Physiological Deterioration Scoring and Mortality Risk Stratification: Machine Learning Models for Early Outcome Prediction in Critical Care Settings

Dr. David Kim, Associate Professor of Cybersecurity, Kookmin University, South Korea

1. Introduction

In recent years, there has been a growing body of work on applications of machine learning models in both the critical care unit and emergency department. It is well established that AI can provide insights beyond traditional mortality prediction tools, such as illness severity stratification, estimation of the impact of acute changes, early warning of sepsis and delirium, and end-of-life care strain, as well as novel ways of understanding the determinants of length of stay in an ICU setting. In particular, interest has focused on predicting patient outcomes to support medical decisions and to enhance the timeliness of patient management. Critical care is expensive to provide; therefore, appropriate selection criteria are required to optimize resources through select surrogate markers. In addition, timely prediction of recovery from critical illness can impact rehabilitation options. Additionally, critical care syndromes are heterogeneous, and patients require patient-centered care instead of guidelines to facilitate personalized and stratified planning for care. Early identification of patient subgroups can enable personalized treatment strategy planning. Timely risk prediction is also imperative in the context of organ dysfunction and critical illness. The purpose of our current review is to primarily highlight studies that aimed at predicting patient outcomes of critical care patients using AI. We will discuss studies that used AI for the prediction of short-term and long-term mortality; acute kidney injury; sepsis; septic shock; adult respiratory distress syndrome; multiple organ failure; early neurologic prognosis; and progression to microcirculatory shock, and we will present relevant pathophysiologic insights gained.

1.1. Background and Significance

One of the hallmarks of successful healthcare delivery is timely and precisely targeted patient care. As patient care increases in complexity and reflects the critical nature of healthcare delivery, strategic solutions are needed for the effective integration of personalized, high-value treatment planning. While advances in molecular diagnostics and image analysis enable the stratification of severe patient cases, the use of machine learning models can complement this approach and assess the most likely outcome in real time. AI has been used to integrate clinical variables to predict outcomes such as mortality in septic patients, shock in suspected sepsis patients in the emergency department, empyema in patients with complicated pneumonia, TB resistance, culture-negative peritoneal TB, flu-related admission in critically ill patients, and multiple other outcomes in critical care.

The integration of machine learning models into existing clinical parameter prediction research offers tremendous potential for the future development of personalized interventions in the critical care setting. However, historically, the use of machine learning algorithms in this manner raises many challenges, including the use of dynamic or personalized parameters estimated in rapid timeframes. The focus on personalized medicine is still in its early stages, which means that despite dynamic data acquisition and continuous monitoring of trends in variables at increasingly frequent time points, special consideration needs to be given to personalized management strategies. Many studies have recently integrated the use of traditional clinical parameters combined with AI approaches in big data studies to enhance the predictivity of various patient outcomes in the intensive care unit. The recent failure of multiple physical parameters to stratify patients in the ICU or at-risk ED makes nonacademic translation of these models difficult. As such, enclosed within exists a personalization granularity for every clinical parameter that can be monitored most effectively, and intelligent models have been shown to enhance infection prediction and stratification of ICU outcomes when combined.

The integration of these models into clinical decision support systems can be used to inform treatment strategies in a timely, cost-effective manner to optimize patient outcomes. However, to date, no study has thoroughly examined how AI and machine learning approaches can be integrated into the decision-making of complex personalization strategies in the acute care setting, the evaluation of these interventions, and finally what has been specifically used to enhance these strategies overall.

1.2. Objectives of the Study

The following is the objective of the study: (i) Machine learning models for patient outcome prediction in critical care: In this study, our primary aim is to evaluate patient outcome prediction through more than ten machine learning models using various imputation techniques and ensemble-based strategies. (ii) AI approaches to assess the risk of critical care patients: Together with intensivists, we seek to utilize explainable AI models to understand how key variables (comorbidities, admission concerns, vital signs) are included in predicting patient outcomes. We will specifically target morbidity-hour scores, a mechanism for assessing patient outcomes that could indicate duration on life support during involvement in research studies. (iii) Exploring personalized treatment strategies using AI in critically unwell patients addicted to alcohol or drugs: Critically unwell patients with either alcohol use disorder, acute methamphetamine intoxication, or chronic methamphetamine use disorder often require organ support during their admission to the hospital. Machine learning can enable the rapid assessment of contextual information, including clinical, comorbidity, and demographic factors, to target the severity of patient scoring systems. Understanding the severity of your patient population benefits benchmarking. Furthermore, understanding the risk for this particular population may alter the threshold for intensive care treatment as this drug- and alcohol-related population is often undertreated or declined admission.

