Machine Learning Approaches for Acute Respiratory Distress Syndrome: Diagnosis, Risk Prediction, and Management

Pawan Kumar Mall[1], Prof Dr Divya Midhun[2], Dr. Rupali Atul Mahajan[3]
[1] Lincoln University ; Lincoln University[2] ; [3] Vishwakarma Institute of Technology, Pune
Email ID pawankumar.mall@gmail.com[1];divya@lincoln.edu.my; rupali.mahajan@viit.ac.in

---

Abstract: Acute Respiratory Distress Syndrome (ARDS) is a serious issue. severe illness affect diffuse alveolar inflammation, terminate in respiratory failure. Since Acute Respiratory Distress Syndrome (ARDS) is a severe and possibly lethal illness that significantly influences patient morbidity and mortality, clinical care strategies are directly impacted. In this study, the use of machine learning (ML) models in diagnosing ARDS, predict risk, and management is investigated. Support Vector Machines (SVM), Random Forest (RF), Decision Trees (DT), and Multilayer Perceptron (MLP) were tested for predictive performance. A clinical and demographic factor dataset was evaluated through these models with performance metrics like accuracy, Matthews Correlation Coefficient (MCC), and F1-score. The performance metrics for training different classifiers show that the SVM with RBF kernel achieved the highest accuracy (0.974), along with a strong MCC (0.931) and F1 score (0.954), indicating excellent predictive capability. The outcomes show that SVM and RF models perform better compared to other classifiers.

Keywords: Machine Learning; Acute Respiratory Distress Syndrome; Risk Prediction; Diagnosis; Random Forest; Support Vector Machine

---

Introduction

Acute Respiratory Distress Syndrome (ARDS) is an acute and potentially fatal condition that largely affects patient morbidity and mortality. There is direct impact on clinical management approaches. ARDS is a frequent condition seen in critical care during the COVID-19 era. It is an enduring challenge in practice for a long time [1]. ARDS is defined by acute diffuse alveolar inflammation and ensuing damage, producing extreme respiratory failure. The inflammatory response compromises the alveolar-capillary membrane, causing pulmonary infiltrates, hypoxemia, and stiff, non-compliant lungs [2]. Usually within seven days of an initial insult, is a characteristic of ARDS. The syndrome was initially described in 1967 and has subsequently been more specifically defined. The most commonly accepted criteria being established by the Berlin Definition in 2012 . This definition, supported by the European Society of Intensive Care Medicine, the American Thoracic Society, and the Society of Critical Care Medicine, describes four main criteria: acute onset over one week, bilateral pulmonary infiltrates, ruling out cardiogenic pulmonary edema, and PaO_2/FiO_2 ratio of <300 mmHg with a positive end-expiratory pressure (PEEP) of >5 cm H_2O.

ARDS is associated with a high mortality, varying from 27% in mild cases to 45% in severe cases based on the degree of hypoxemia. ARDS is a frequent complication of critically ill patients and occurs in up to 25% of intubated patients. Damage to alveolar and capillary structures underlies the pathophysiology of ARDS and leads to deranged gas exchange and hypoxemia. Inflammation and alveolar-capillary membrane increased permeability enable fluid and cellular waste to collect in the alveoli and interstitial space, further impairing lung function. Pulmonary hypertension and intrapulmonary shunting also lead to systemic hypoxemia and multi-organ failure, which are the major causes of death in ARDS patients [3].

It is important to know the underlying causes of ARDS in order to manage it effectively. The condition may result from direct (primary) or indirect (secondary) insults. Direct insults are pneumonia, aspiration, pulmonary contusion, and toxic inhalation, which cause direct damage to the alveolar epithelium [4]. Indirect insults like sepsis, trauma, pancreatitis, and transfusion-related acute lung injury (TRALI) induce systemic inflammation that causes damage to the capillary endothelium. Both routes eventually result in the typical features of ARDS, such as alveolar collapse, fluid buildup, and compromised gas exchange [5].

Related work

Table 1. Highlighting advancements, methodologies, advantages, and limitations of AI-driven approaches in ARDS.

