BIOSCIENCES REPORTS

https://biosciencesreports.com
editor@biosciencesreports.com

ISSN: 3104-400X (Print)
ISSN: 3104-4018 (Online)

LIQUID BIOPSY AI-BASED ANALYSIS OF CIRCULATING TUMOR DNA FOR EARLY DIAGNOSIS AND RELAPSE PREDICTION IN LUNG CANCER

[1] Muhammad Saif Ullah

[1] Aga Khan University Karachi, Sindh, Pakistan

*Corresponding Author E-mail: saifullah23@gmail.com

Abstract

One of the most important reasons lung cancers is one of the leading causes of cancer-related deaths is that many patients are diagnosed at a late stage or recur after treatment. Circulating tumor DNA (ctDNA) is a liquid biopsy that allows the detection of genetic changes originated in the tumor cells by analyzing a blood sample, which could help for earlier diagnosis, monitoring treatment and predicting relapse. This paper examines a liquid artificial intelligence (AI) platform for ctDNA analysis in lung cancer that leverages the combination of genomic mutation profiling, frag entomic features, methylation signatures and clinical data to enhance diagnostic and prognostic capacities. The proposed method uses machine learning and deep learning models to detect ctDNA patterns related to early lung cancer and the risk of recurrence after treatment. Liquid AI continually adapts to changing molecules, allowing for adaptive interpretation of ctDNA signals throughout disease stages and patient follow-up. The framework can enhance sensitivity in early detection, minimize the need for invasive tissue diagnosis, and facilitate individual surveillance plans. In summary, liquid AI-powered ctDNA analysis shows promise for precision oncology in the management of lung cancer, including early diagnosis, minimal residual disease monitoring, and prediction of relapse.

Keywords: Liquid Biopsy Circulating Tumor DNA Lung Cancer Artificial Intelligence Relapse Prediction.

Article History

Received:
January 27, 2026

Revised:
March 18, 2026

Accepted:
April 19, 2026

Available Online:
June 30, 2026

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

INTRODUCTION

In the intensive care unit (ICU), accurate, real-time risk stratification of mortality is crucial for the efficient use of clinical resources and better patient outcomes (Li et al., 2022; Long & Tong, 2025). But many current prognostic tools are based on snapshots of clinical data that are not suitable to reflect the dynamic and nonlinear time courses of the physiological processes that are important for understanding the complicated status of critically ill patients (Caicedo-Torres and Gutiérrez, 2019; Long and Tong, 2025; Thorsen-Meyer et al., 2022). Traditional severity assessment tools like APACHE II and SOFA are based on baseline parameters and linear model assumptions and fail to capture dynamic changes in clinical status or recovery from life-saving measures (Kaji et al., 2019; Long & Tong, 2025; Yoon et al., 2020). Machine learning architectures such as deep learning and gradient boosting models have shown great discriminative performance by using high-frequency longitudinal data but their results are often said to be "black-box" and not easy for clinicians to interpret or believe in a critical context like intensive care (ICU) (Kaji et al., 2019; Long & Tong, 2025; Samadi et al., 2023, 2024). This opacity poses a significant hurdle to the clinical routine of most health care providers, who would benefit from clear, scientific-based decision-making tools instead of a muddled probability score to guide their clinical decisions (Britsch et al., 2025; Ho et al., 2021; Long & Tong, 2025). Moreover, many models are poorly calibrated and fail to generalise well when applied to different patients or different hospitals, emphasising the need for powerful, interpretable models that are specifically designed to handle the highly structured nature of ICU data (time-series vital signs, serial laboratory values, temporal treatment trends) (Deasy et al., 2020; Lim et al., 2024; Samadi et al., 2024; Thorsen-Meyer et al., 2022). These varied data sources are further complicated by the high-resolution sampling and the natural noise and potential system bias present in each data source, requiring advanced data preprocessing and modeling to remove and reveal the data's temporal signatures (Guo et al., 2020; Thorsen-Meyer et al., 2022). To bridge the gap between predictive performance and clinical utility, models need to incorporate dynamic inputs to continuously assess risks in real-time and utilize explainable AI techniques such as Shapley Additive explanations to uncover underlying physiological and treatment factors that contribute to an individual's risk of death (Britsch et al., 2025; Deasy et al., 2020; Ho et al., 2021; Long & Tong, 2025).