In this work, we are hoping to answer the following questions: (1) Do machine learning models outperform traditional multivariable regression models in predicting patient outcomes? (2) Do machine learning models require more complex imputation techniques than traditional modeling? (3) Do AI-enhanced prognostic scores outperform established prognostic scores? (4) How important are the following variables on hospital mortality prediction in adult critical care patients: the patient's illness, the active comorbidities, and the number of health care issues present on the day the patient presents to the hospital?

2. Fundamentals of Critical Care

Critical care is a medical discipline that provides care for patients suffering from life-threatening illnesses or injuries. This discipline is characterized by the provision of life support and kindness for patients who are unable to function on their own; therefore, critical care does not only focus on how to cure patients but also on giving compassion and improving patients' quality of life. The medical team in critical care is arranged in a multi-disciplinary manner so that all needs (medical-physical, psychological-mental, social, and spiritual) of the patient can be identified and fulfilled well and continuously as long as the patient is in the unit. Today, the rapid increase in the aging population has led to an expansion in the demand for critical care, making it difficult for critical care services to meet the growing patient population. Critical care is also becoming more complex, and the patient populations it cares for are becoming sicker and have a higher level of comorbidity. In this context of critical care delivery, one of the main principles is the early comprehensive advanced monitoring of the patients to diagnose and, furthermore, to allow timely escalation of therapy and intervention.

Limited resources in the ICU need to be allocated to those who are likely to do well, and those whose prognosis is such that continuing with intensive treatment is futile should have their care limited. One of the difficulties in critical care is identifying to whom the request for additional support is necessary. This is because patient wishes or prognosis are often unclear upon hospitalization, although some people would want everything possible to be done for themselves. Ethical considerations inevitably limit the ability to provide critical care for all. For a patient with diseases such as chronic obstructive pulmonary disease with severe carbon dioxide retention, mechanical ventilation does not seem proportionate to the need for the recovery expectations of such patients. The purpose of the limitation of health services is to recommend providing treatment that prioritizes value and to do so according to each patient's situation. Critical care should continue to respond to appropriate social, economic, philosophical, religious, and ethical dilemmas and conflicts while still carrying out its humanitarian and charitable obligations. Therefore, early recognition of patients' illnesses is absolutely necessary in critical care. Such patients usually have a severe clinical condition with shock and impaired respiratory function, requiring rapid treatment to avoid fatal outcomes.

2.1. Definition and Scope

Critical care generally refers to the intense and specialized level of care provided to patients who are critically ill, have life-threatening conditions, and require close monitoring and attention from healthcare professionals to ensure their survival. There is a great diversity of medical illnesses that can lead the patient into critical care, and these medical illnesses can affect any of the organs, including those of the cardiovascular, central nervous system, genitourinary, gastroenterology, endocrinology, infectious diseases, poisoning, manifested by systemic illness, or a number of different combinations of organ systems. These organ system-related problems can occur secondary to medical and surgical conditions and are part of a holistic and inclusive practice of medicine in the twenty-first century. These observations guide the development of critical care facilities within general or specialist hospitals' ICUs, high dependency units, resuscitation rooms in the emergency departments, and operating rooms where such types of illness may present. The intensive care of patients has to be multifaceted in approach, requiring a combination of the skills of many healthcare professionals as specialists or generalists, and a critical care-trained nursing staff and advanced respiratory therapists.

The assessment and care of the patients within the above environment, therefore, need to be systematic, incorporating both disease-related and demographic patient factors to facilitate the best possible and appropriate integrated care in reducing morbidity and optimizing the survival of patients presenting with physiological abnormalities. The physiological derangements in these individuals are complex and unpredictable but usually remain within defined parameters, which can sometimes manifest as a degree of physiological decompensation and is, therefore, an ideal setting for developing and integrating specific technologies such as machine learning that may attempt prediction and patient management, particularly patient-tailored treatments. So, for the purpose of this material, our focus will be on machine learning approaches in critically ill patients and not in the general population.

2.2. Key Challenges in Critical Care

In critical care environments, clinicians are faced with a multitude of complex challenges that affect the delivery of care in the most effective and efficient way. These challenges are primarily on the clinical level where patients present with disease or injury involving one or multiple organ systems. Conditions are further complicated by high-acuity patients who may have severe physiological derangements requiring prompt attention or intervention to prevent catastrophic organ failure. Yet, in many cases, the use of highly limited medical resources is crucial for healthcare professionals. In addition, clinical data are typically multimodal, high-dimensional, and complexly interrelated, reflecting disruption and derangement at the molecular, cellular, and systemic levels. Associated decision-making is further complicated by the chronicity of many critical illnesses and the frequent occurrence of shifts in ill patients' conditions.