Reference	Objective	Methodology	Advantages	Limitations

[6] Li et al. (2025)	To explore how AI can transform sepsis management (early detection, personalized treatment, real-time monitoring).	ML techniques like random forest and deep learning were used to analyze EHR data for early detection, risk prediction, and treatment personalization.	Improved early detection, personalized therapies, real-time complication monitoring, dynamic risk assessment.	Ethical challenges (data privacy, algorithmic bias), limited dataset diversity.
[7] Daramola et al. (2025)	To assess the performance of ML models for predicting COVID-19 mortality risk with cross-validation and PCA.	ML models (Deep MLP, SVM, XGBoost) with CV, SMOTE, and PCA applied to ICU patient data. SHAP used for model interpretability.	High accuracy (Deep MLP: F1=0.92, AUC=0.94), identification of critical variables like LOS, D-dimer, CRP, etc.	Limited dataset (single hospital), minimal exploration of cross-regional model performance.
[8] Garrido et al. (2025)	To identify vital biomarkers for prognosis during emerging pandemics using AI.	Random Forest algorithm analyzed 89+ variables to identify key mortality biomarkers.	High accuracy (92.61%), rapid identification of biomarkers (e.g., PCT, LDH, CRP), improved prognosis and treatment.	Focused on COVID-19; generalizability to other diseases needs validation.
[9] Groenland et al. (2025)	To investigate cardiac and inflammatory biomarkers predicting extubation failure in COVID-19 ARDS.	Logistic regression analyzed biomarkers (NT-proBNP, Hs-TnT, PCT) for extubation failure in ICU patients.	Identified Hs-TnT as an independent predictor, providing objective measures for extubation readiness.	Single-center study, limited sample size, focus on specific biomarkers.
[10] Zhao et al. (2025)	To evaluate SIRI and APACHE II score in predicting ARDS prognosis caused by wasp stings.	Retrospective study using logistic regression and ROC curve analysis of clinical data.	Identified significant predictors (APACHE II score, SIRI) with high predictive value (AUC: 0.821, 0.809).	Limited to wasp sting-related ARDS, small sample size (46 patients).
[11] Sarma et al. (2025)	To explore ML applications in cardiac ICU for improving care, education, and personalized therapies.	Reviewed evidence on ML applications like ECG interpretation, dynamic risk stratification, and clustering sub-phenotypes.	Enhanced diagnostics, individualized therapies, optimized triaging, and training opportunities.	Practical and ethical challenges, reliance on clinician understanding of AI.
[12] Pan et al. (2025)	To develop an ML-based model for predicting mortality in SCAP ICU patients.	LightGBM model validated using nested cross-validation; 23 key features selected via Recursive Feature Elimination.	High AUC (0.842), validated web calculator for real-time clinical use.	Dataset limited to two centers; requires broader validation.
[13] Pedarzani et al. (2025)	To develop a dynamic ML mortality prediction system for ICU patients using repeated measurement data.	MixRFb algorithm integrating RF and mixed-effects models with RDW and SAPS scores.	Superior performance (AUC=0.882), dynamic prediction using repeated measures, ethical trial selection.	Limited to SAPS scoring framework, requires broader application in diverse ICUs.

Key Contribution

This paper provides a comprehensive comparative analysis of ARDS using the latest machine learning (ML) models applied to diagnosis , risk prediction, and management. By systematically evaluating state-of-the-art ML techniques.

Method, Experiments and Results

Dataset: This dataset contains comprehensive information on patients with lung cancer, including demographic details, lifestyle factors, environmental exposures, and clinical symptoms. The dataset is particularly useful for building predictive models to identify high-risk individuals and tailor personalized interventions.