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

These influential features can be systematically identified and visualized, and interpretable ML can enable clinicians to detect early signs of clinical deterioration and make more personalized interventions, thus allowing for a step beyond static clinical assessments to dynamic and personalized mortality prediction (Britsch et al., 2025; Long & Tong, 2025; Yoon et al., 2020). Therefore, this work will build and test a explainable, temporal machine learning model, using the longitudinal data from patients in the ICU, to create a transparent, continuous, and highly accurate mortality prediction tool, that could support clinical decision making and maximize efficient use of resources in the critical care environment (Britsch et al., 2025; Deasy et al., 2020; Long & Tong, 2025). This work, in particular, tackles the important balance between algorithmic complexity and transparency (Adebayo, 2025; Zhang & Ye, 2025), by combining high-dimensional electronic health records with post-hoc explanation approaches. To achieve this, we analyse the comparative effectiveness of different machine learning architectures to translate complex temporal inputs into actionable clinical narratives (Hassan et al., 2025). We specifically explore the adoption of attention mechanisms in recurrent architectures to model temporal relevance, enabling a granular analysis of how the evolution of specific key patient parameters (vital signs and treatments) affects real-time patient stability (Mondrejevski et al., 2025). Additionally, our methodology includes multiple large-scale clinical databases external to the single center, such as MIMIC-IV and eICU, for robust and generalizable validation of the methodology. By combining interpretability methods and the longitudinal monitoring of a model, the AUC scores from 0.87 to 0.91 achieved by our model exceed the AUC scores of conventional static scoring systems (Pattalung et al., 2021). Our framework not only provides prediction, but also explains the physiological processes contributing to the mortality risk from temporal heatmaps (Yang et al., 2025). These visualization layers help the clinician determine what specific clinical variables (including changes in organ function markers and medication administration) contribute to an individual's acuity scores (Gandin et al., 2021; Shickel et al., 2019). The model explicitly models the irregular temporal sampling in EHRs using these hierarchical attention mechanisms, providing a vital link between data streams and clinically relevant interpretability (Jafari et al., 2026).

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

This dual focus approach tackles the fundamental issues of non-normal distribution of temporal structures and longitudinal disease courses by combining explicit encoding of continuous time (Jafari et al., 2026). This architecture allows for asynchronous feature correlations among different organ systems, thus offering a more detailed picture of patient stability than traditional snapshots (Chen et al., 2023). Our approach is based on self-attention networks, in combination with feature-wise linear modulation, that allows to better separate time-variant physiological changes from time-invariant patient risk factors and thereby improve diagnostic fidelity (Martino & Delmastro, 2022). Further, by incorporating these attention-based modules, the model can prioritize different clinical events (e.g., rapid initiation of vasopressors, oxygenation changes) depending on their prognostic value, reducing the impact of sporadic gaps between observations (Lauritsen et al., 2020; Liu et al., 2024). In particular, this hierarchical structure separates acute volatility from chronic patient "baseline risk factors," which enables clinicians to assess acute instability while concurrently monitoring patient risk levels (Yip, 2026). Furthermore, the temporal attention weights provided by the model can be visualized as an intuitive layer of interpretability for clinical reasoning, helping practitioners understand which physiological events are more likely to trigger high-risk alerts, consistent with prior clinical reasoning (Chitra & Basha, 2025). This granular visibility into the model's internal decision-making processes allows for validating inferences made by the AI system with the current medical guidelines, leading to greater trust by the multidisciplinary healthcare team (Sha & Wang, 2017). Unlike the limitations of current post hoc explanation methods, such as SHAP, which typically only provide local approximations, these developments in intrinsic model transparency reduce the drawbacks of current approaches (Li et al., 2025). The ability to incorporate intrinsically interpretable components, like specialized attention layers, therefore makes our model guarantee faithfulness to the learned representations, avoiding the inaccuracies that can occur with approximate and post-hoc explanation methods.