Ignoring the concerns inherent to critical care may lead to very poor patient outcomes, translating into both poor morbidity and mortality results and demonstrating significant expenditures. Disparities are especially pronounced for critically ill patients, where organ dysfunctions contribute to a high mortality rate, and emergency surgery risks are indicated, for example, in abdominal surgery. An additional systemic factor includes faulty coordination and communication between different care teams involved and the use of inadequate end-of-shift care handover, which has been associated with treatment delays and the occurrence of preventable complications. Communication barriers leading to delays and condition aggravation have also been reported within and across care levels. Not least, ethical and legal responsibilities frame the activities of medical professionals and considerably contribute to effective clinical decision support, including end-of-life and long-term care.

3. Machine Learning in Healthcare

In essence, machine learning uses mathematical models to identify and make sense of patterns in data. This technique contrasts with traditional computational methods, which rely on human input to solve complex models. Supervised machine learning models, specifically, are trained to learn the best patterns to identify outcomes in their training data and then apply those patterns to unseen testing data. These models usually have to be "tuned" to recognize specific characteristics of the data and can be applied to both structured data and unstructured data. In the context of critical care, machine learning finds promising applications for identifying patterns in large amounts of data to predict patient outcomes, recognizing risk factors, and planning diagnoses and/or treatments. The goal of machine learning in critical care, along with other domains, is to improve decision-making for medical providers and, ultimately, patient outcomes.

Healthcare has been marked by the increasing use of machine learning and data analytics in clinical care, treatments, and other health services. For example, machine learning algorithms have been applied to patient health data for prediction, diagnosis, and personalized treatment plans. Reinforcement learning methods have been developed to provide personalized treatment strategies. In addition, big data has been leveraged from distributing diagnostics and treatments, with the potential to impact healthcare. Despite these significant strides, some potential barriers to machine learning development within medical care and healthcare include resistance from healthcare providers and management, the need for creating effective and directive clinical guidelines for medical providers, advanced training for medical providers to help make decisions, and data privacy issues.

3.1. Applications in Critical Care

Clinical care for patients in critical conditions is as challenging as it is multifaceted, requiring constant data acquisition and transmission as well as meticulous patient monitoring. The mechanical application of machine learning algorithms to the analysis of the data collected in critical care settings has proved to be both powerful and informative. Numerous possible applications of AI exist within these areas, including, most importantly, the monitoring of vital signs. The continuous analysis of patient data allows machine learning models to detect trends that might have escaped human operators, with obvious ethical consequences arising from decisions to withhold this information. Risk stratification, that is, the determination of the need and intensity of monitoring strategies, is another application, the results of which can guide the deployment of intensive care resources.

One popular area for applying machine learning models is early, intermediate, and long-term patient outcome prediction. It is hoped that by algorithmically identifying patients, facilities can intervene to differentially influence identified outcomes. Predictive models exist that focus on end-point prediction and those that provide a risk assessment over time. The holy grail of critical care is personalized medicine, in which the therapy is tailored according to the future behavior and individual patient phenotype to optimize results. AI models might help to drive personalized treatment strategies. Machine learning models can support the generation of decision rules or heuristics that capture the effect of increasing numbers of patient variables on clinical activity. The models discussed above are dependent on the availability of data to train models that are then deployed in decision support. Most relevant for the middle- and late-range predictive models, the minimal data set that has been shown to be most effective is demographics, blood gas data, and condition at arrival in the emergency room.

3.2. Challenges and Opportunities

Several challenges exist when implementing machine learning in critical care. One common problem is the limited quality of available data, comprising missing values and noisy signals. Furthermore, the integration of new tools with existing clinical workflows and electronic health records is not trivial because it requires proper access to computational infrastructures. Moreover, healthcare professionals should acquire new skills, and organizational procedures need to be adapted to make use of this technology. In some cases, machine learning black-box models for risk stratification or complete treatment policy optimization may not be accepted owing to a lack of trust. The risk of biased predictions leading to discrimination or mistreatment of certain individuals due to socio-economic factors or gender introduces potential ethical problems. The use of sensitive patient data needs to comply with privacy regulations, and more generally, the interfaces between machine learning and human clinical reasoning require a transparent, interpretive care provider interaction and associated ethical problems. In contrast to these challenges, intelligent data-driven predictive algorithms can improve the decision-making process in intensive care. It was shown that a combination of logistic regression with a relatively small neural network model indeed delivers more accurate predictions for the risk of arrhythmia than logistic regression alone. Risk prediction is a main focus of critical care medicine, but the complexity and stochasticity associated with the decision-making process complicate the development of good prognostic classifiers. The use of machine learning for these tasks can only be expected to do well in the highly complex problems encountered in medicine when better trained. Several promising machine-learning-based health predicting risk stratification and treatment recommendation systems have been developed for other applications outside of the critical care unit, and some have been put into practical use.