Table 2: Dataset description

Feature	Data Type	Description
Age	Numeric	Represents the patient's age.
Gender	Categorical	Specifies the patient's gender.
Air Pollution	Categorical	Indicates the patient's exposure level to air pollution.
Alcohol Use	Categorical	Represents the frequency or level of alcohol consumption.
Dust Allergy	Categorical	Indicates whether the patient has an allergy to dust.
Occupational Hazards	Categorical	Represents the level of work-related health risks.
Genetic Risk	Categorical	Indicates hereditary susceptibility to diseases.
Chronic Lung Disease	Categorical	Identifies whether the patient has a chronic lung condition.
Balanced Diet	Categorical	Reflects the patient's dietary habits and nutritional balance.
Obesity	Categorical	Represents the patient's level of obesity.
Smoking	Categorical	Indicates the patient's smoking habits.
Passive Smoker	Categorical	Represents exposure to secondhand smoke.
Chest Pain	Categorical	Identifies the presence and severity of chest pain.
Coughing of Blood	Categorical	Indicates whether the patient has experienced blood in cough.
Fatigue	Categorical	Represents the patient's level of tiredness or exhaustion.
Weight Loss	Categorical	Identifies any unintentional weight loss.
Shortness of Breath	Categorical	Indicates breathing difficulties.
Wheezing	Categorical	Represents abnormal high-pitched breathing sounds.
Swallowing Difficulty	Categorical	Identifies any issues in swallowing food or liquids.
Clubbing of Finger Nails	Categorical	Indicates nail deformities linked to respiratory/cardiac conditions.
Frequent Colds	Categorical	Represents the frequency of common colds.
Dry Coughs	Categorical	Identifies occurrences of persistent dry coughing.
Snoring	Categorical	Indicates whether the patient experiences snoring.

Pre-processing:

The bar chart in figure 1 shows the association between difficulty with swallowing and alcohol consumption, and there are some interesting trends. The highest alcohol consumption is among those with a level of swallowing difficulty of 7, and the lowest is among those at level 5. This indicates that the severity of swallowing difficulties could be affecting drinking behavior, but the association is not linear. The range in alcohol consumption between levels, represented by the error bars, demonstrates considerable variability of experience among individuals. These data suggest a rich interaction between bodily distress and drinking that deserves study of the possible underlying psychological or social processes.

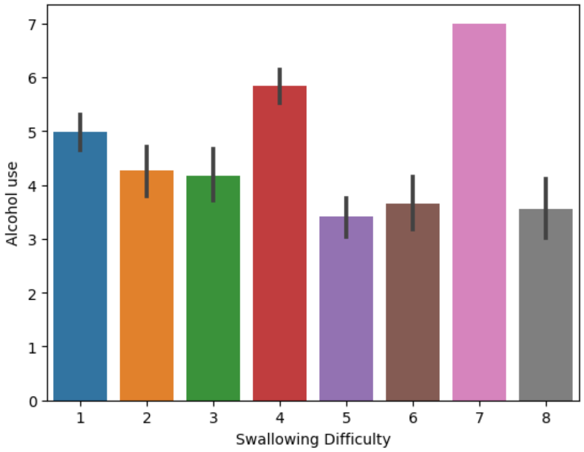

Figure 1: The relationship between swallowing difficulty and alcohol use,

The bar chart in figure 2 shows the connection between weight loss and smoking, with significant differences between the different levels of smoking. Weight loss is greatest at smoking level 1, implying that non-smoking or minimal smoking could be linked to increased weight loss. At smoking level 8, the least weight loss is experienced, which may indicate a connection between heavy smoking and lower weight loss. The trend is not linear, with moderate levels of smoking (6 and 7) having comparatively high weight loss than the other groups. Error bars represent variability within each group, implying differences between individuals in weight loss experience. These results suggest a multifaceted interaction between smoking behavior and weight control, and further investigation into physiological or lifestyle determinants of this relationship is needed.

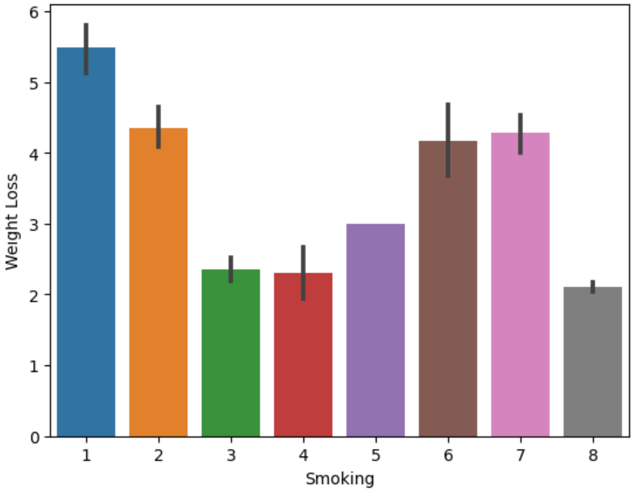

Figure 2: The connection between weight loss and smoking

The bar graph in figure 3 shows how smoking relates to varying levels (Low, Medium, and High). Smoking is most prevalent in people who are in the High level group, whereas those who are in the Medium and Low groups have successively lower smoking levels. The gap between the two groups is quite different, suggesting a clear gradient in smoking prevalence across levels. Error bars indicate some diversity within each category but are themselves quite small and indicate uniform patterns of smoking amongst individuals in each category. What these results bring to light is a strong linkage between the classed levels and smoking intensity implying that with a rise in the level smoke activity becomes more likely.

