model's predictions concerning BT variations. Specifically, the experimental results for HR prediction for

a specific quantile (0.05, 0.25, 0.50, 0.75, 0.95) are presented, respectively, in each row. A color gradient alongside each scatter plot translates the spectrum of colors to specific BT values, offering an instant visual guide. The scatter points, color-coded to represent varying BTs, effectively differentiate the HR predictions according to BT variations, thereby enriching the visual representation and enhancing the interpretability of the model's capacity to account for temperature-dependent HR variations. Consequently, analyzing the scatter plots across all the quantiles reveals a pattern: as BT increases, so does the predicted HR, while an inverse relationship is observed with increasing age, where the predicted HR generally declines. This visualization is not merely a reflection of the model's prediction accuracy; it also communicates critical insights for clinical applications, particularly as it highlights the expected HR values within the commonly encountered 5th to 95th percentile range in a clinical setting.

To enable clinicians at CHUSJ to utilize our QR prediction model implemented with GBM, we have developed an example of a user interface (UI) integrated into the hospital's CDSS, as shown in Fig. 6. This UI processes three input parameters: the current HR, BT, and the patient's age, then provides HR predictions across various percentiles (5th to 95th) specifically tailored to individual patients. Importantly, these results are derived from the inference process of our trained model, which defines a normal HR range within the 5th to 95th percentile. Within this range, a green dot indicates that the patient's HR falls within the expected parameters, signaling normalcy. Conversely, a red dot is an immediate visual indicator of deviation from the norm, prompting timely attention and intervention. Through this interface, the CDSS can assess in real-time whether a patient's HR is within the predicted safe range or requires further evaluation, aiding caregivers in their clinical decision-making process.

This UI will be integrated into a personalized dashboard, for which we designed a prototype to optimize the representation of clinical data collected from existing sources (e.g., EHR, clinical systems, and devices) via a structure that supports the integration of a home-developed CDSS in the PICU [47]. The personalized dashboard was based on analyzing end-user needs and their clinical workflow. It is a structure with three levels of abstraction—unit level, patient level, and system level—to optimize clinical data representation and display for efficient patient assessment and to provide a flexible platform to host the internally developed CDSS. However, it should be noted that integrating this tool in the PICU of CHUSJ does not require caregivers to enter patient information manually. The data representation for the CDSS consists of three levels: unit, patient, and system levels. The unit level allows visualization of all admitted patients, while the patient level enables assessment of the patient's status and progress [47]. Lastly, the system level presents many "groups of indicators related to human body systems and provides access to decision support tools developed for specific issues involving these systems [47]". Therefore, the patient's information is already integrated into the system automatically. Consequently, our predictive model will be generated and displayed at the system level of the CDSS.

Our model's limitation is that no HR prediction is provided if the inputs fall outside the ranges covered by our dataset. This

underscores the importance of ensuring that the input parameters remain within the dataset's scope to obtain accurate predictions. Furthermore, clinicians must evaluate the predictions' interpretation and validation. This highlights the critical role of clinical expertise in assessing and confirming the accuracy and clinical relevance of the predicted HR values in this population.

IV. CONCLUSION

In summary, this study represents a comprehensive and innovative approach to understanding the relationship between HR, BT, and age in children within the PICU. By employing a meticulous multi-stage data preprocessing strategy, the research aimed to reveal complex patterns that conventional models might overlook.

The findings of this study align with prior research by confirming the expected downtrend in HR with increasing patient age. However, the nuanced approach of grouping data by temperature rather than age allowed for a more granular exploration. Notably, the performance evaluation of various ML algorithms yielded insightful results. Unlike earlier studies that suggested linear correlations, conventional linear regression demonstrated limited effectiveness in accurately capturing the relationship within the data. In contrast, QR implemented with advanced ML techniques, such as the GBM model, exhibited superior performance by adjusting more accurately to variations in HR across a broad BT range from 33 to 40.9 °C, particularly in regions where linear models tended to underestimate or overestimate HR. This highlights the potential limitations of assuming a strictly linear association in such datasets. Furthermore, the HR model predictions clearly show a downward HR trend with age and an upward trend with BT between the 5th and 95th percentiles. Based on that model, we created a simple user interface for caregivers. Based on age and BT, they can quickly determine in real-time whether a patient's HR falls within the normal range or not.