4. Predictive Modeling in Critical Care

Machine learning encompasses several methodologies that can forecast patient outcomes in the critical care setting with different purposes and characteristics. In this context, appropriate data collection and preprocessing are concerns before using any machine learning strategy. We seek to perform predictive modeling considering all available types of data, such as clinical records, laboratory and other clinical results, demographic information, and chemoinformatics. We then extract only one set from that, aiming at only demographic and laboratory results, to be used in a machine learning strategy example. In the preprocessing approach, we conduct feature selection and engineering, choosing relevant variables and generating additional data or information associated with each patient that could improve model performance. We describe the feature selection approach considering the computational expense, the sparsity in some data sources, the relevance of relevant variables, and how important they are in clinically relevant features. We also create a method that deals with time course features.

Once the data preprocessing is performed, we can develop models for the prediction of patient outcomes from those specific sets of data. Several algorithms can be used with different data and related feature preprocessing. In general, the models can be divided into two main groups: unsupervised or supervised. Unsupervised learning groups similar patients using clustering methodology. In contrast, supervised models use data from patients with known outcomes to learn a mapping from input to output. Because there are many available algorithms and approaches, with most of them providing multiple settings or parameter options, it is difficult to compare them in many clinical applications. Discussing the comprehensive options is unnecessary at many times. The outcome and the settings of the models are evaluated using standard metrics, such as accuracy, sensitivity, F-measure, among others. It is important to discuss the pertinent outcome settings in clinical applications and validate the developed models to avoid overfitting and improve model generalization in new sets of patients or in different clinical environments. Model validation and clinical use, as well as the limitations and the future of the models, are discussed.

4.1. Data Collection and Preprocessing

Data is the key factor for developing machine learning models for applications in critical care, as it is used to train algorithms in the pre-implementation process. Collected data is either part of electronic health records, patient monitoring systems, or the outcome of laboratory tests. Given the inheritance of resources, particularly medical history records, clinicians are often not able to predict which information will be used to train a machine learning model; it is rather collected and stored over time. Collecting high-quality data is one of the most demanding steps in critical care facilities due to the urgency of the situation and the unconscious state of patients. Similarly, in multiple studies, the data has been obtained using a pragmatic approach where the data were extracted using a customized script and prospectively reviewed for accuracy. Preprocessing the data starts at the time the data is collected and can be divided into two parts: preprocessing and postprocessing. The preprocessing phase involves the elimination of missing, inconsistent, or uncertain data in the dataset that helps convert the dataset into an informative dataset. This will reduce the noise and inconsistency in the input dataset, where the postprocessing or data preparation step transforms the fully preprocessed data into a form that can be handled by machine learning algorithms. It includes steps such as partitioning the dataset into training and testing, or cross-validation test sets. Partitioning the dataset into training, validation, and testing sets in the correct ratio is an essential step to developing and building robust predictive models. All subsequent steps of modeling will rely on the robustness of the earlier steps.

4.2. Feature Selection and Engineering

This relies on the identification of patient data that are most relevant or predictive of outcomes before predictive models can be developed. This is particularly relevant to the analysis of clinical data for outcomes in critical care populations when hundreds of basic measurements are usually taken from each patient. It is suggested that there is a relatively small amount of patient data that are most relevant to mortality. Five variables have been identified as the most relevant to predicting 30-day mortality, and minimum diastolic blood pressure and mean heart rate were found to be most highly associated with 28-day mortality. It is mentioned that statistical tests are the most common method for selecting features. However, a survey of feature selection suggests that different types of methods will select different subsets of relevant variables.

This is a free text field for any other critical analytical tools and choices to include back design. Genetic programming feature selection was used to create a rule set for predicting intensive care mortality. A support vector machine was used as another model type to assess performance, suggesting that feature selection will be important when using different models. Feature engineering is another approach to feature selection, whereby new variables are defined that identify underlying structure in the data that may not be represented in its original form. This approach is an essential part of model building when domain expert knowledge may enable new variables to be constructed that are relevant or contain information. Knowledge of physiology was used to derive a variable to be used in a prediction model. Projected mean velocity and parathyroid hormone variability were derived from measurements of corrected flow time in cardiac output measurement. Respiratory rate was measured from a database of vital signs in critically ill patients before site or organ failure levels were identified.