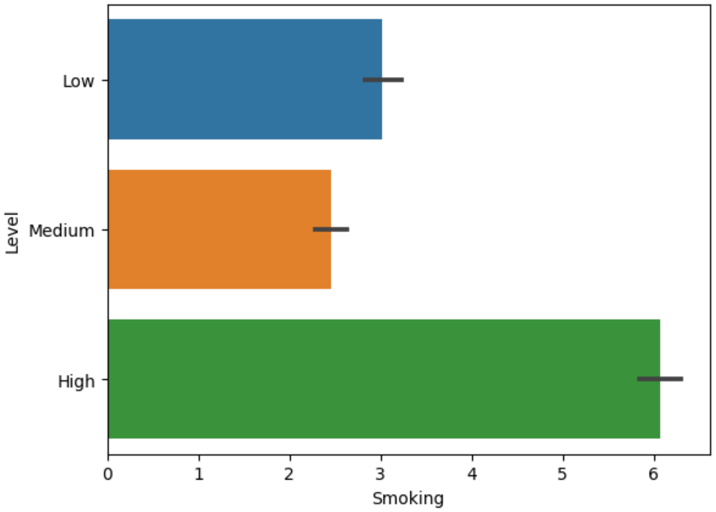

Figure 3: The relation between smoking relates to varying levels

Different ML Models:
comparison of the K-Nearest Neighbors (KNN), Support Vector Machine with RBF Kernel (SVM_RBF), Decision Tree (DT), Random Forest (RF), and Multilayer Perceptron (MLP) algorithms in terms of their key characteristics, advantages, and limitations:

Table 3:Different ML Models

Algorithm	**Key Characteristics**	**Advantages**	**Limitations**
KNN	- Instance-based learning - Non-parametric - Lazy learner	- Simple to implement and understand - No training phase - Effective for small datasets	- Computationally expensive for large datasets - Sensitive to irrelevant features and noise
SVM_RBF	- Kernel-based method - Effective for non-linear data - Margin maximization	- High accuracy for complex datasets - Robust to overfitting in high-dimensional spaces	- Computationally intensive - Requires careful tuning of hyperparameters (e.g., C, gamma)
DT	- Tree-based model - Splits data based on feature values - Interpretable	- Easy to visualize and interpret - Handles both numerical and categorical data	- Prone to overfitting - Sensitive to small changes in data
RF	- Ensemble of decision trees - Bagging technique - Reduces overfitting	- High accuracy and robustness - Handles missing data and outliers well	- Computationally expensive - Less interpretable than single decision trees
MLP	- Feedforward neural network - Multiple layers of neurons - Non-linear mapping	- Can model complex, non-linear relationships - Scalable to large datasets	- Requires large amounts of data - Computationally expensive and hard to interpret

Table 4: ML Model Metric

Metric	**KNN**	**SVM_RBF**	**DT**	**RF**	**MLP**
Accuracy	Moderate	High	Moderate	High	High
Interpretability	Low	Low	High	Moderate	Low
Training Speed	Fast (no training)	Slow	Fast	Moderate	Slow
Scalability	Poor (large data)	Moderate	Moderate	High	High
Overfitting Risk	Low	Low	High	Low	Moderate

Result:

1. Accuracy: Measures overall correctness of the model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

Measures the proportion of correct predictions.

2. Matthews Correlation Coefficient (MCC):

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \qquad (2)$$

A balanced metric for binary classification, even when classes are imbalanced.

3. F1 Score: Balances precision & recall for better performance evaluation.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (3)$$

The harmonic mean of precision and recall, balancing false positives & false negatives.