V. FUTURE WORKS

We recognize the importance of evaluating the generalizability of our model on external datasets. In this study, we initially developed our model from a specific dataset. However, to enhance the model's applicability, we propose future validation on separate datasets from other PICUs. Additionally, utilizing datasets with multiple observations per patient would allow us to further assess the model's robustness and adaptability. By testing the model across different patient populations and clinical settings, we aim to ensure its performance extends beyond the initial dataset used in this study.

Additionally, the exact nature and extent of the non-linear relationship require more detailed quantitative analysis. However, qualitative insights indicate that linear models often underestimate HR at lower BT levels and overestimate it at higher BT levels, especially among younger age groups. This limitation underscores the need for a more comprehensive understanding of the interaction between HR, BT, and age. Future research should prioritize rigorous quantification of this nonlinearity and investigate the specific conditions where linear

models diverge from observed trends to enhance predictive accuracy.

Furthermore, selecting the observation closest to the population-level median HR within each BT group helped reduce data redundancy and maintain a manageable dataset. However, this approach may introduce bias by prioritizing group-based patterns over capturing individual-specific nuances, potentially overlooking unique HR-BT relationships for each patient. Future work should explore alternative strategies, such as using patient-specific medians as references or incorporating all available data points through advanced modeling techniques, to capture individual variability better and improve the model's robustness and generalizability across diverse datasets.

Looking forward, the paper suggests promising future directions and areas for improvement. One intriguing prospect involves observing a cohort of subjects upon whom clinicians have implemented interventions guided by the predicted HR model. These observations could involve observing trends and assessing the long-term effects of interventions. Extracting insights from the resulting conditions and outcomes would be interesting in refining future studies. Moreover, if the HR falls outside the 5th and 95th percentiles, the time spent in this range may be associated with bad outcomes. It would be valuable to investigate whether there are treatments that could reduce the duration of time spent in these extremes to enhance patient outcomes. Furthermore, while most of our patients are calm, we suggest a subgroup analysis to compare the results for agitated versus non-agitated patients to ensure the relationship remains consistent. Another direction for future research is to explore the interplay of variables, focusing on gender-based differences, considering that HR tends to be higher in women than men. Also, we suggest enhancing the predictive UI tool by integrating additional parameters such as respiratory rate, blood pressure, and oxygen saturation. Including these variables will help improve the tool's predictive accuracy and clinical utility, making it more comprehensive in addressing the complex needs of critically ill pediatric patients. Finally, incorporating other neural network architectures, such as Convolutional Neural Networks (CNN) or Transformers, could offer valuable insights.