Pitfalls of including features that are not relevant include including variables that do not contain clinically relevant information about patient status, which will dilute the predictive accuracy of a model. Features that are directly predictive of outcome may become weaker with irrelevant features included. Additionally, redundant features can increase the complexity of models. Including an irrelevant set of features may lead to a model not transferring to new, unseen data, as it is overfit. Redundant features can also make models less comprehensible. For example, in the area of hyperglycemia, the creation of new parameters from routine blood glucose measurements using measures of glycemia variability, skewness, kurtosis, glycemia slope measurements, and daily mean slope changes in addition to meal times has been proposed. A significant expansion of the operational capacity would be needed to add the wider ranges of data associated with these new parameters. Nor would expanding the number of measured variables necessarily improve patient outcome prediction. Additional parameters might provide little new information if the underlying structure or model structure is similar to existing models developed using feature selection or feature extraction, for example. In this way, including a range of additional but irrelevant parameters increases model complexity without improving predictions.

4.3. Model Development and Evaluation

Model development and evaluation in critical care predictive modeling involves the choice of machine learning model for the task of interest, such as logistic regression or decision tree models for binary prediction, and the use of state-of-the-art models such as neural networks or random forests. Model development is conducted iteratively, whereby multiple models and feature sets are developed and tested before one is selected as the "final" model. The choice of metric used to evaluate model success depends on the primary objective. For diagnosis, it is often accuracy, but for prognostic modeling, it is often sensitivity, specificity, or the area under the receiver operating characteristic curve.

When a model is tested on data other than the data used to train it, it can be expected to achieve a metric similar to that achieved on the training data, which motivates the use of validation techniques. In machine learning, cross-validation is a particularly useful technique for use with small datasets. In healthcare, typically 5- or 10-fold cross-validation is used. Larger datasets may utilize additional hold-out sets for use in testing after some (or sometimes all) model development has been conducted. The key challenges in model development include overfitting, where a model describes the training data too well, leading to high variance and poor performance on unseen data. Underfitting is also a risk, where a model is too simplistic to capture the underlying structure of the data. In techniques, these issues can be addressed by reducing model flexibility, increasing flexibility, and by increasing the quality of the training data and the size of the available dataset. A final important part of model development is in the testing and intended use of the model; in general, the more complex the model, the more data is required. Setting a threshold for acceptable performance is a crucial step before a given model can proceed to the clinical testing phase.

5. AI Approaches for Risk Assessment

There are different AI approaches for risk assessment in critical care that we will address separately for the prediction of patient adverse outcomes and for the personalization of critical care therapies. In the past few years, various machine learning models have been developed to predict adverse patient outcomes based on historical patient data. These models are based on different classification algorithms like decision trees, logistic regression, support vector machines, neural networks, boosted methods, and ensemble methods, which have shown to be effective for risk stratification in various applications. To correctly identify an appropriate algorithm for building risk stratification models, one should consider the nature of clinical data, the processing of which will result in a suitable method of data representation and the predictive algorithm or computer-driven model that best fits the data for better prediction.

A different way to evaluate the transfer of knowledge from the data to the developed models has been proposed using two separate evaluation sets: one for predicting the independent targets and the other for the prediction of adverse outcomes, where an ensemble model combining clinical expertise and algorithmic prediction on the second representative evaluation group was considered. Modeling performance was evaluated using different performance measures such as precision, recall, F1 score, the area under the receiver-operator characteristic curve, and Net Reclassification Index when model outputs were interpreted as probabilities. Since clinicians have to be able to interpret machine learning outputs while making critical care decisions, machine outputs are generally interpreted as classification or regression predictions.