The performance metrics for training different classifiers show that the SVM with RBF kernel achieved the highest accuracy (0.974), along with a strong MCC (0.931) and F1 score (0.954), indicating excellent predictive capability. The Random Forest (RF) model also performed exceptionally well, with high accuracy (0.954), the highest MCC (0.941), and the best F1 score (0.974), suggesting robust classification power. Decision Tree (DT) followed closely with slightly lower accuracy (0.964) and MCC (0.911). The Multi-Layer Perceptron (MLP) and k-Nearest Neighbors (kNN) showed comparatively lower but still reliable performance, with MLP slightly outperforming kNN across all metrics. Overall, SVM and RF demonstrated superior classification effectiveness.

Table 5:Training result of different ML models

	Accuracy	MCC	F1
knn	**0.914286**	**0.871977**	**0.91363**
svm_rbf	**0.974176**	**0.93109**	**0.953961**
dt	**0.964176**	**0.91109**	**0.943961**
rf	**0.954176**	**0.94109**	**0.973961**

mlp	**0.925714**	**0.89109**	**0.923961**

The performance metrics for testing the classifiers reveal that the Multi-Layer Perceptron (MLP) achieved the highest accuracy (0.913) and F1 score (0.913), indicating strong predictive performance.Random Forest (RF) also yielded good results with high accuracy (0.905) and best MCC (0.897), indicating strong classification ability. The Decision Tree (DT) classifier was found to be consistent in performance with 0.875 accuracy and an MCC of 0.847. k-Nearest Neighbors (kNN) was found to be moderately accurate (0.85) and MCC (0.773), being consistent but less powerful than MLP or RF. On the other hand, the SVM using RBF kernel had poor performance in all the metrics, with the lowest accuracy of 0.403, MCC of 0.053, and F1 score of 0.236, implying classification challenges in this dataset. MLP and RF exhibited the best classification performance overall.

Table 5:Testing result of different ML models

	Accuracy	**MCC**	**F1**
knn	**0.85**	**0.773399**	**0.847849**
svm_rbf	**0.403333**	**0.05334**	**0.23572**
dt	**0.875463**	**0.84657**	**0.86087**
rf	**0.905463**	**0.89657**	**0.84087**
mlp	**0.913333**	**0.869141**	**0.913156**

Discussions

The findings of this study highlight the potential of machine learning (ML) in diagnosing and predicting outcomes for ARDS. Various ML models such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Multilayer Perceptron (MLP), and k-Nearest Neighbors (KNN) were assessed for their effectiveness in ARDS classification. The SVM and RF emerged as the most robust models, demonstrating superior accuracy, Matthews Correlation Coefficient (MCC), and F1-score.

The high performance of SVM and RF can be attributed to their ability to handle complex, high-dimensional clinical data, making them well-suited for identifying patterns in ARDS-related variables. The KNN exhibited lower predictive power due to its sensitivity to noisy data and computational inefficiency with large datasets. The Decision Tree models faced challenges with overfitting, reducing their generalizability.

One of the key observations in this study is the variability in ML model performance across training and testing phases. While SVM and RF performed exceptionally well in training, their performance slightly declined during testing, emphasizing the need for external validation on diverse datasets. Additionally, MLP demonstrated competitive results but required extensive computational resources and hyperparameter tuning.

The integration of ML in ARDS management has significant clinical implications. Early diagnosis facilitated by ML models can enhance patient stratification, leading to timely interventions and improved survival rates. Moreover, personalized treatment plans based on predictive analytics can optimize resource allocation in intensive care units (ICUs). However, the ethical and practical challenges of implementing AI-driven models in healthcare settings must be addressed, including data privacy concerns, algorithmic bias, and the need for clinician interpretability of model outputs.

Despite its promising results, this study has limitations. The dataset used, while comprehensive, may not be fully representative of global ARDS patient populations, necessitating further research with larger, multi-institutional datasets. Additionally, future studies should explore hybrid ML models and deep learning techniques to enhance predictive accuracy. The development of real-time AI-assisted diagnostic tools could further revolutionize ARDS management, paving the way for AI-driven decision support systems in critical care settings.