ACKNOWLEDGMENT

REFERENCES

[1] T. M. Slusher et al., "Pediatric critical care in resource-limited settings—Overview and lessons learned," *Front. Pediatrics*, vol. 6, 2018, Art. no. 49.
[2] J. A. Heneghan et al., "Characteristics and outcomes of critical illness in children with feeding and respiratory technology dependence," *Pediatr. Crit. Care Med.: J. Soc. Crit. Care Med. World Federation Pediatr. Intensive Crit. Care Societies*, vol. 20, no. 5, pp. 417–425, 2019.
[3] "CDSS grouping in AI acute care for the child," Accessed: Apr. 16, 2024. [Online]. Available: https://chusj-sip-ia.ca/home
[4] I. J. Brekke et al., "The value of vital sign trends in predicting and monitoring clinical deterioration: A systematic review," *PLoS One*, vol. 14, no. 1, 2019, Art. no. e0210875.
[5] M. M. Jensen and M. Brabrand, "The relationship between body temperature, heart rate and respiratory rate in acute patients at admission to a medical care unit," *Scand. J. Trauma, Resuscitation Emerg. Med.*, vol. 23, no. 1, 2015, Art. no. A12.
[6] P. Davies and I. Maconochie, "The relationship between body temperature, heart rate and respiratory rate in children," *Emerg. Med. J.: EMJ*, vol. 26, no. 9, pp. 641–643, 2009.
[7] S. Liu, J. Yao, and M. Motani, "Early prediction of vital signs using generative boosting via LSTM networks," in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, 2019, pp. 437–444.
[8] B. W. Nelson et al., "Guidelines for wrist-worn consumer wearable assessment of heart rate in biobehavioral research," *NPJ Digit. Med.*, vol. 3, no. 1, 2020, Art. no. 90.
[9] S. Fleming et al., "Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: A systematic review of observational studies," *Lancet*, vol. 377, no. 9770, pp. 1011–1018, 2011.
[10] C. Daymont, C. P. Bonafide, and P. W. Brady, "Heart rates in hospitalized children by age and body temperature," *Pediatrics*, vol. 135, no. 5, pp. e1173–e1181, 2015.
[11] S. Sidhu and J. E. Marine, "Evaluating and managing bradycardia," *Trends Cardiovasc. Med.*, vol. 30, no. 5, pp. 265–272, 2020.
[12] C. Heal et al., "The association between temperature, heart rate, and respiratory rate in children aged under 16 years attending urgent and emergency care settings," *Eur. J. Emerg. Med.*, vol. 29, no. 6, pp. 413–416, 2022.
[13] H. K. Walker, W. D. Hall, and J. W. Hurst, "Clinical methods: The history, physical, and laboratory examinations," Butterworths, 1990.
[14] M. Thompson et al., "Deriving temperature and age appropriate heart rate centiles for children with acute infections," *Arch. Dis. Childhood*, vol. 94, no. 5, pp. 361–365, 2009.
[15] G. W. Kirschen et al., "Relationship between body temperature and heart rate in adults and children: A local and national study," *Amer. J. Emerg. Med.*, vol. 38, no. 5, pp. 929–933, 2020.
[16] M. E. Broman et al., "The relationship between heart rate and body temperature in critically ill patients," *Crit. Care Med.*, vol. 49, no. 3, pp. e327–e331, 2021.
[17] D. Brossier et al., "Creating a high-frequency electronic database in the PICU: The perpetual patient," *Pediatr. Crit. care Med.*, vol. 19, no. 4, pp. e189–e198, 2018.
[18] D. Kinaneva et al., "Machine learning algorithms for regression analysis and predictions of numerical data," in *Proc. 3rd Int. Congr. Hum.-Comput. Interact., Optim. Robot. Appl.*, 2021, pp. 1–6.
[19] L. D. Cloedt et al., "The impact of implementing a "pain, agitation, and delirium bundle" in a pediatric intensive care unit: Improved delirium diagnosis," *J. Pediatr. Intensive Care*, vol. 11, no. 03, pp. 233–239, 2021.
[20] A. Amigoni et al., "Recommendations for analgesia and sedation in critically ill children admitted to intensive care unit," *J. Anesthesia, Analg. Crit. Care*, vol. 2, no. 1, 2022, Art. no. 9.
[21] P. Dolibog et al., "Comparative analysis of human body temperatures measured with noncontact and contact thermometers," *Healthcare*, vol. 10, no. 2, 2022, Art. no. 331.
[22] M. Meyer, M. Rambod, and M. LeWinter, "Pharmacological heart rate lowering in patients with a preserved ejection fraction—Review of a failing concept," *Heart Failure Rev.*, vol. 23, pp. 499–506, 2018.
[23] G. S. King et al., "Antiarrhythmic medications," StatPearls Publishing, 2024.
[24] "Pediatric medical devices," Accessed: Apr. 16, 2024. [Online]. Available: https://www.fda.gov/medical-devices/products-and-medical-procedures/pediatric-medical-devices
[25] V. K. Patidar et al., "Quantile regression comprehensive in machine learning: A review," in *Proc. IEEE Int. Students' Conf. Elect., Electron. Comput. Sci.*, 2023, pp. 1–6.
[26] G. D. Hutcheson and L. A. M. Moutinho, "The SAGE dictionary of quantitative management research," in *SAGE Dictionary Quantitative Manage. Res.*, pp. 1–344, 2011.
[27] S. Fafalios, P. Charonyktakis, and I. Tsamardinos, "Gradient boosting trees," *Gnosis Data Anal. PC*, vol. 1, pp. 1–3, 2020.
[28] N. Aziz et al., "A study on gradient boosting algorithms for development of AI monitoring and prediction systems," in *Proc. Int. Conf. Comput. Intell.*, 2020, pp. 11–16.