5.1. Classification Algorithms

A large fraction of AI applied to critical care problems centers around risk assessment. Predicting binary or ordinal outcomes like hospital mortality, near-death experiences, discharge to nursing homes, and 30-day rehospitalization does not impose the need for time-specific algorithms and is hence simpler to model. However, due to the large gamut of classification algorithms in AI, a multitude of options are available to the investigator. A foremost, time-tested algorithm in the critical care domain is the decision tree algorithm. Decision trees and their extensions are very appealing to clinicians and data analysts due to their simplicity and interpretability. Decision trees can be readily interpretable, and with their intuitive structure, are akin to the multiple decision-making processes of experienced clinicians. Random forests are a simple extension of decision trees with a bootstrapping algorithm that allows for predictions made by a plurality of trees. Boosted trees use a different sort of weighting process in their bootstrapping and can be run in sequence rather than parallel. Neural networks, in turn, utilize a system of nodes in layers; the first layer contains an input layer, with an output layer containing a score. The layers in between the input and output are called hidden layers. Several different strategies exist for building a neural network model, such as deep feedforward networks, deep learning, and convolutional networks. Each of these algorithms has its own strengths and weaknesses, the most important of which for critical care modelers is predictive accuracy versus interpretability. Subsequently, the choice of algorithm should depend upon the clinical question and purpose of building the model. In addition, unique factors such as volume and downstream applications may also come into play when selecting an algorithm. A deluge of examples demonstrating the successful application of AI in predicting patient outcomes in critically ill patients exists, demonstrating a wide gamut of algorithms. Ultimately, by utilizing these algorithms, we can bridge the gap between understanding the patient's overall physiological deterioration and the effect on specific shared biomarkers. Along these lines, we can utilize machine learning to tailor personalized treatment strategies that may ultimately save lives. Additionally, many have added complexity to their models in the name of accuracy at the expense of clinical interpretability with diminished added clinical value.

Such a fact indicates the necessity of balancing model complexity and clinical impetus so that the developed models can be integrated into clinical applications and/or guidance. A majority of the research does not tie back to changes at the bedside. Furthermore, many of the aforementioned findings may not easily translate into any visible outcomes in terms of diagnosing and treating patients in real-world settings. For example, while the timing of surgery can altogether be avoided in 90% of patients if operations are deliberately postponed by 24 hours, discerning such patients with accuracy is the foreign concept that can be answered by AI and utilized to provide guidance on treatment.

5.2. Performance Metrics

In classification problems for critical care patients, it is essential to identify model performance by means of performance metrics. Accuracy measures the proportion of correctly classified instances. Precision indicates the proportion of true positive cases out of all the predicted positive cases, which is also called positive predicted value. Recall, also known as sensitivity, shows the fraction of actual positive instances correctly predicted by the model. F1 score, also known as the F-score or the F-measure, is a measure of a test's accuracy. The area under the receiver operating characteristic curve (AUC-ROC) is a plot of true positive rate against the false positive rate. It tells how much the model is capable of distinguishing between classes.

The choice of the right metric also depends on what we want to optimize in terms of risk management and treatment strategies. Selecting an appropriate metric, depending on the specific clinical context and practical risk implications, is extremely important. Accuracies might not be appropriate in imbalanced classification problems, where the distribution of classes is not equal, for instance in the prediction of mortality, where the number of deceased patients is much lower than that of the healed patients. Thus, advice on which metrics to choose for the evaluation of critical care prediction models is needed. Guidelines on AI in critically ill patients recommend ongoing monitoring and evaluation, including external validation in clinical settings, with systems calibrated and risk factors verified for model understandability.

6. Personalized Treatment Strategies

Treatment paradigms in critical care, as in many areas of medicine, have traditionally been based on generalized population-level descriptions of risk. Thankfully, the last decade or so has seen a dramatic shift in the direction of personalized treatment strategies—that is, treatments that are designed and intended for each unique patient affected by a given axis of risk (be it cancer, heart disease, and so forth) in a way that aligns with that patient's unique risk profile and general condition as denoted by various high-dimensional or low-dimensional molecular, imaging, laboratory, and clinical characteristics. More generally, as pharmacologic and other medical therapy options continue to diversify in number and type, so too does the impetus to utilize the vast amounts of health-related data in describing the evolution and outcomes of unique individuals. Moreover, patient-centered care that protects the true autonomy exhibited by each patient affected by disease should not be conflated with more general population-level considerations in developing treatments.

In principle, one could perform a deep query into heterogeneous patient data sets and emerge with exquisite descriptions and categorizations of each unit along myriad health informatics relevant dimensions. Explicit treatment assignments, or at minimum, treatment selection indices could then be developed using the patient-specific profiles. In theory, these patient-specific indices would reflect not only the scientifically and clinically driven model of the outcomes for that unit under each treatment option, but would do so in a way that emphasizes what interventions are beneficial for such units. Integration of these AI-driven predictive models with clinical decision support systems has the potential to leverage this valuable data to further tailor treatment recommendations to individual patients, for instance by predicting adverse events or high-risk time windows where treatment might be particularly hazardous.