Conclusions

This study demonstrates the efficacy of machine learning models in predicting and managing ARDS. The evaluation of various ML techniques highlights that Support Vector Machines and Random Forest models achieve the highest predictive accuracy. By leveraging clinical and demographic data, ML approaches enable early diagnosis and personalized treatment, ultimately improving patient outcomes. Future work should focus on expanding datasets and refining algorithms to enhance generalizability across diverse patient populations. The integration of AI-driven tools in clinical settings has the potential to revolutionize ARDS management, offering data-driven decision-making for healthcare professionals.

References

[1] Matthay, M. A., Zemans, R. L., Zimmerman, G. A., Arabi, Y. M., Beitler, J. R., Mercat, A., ... & Calfee, C. S. (2019). Acute respiratory distress syndrome. Nature reviews Disease primers, 5(1), 18.

[2] Force, A. D. T., Ranieri, V. M., Rubenfeld, G. D., Thompson, B., Ferguson, N., Caldwell, E., ... & Slutsky, A. S. (2012). Acute respiratory distress syndrome. Jama, 307(23), 2526-2533.

[3] Bos, L. D., & Ware, L. B. (2022). Acute respiratory distress syndrome: causes, pathophysiology, and phenotypes. The Lancet, 400(10358), 1145-1156.

[4] Aslan, A., Aslan, C., Zolbanin, N. M., & Jafari, R. (2021). Acute respiratory distress syndrome in COVID-19: possible mechanisms and therapeutic management. Pneumonia, 13(1), 14.

[5] Beitler, J. R., Thompson, B. T., Baron, R. M., Bastarache, J. A., Denlinger, L. C., Esserman, L., ... & Calfee, C. S. (2022). Advancing precision medicine for acute respiratory distress syndrome. The Lancet Respiratory Medicine, 10(1), 107-120.

[6] Li, F., Wang, S., Gao, Z., Qing, M., Pan, S., Liu, Y., & Hu, C. (2025). Harnessing artificial intelligence in sepsis care: advances in early detection, personalized treatment, and real-time monitoring. Frontiers in Medicine, 11, 1510792.

[7] Daramola, O., Kavu, T. D., Kotze, M. J., Marnewick, J. L., Sarumi, O. A., Kabaso, B., ... & Nyasulu, P. S. (2025). Predictive modelling and identification of critical variables of mortality risk in COVID-19 patients. Scientific Reports, 15(1), 2184.

[8] Garrido, N. J., González-Martínez, F., Torres, A. M., Blasco-Segura, P., Losada, S., Plaza, A., & Mateo, J. (2025). Role of Artificial Intelligence in Identifying Vital Biomarkers with Greater Precision in Emergency Departments During Emerging Pandemics. International Journal of Molecular Sciences, 26(2), 722.

[9] Groenland, C. N., Siemers, A. H., Dubois, E. A., Gommers, D., Heunks, L., Wils, E. J., ... & Endeman, H. (2025). The prognostic role of cardiac and inflammatory biomarkers in extubation failure in patients with COVID-19 acute respiratory distress syndrome. Annals of Intensive Care, 15(1), 2.

[10]Zhao, Z., Zhang, X., Liu, D., Tao, W., Tong, C., Wang, Y., & Fang, Z. (2025). Predictive Value of the Systemic Inflammatory Response Index and APACHE II Score for Acute Respiratory Distress Syndrome Caused by Wasp Stings. Toxicon, 108252.

[11] Sarma, D., Rali, A. S., & Jentzer, J. C. (2025). Key Concepts in Machine Learning and Clinical Applications in the Cardiac Intensive Care Unit. Current Cardiology Reports, 27(1), 30.

[12] Pan, J., Guo, T., Kong, H., Bu, W., Shao, M., & Geng, Z. (2025). Prediction of mortality risk in patients with severe community-acquired pneumonia in the intensive care unit using machine learning. Scientific Reports, 15(1), 1566.

[13] Pedarzani, E., Fogangolo, A., Baldi, I., Berchialla, P., Panzini, I., Khan, M. R., ... & Azzolina, D. (2025). Prioritizing Patient Selection in Clinical Trials: A Machine Learning Algorithm for Dynamic Prediction of In-Hospital Mortality for ICU Admitted Patients Using Repeated Measurement Data. Journal of clinical medicine, 14(2), 1-18.