[29] B. Boehmke and B. M. Greenwell, *Hands-on Machine Learning With R.* Boca Raton, FL, USA: CRC Press, 2019.

[30] A. Cutler, D. R. Cutler, and J. R. Stevens, "Random forests," in *Ensemble Machine Learning: Methods and Applications*. Berlin, Germany: Springer, 2012, pp. 157–175.

[31] S. Sivamohan, S. Sridhar, and S. Krishnaveni, "An effective recurrent neural network (RNN) based intrusion detection via bi-directional long short-term memory," in *Proc. Int. Conf. Intell. Technol.*, 2021, pp. 1–5.

[32] B. N. Saha and A. Senapati, "Long short term memory (LSTM) based deep learning for sentiment analysis of English and Spanish data," in *Proc. Int. Conf. Comput. Perform. Eval.*, 2020, pp. 442–446.

[33] D. Molina Estren, A. K. De la Hoz Manotas, and F. M. Palechor, "Classification and features selection method for obesity level prediction," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 11, pp. 2525–2536, 2021.

[34] G. Varoquaux and O. Colliot, "Evaluating machine learning models and their diagnostic value," in *Machine Learning For Brain Disorders*. Berlin, Germany: Springer, 2023, pp. 601–630.

[35] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, 2021, Art. no. e623.

[36] D. M. Keer, H. Lohiya, and S. Chouhan, "Goodness of fit for linear regression using R-squared and adjusted rsquared," *Int. J. Res. Pub. Rev.*, vol. 4, no. 3, pp. 2431–2439, 2023.

[37] K. Tyagi et al., "Regression analysis," in *Proc. Artif. Intell. Mach. Learn. EDGE Comput.*, 2022, pp. 53–63.

[38] C. Rane et al., "Optimal input gain: All you need to supercharge a feed-forward neural network," 2023, *arXiv:2303.17732*.

[39] I. Steinwart and A. Christmann, "Estimating conditional quantiles with the help of the pinball loss," *Bernoulli*, vol. 17, pp. 211–225, 2011.

[40] J. A. Machado and J. Silva, "Quantile regression and heteroskedasticity," Dept. of Econ., Univ. of Essex, Colchester, England, U.K., 2013. [Online]. Available: https://jmcss.som.surrey.ac.uk/JM_JSS.pdf

[41] T. Narayan et al., "Expected pinball loss for quantile regression and inverse CDF estimation," *Trans. Mach. Learn. Res.*, vol. 2024, Accessed: Feb. 18, 2024. [Online]. Available: https://openreview.net/forum?id=Eg8Rnb0Hdd

[42] A. Alutaibi and S. Ganti, "Network traffic prediction using quantile regression with linear, tree, and deep learning models," in *Proc. IEEE 45th Conf. Local Comput. Netw.*, 2020, pp. 421–424.

[43] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[44] O. M. Momo and Y. Morikawa, "Measuring the association between temperature and heart rate in children: Be aware of the uncertainties," *Eur. J. Emerg. Med.*, vol. 30, no. 2, pp. 147–148, 2023.

[45] C. Heal et al., "Response to 'measuring the association between temperature and heart rate in children: Be aware of the uncertainties'," *Eur. J. Emerg. Med.*, vol. 30, no. 2, pp. 148–148, 2023.

[46] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, pp. 1189–1232, 2001.

[47] N. Yakob et al., "Data representation structure to support clinical decision-making in the pediatric intensive care unit: Interview study and preliminary decision support interface design," *JMIR Formative Res.*, vol. 8, 2024, Art. no. e49497.