One central issue here is the potential implementation of this giant leap within critical care, particularly with respect to its respectful integration into contemporary care schemas. A significant amount of prior work and algorithm development has addressed at least some of these issues. Machine learning and predictive modeling as quantitative strategies to tailor interventions have indeed been successfully incorporated in a diversity of areas of treatment within critical care scenarios. While the possibilities are vast, and the technology is advancing, technological prowess and capacity are still necessary but insufficient in these considerations. Similarly, aside from the clinical or biological scientific validation needs for these strategies, attention must also be paid to the ethical and societal dimensions of developing powerful new capabilities, and must attend to the development of safeguarding procedures, if it is reasonable to expect that such strategies will eventually be implemented into routine clinical use in the not distant future. This review attempts to explore some of these foundational dimensions, uncover existing and persistent gaps in our understanding, and discuss case studies to provide illustration into the principles explored.

6.1. Precision Medicine

In critical care, precision medicine is defined as the delivery of personalized treatment strategies for a selected group of cases based solely upon a robust validation of probability gained by powerful in-silico strategies to predict patient response to a given treatment. Personalized treatment may be shaped by a patient's individual data, such as records of disease, genetic profiles, or other markers like serum protein or metabolomic patterns. This definition captures not only possible genetic-based predictions of treatment response but can also include phenotypic predictions based on any combination of a patient's attributes, such as clinical history, laboratory results, and imaging.

Due to their power for prediction, machine learning algorithms offer robust statistical tools to underpin personalized medicine strategies by predicting patient or organ response to a given intervention. Despite this capability, many critical care questions cannot be answered using only clinical data, and a personalized medicine paradigm will need to integrate diverse data sets. The data used will depend on the prediction problem and may include data such as transcriptional, genomic, proteomic, metabolic, or other relevant markers of the disease process under investigation. This will ensure that identified strategies and treatments will bring the correct influence to address the patient's underlying cell dysfunction, irrespective of clinical presentation. Vital in the optimal pursuit of precision medicine in critical care is patients' consent to access all the needed data, along with mechanisms providing security to ensure privacy when storing and sharing data.

While demonstrating clear potential, several challenges must first be navigated if this approach is to be incorporated into routine clinical care in critical care: the availability of resources to interact with patients outside normal working hours to obtain consent would be substantial and may prove a significant barrier to implementation. Similarly, the potential for variations in access according to geographical location, race, or social status would need to be addressed. Finally, while the algorithms developed using these data will predict responses at population levels, they do not guarantee the effect on an individual, which may or may not be the same. The time for clinical implementation of these techniques has not yet arrived.

In the development and optimization of this practice, researchers have suggested that the use of machine learning performs best for patients on therapy that were similar, such as those on low mechanical ventilation after cardiac surgery and those with a high IL-6. More comprehensively, serine and arachidonic acid metabolomics could be identified as a potential personalized medicine strategy for patients critically ill with SIRS following cardiac surgery.

6.2. Decision Support Systems

Personalized treatment strategies bring new opportunities to critical care, particularly for interventions promoting individual patient outcomes. In the ICU, decision support systems are tools designed to inform healthcare providers in their decision-making processes by providing computations and predictions made from studying the patient over time, enhanced by predictive analytics. Personalized treatment strategies can be expected to be highly valuable in many areas of critical care settings, as they permit adaptation of interventions to the current evolution of the patient's state, thereby increasing the therapeutic index. A substantial proportion of decision support systems are based on machine learning models that aim to provide several features, such as improved diagnosis, prediction, selection of interventions, prediction of outcomes, severity of illness estimation, or longer-term classification and prediction. The added value of these new models developed from large and diverse datasets is that they provide better accuracy, such as area under the receiver operating characteristic curve, area under the precision-recall curve, area under the time-dependent receiver operating curve, or area under the time-dependent precision-recall curve, or calibration in the expected setting, identifying interactions and selection of suboptimal patients from existing treatment and diagnostic facilities.

Additional information can be available on the user interface, such as prognostic prediction consisting of the risk of the date-specific patient population; regulations compatible with healthcare and safety; the availability of diagnostic data on a state-of-the-art injury or the latest waiting days from an infection; or personalized therapy management, updated illness tips, plus biomarkers predicting the duration of a pathological state or the expected long-term prognosis. As decision support systems aim to intervene in clinical decision-making processes, their added value is of a substantially higher standard of care and predictive model in terms of reducing between-center variation, facilitating best practices, or helping meet internationally recognized guidelines. Under the introduction of personalized treatment, the patient decision-making guideline requires a more feasible and accepted impact on patient care, including the end-of-the-year risk reduction, patient experience, and overall burden of the healthcare system. An effective comprehensive treatment for septic patients would limit anti-infective consumption, thereby delaying or moderating the severity of any additional healthcare-acquired conditions and reducing hospital length of stay. Also, selection of both best practices or first-in-patients with particularly long survival rates introduced early or later in the treatment algorithm could substantially underpower trial or registry designs.

As outlined in the introduction, the introduction of personalized markers or therapeutic options considerably extends the development and use of existing biomarker-based diagnostic capacity technologies. One should keep in mind that resistance of the existing practice to derisking diagnostic or therapy decisions is mainly connected with the previously known detrimental evolution of the local and systemwide patient/professional/community decision-making. More specific challenges are the need for a major reimbursement strategy for any supportive cockpit operation decision support systems, including maintenance and upgrade, and potential costs for system-wide electronic health record adaptations and coordinated clinical/HR system meta-retraining courses.

7. Case Studies and Applications

Several applications of machine learning models for predictive analytics in critical care have been introduced throughout this review. From intensive care units in hospitals to outpatient settings during natural disasters or outbreaks, AI can support medical professionals in their efforts to predict patient outcomes and therefore tailor healthcare pathways more effectively. With implementations that range from the prediction of mortality, severity of sepsis, and nosocomial infection to personalizing therapeutic strategies in critically ill patients, the covered studies offer diverse insights into the potential predictive value of machine learning models and explain real-life problems in critical care that can be supported by such methods. An interdisciplinary collaboration between medical doctors, data scientists, and statisticians has been a pivotal factor of success for several of the above-mentioned case studies. Commitment to thoroughly preparing clinical data before the model can be trained and applied, thus ensuring data quality, emerged as another key factor in the expected success of the model deployment, especially given the sensitivity of critical care data. However, several studies also discussed challenges along the way, especially regarding difficulties in training models on data from heterogeneous healthcare settings and implementing machine learning models into clinical practice. In this review, the four cornerstones of machine learning model deployment in critical care will be explored, summarized, and discussed in detail.

8. Future Directions and Emerging Trends

8.1. Continued Research Critical care offers a unique environment for promising AI research applied to patient populations with a high projected rate of adverse events. Active as critical care has been and remains in the area of clinical decision support research, many questions and ongoing investigations remain. Many critical care patients have developing trajectories that will place them at a higher risk of dramatic worsening before the next scheduled contact with a care provider. How can we maximize our chance of identifying these patients ahead of time and obtaining optimal predictions about what their trajectories will involve? Continued focus on these questions is a well-justified—and still-exploding—research direction.

8.2. Methodological Advances The intensified interest in critical care decision support has paralleled the exploding capacity for machine learning in operational healthcare. Many old and newly blossoming techniques are being tested for their potential in risk stratification and therapeutic guidance. The potential for techniques like deep learning to be used for fast, real-time analytics could help circumvent the "tar pit" of health research with currently available databases.

8.3. Interdisciplinary Benefits in AI and Clinical Sciences While AI techniques for critical care offer much promise, there is also ample opportunity to improve AI by interfacing it with expert insights from a variety of clinical domains and enacting more aggressive forms of evidence-based medicine. Given the traditional early involvement of interdisciplinary distant collaborators present in new initiatives for data-driven healthcare, the curiosity of medical professionals is not only in the potential to directly improve patient care but also to transform the ways we answer clinical questions.

8.4. Policy and Economic Considerations AI for health has the potential to greatly outperform human experts, provided appropriate data are also available. The main obstacle to this vision is our inability to effectively transform the mass of available data. Policies should be designed to leverage work on the optimal integration of human intuition with AI techniques capable of handling increasingly high-dimensional and massive data, and to facilitate accelerated clinical validation and deployment of AI-based technologies.

8.5. Ethical Considerations in Evolving AI As AI continues to proliferate and evolve within the healthcare realm, it is important to remain critical of the potential ethical pitfalls this can present. More studies should be performed to identify and address the biases and accuracy in prediction models. There is evidence that few model development studies for quality of care assessment make it to practical application, and there are major caveats with the most frequently applied models. This finding underscores the needed next steps if machine learning tools and innovations are to continue successfully transforming medicine for the next decade.

8.6. Personalized Therapies Data-driven insights can help us improve critical care and inform the choices most likely to affect patient outcomes. In some years, we may even find deviations in therapies regulated by predictive data rather than by guiding randomized trials. Experimenting with the conterminous design of trials with big health data is well underway, and it will be interesting to watch the results of such research endeavors emerge in the future.