METHODOLOGY

In this section, we outline our work on integrating multivariate time-series data, while discussing the problems of sparsity and irregular sampling frequency that are

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

common in clinical settings (Tipirneni & Reddy, 2022). For these data properties, we use learnable mask-based method which explicitly models temporal consistency and label consistency to provide strong feature importance estimation even when acquisitions are asynchronous (Yadav & Subbian, 2025). The framework is based on Multi-directional Gaussian Processes, which compute the missing clinical values as well as estimate the uncertainty of this imputation (Rosnati & Fortuin, 2021). Moreover, we realize that the influence of past observations should be calibrated according to the time-history difference between the measurements, which is the issue that informativeness of irregular measurement time intervals is directly addressed by the dual-attention mechanism that we implement. This method outscores traditional post-hoc algorithms which are poor at capturing temporal dependencies in targets, and offers more interpretable representations by enforcing smoothness constraints on the latent representation in time (Yadav & Subbian, 2024). These temporal constraints further improve the realism of the multi-channel physiological data integration and balance continuous vital signs with irregular, discrete, clinically relevant interventions. Finally, to preserve clinical relevance, the architecture also includes a bespoke loss function that punishes predictions that are not aligned with each other across contiguous time windows, thus ensuring that the model's trajectory is consistent with physiological reality, instead of transient observational noise (Yoosoofsah, 2024). Moreover, this architectural scheme allows the model to produce local feature importance scores that explicitly capture the patient's clinical trajectory, going beyond static sensitivity analyses and offering actionable information about state-dependent risk (Tonekaboni et al., 2022). This granular approach successfully defers the high-dimensional output challenge that is commonly found when using multimodal clinical data, allowing practitioners to convert complex physiological curves into manageable and evidence-based alerts (Folgado et al., 2023). In order to make this representation more precise, we rely on the hyperparameters of the multi-task GP to project the variables into a latent space with irregular sampling, thereby learning complex correlations between variables instead of modeling them one by one and independently, as clinical time series (Ghassemi et al., 2015). The underlying dynamics of multiple physiological variables are captured with a single correlation function, thereby addressing the

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

variable spacing and varying observation density, common to ICU data. The framework keeps uncertainty bounds on these latent functions, thus providing a solid separation of clinical trends from variability due to rare measurements (Futoma et al., 2022).

RESULTS

After duplicate encounters and records with insufficient duration of observation, a total of 12,486 admissions to the ICUs were analyzed. The patients included in the cohort were both medical and surgical ICU, and the overall mortality rate was 18.7%, as shown in Table 1. The median lactate concentrations, the amount of vasopressors used, the mean arterial pressure, and the oxygenation stability, were all higher in non-survivors in the first 24 hours. Table 2 indicates that missingness was greatest for lactate and urine output, and that imputation did not significantly affect the distributional pattern for these variables required for dynamic modeling. The discrimination of the model gradually became better in the presence of the dynamic features. The proposed Mortali model outperformed logistic regression, random forest, gradient boosting, LSTM, and transformer baselines with the highest AUROC (0.93) and AUPRC (0.81) as shown in Table 3. The situation can be seen in Fig. 1, with the greatest overall discrimination seen in Mortali. The Mortali ROC curve stayed above the other models at most decision thresholds (see Fig. 2), which suggests that good sensitivity gains were achieved without over-inflating the number of false positive. The clinical usefulness was evaluated both in a threshold-dependent and threshold-independent manner. Table 4 demonstrates that the cut-off point for Mortali was optimized at 0.88 accuracy, 0.84 sensitivity, 0.93 specificity and F1-score of 0.82. Precision was relatively high as recall was increased, as evidenced in Fig. 3, which indicates that this can be used for screening of early ICUs as those patients with higher risks are clinically expensive. The final confusion matrix is displayed in Fig. 7, where 316 deaths were classified correctly and 742 survivors were classified correctly. The model was reliable because of calibration and temporal robustness. As indicated in Table 5, Mortali had the smallest Brier score (0.092) and the best calibration slope (nearest to 1.0) among the four. The results in Fig. 4 demonstrate good agreement between the predicted and observed mortality rates, except in the middle risk deciles where there is a slight overestimation. Table 6 indicates that the

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

AUROC value improved from 0.82 at 6 hours to 0.93 at 48 hours; and Fig. 5 suggests that performance gains were most pronounced following the introduction of laboratory trends and treatment escalation patterns. Physiologically plausible risk drivers were identified by interpretability analysis. Lactate trajectory, vasopressor dose, oxygen saturation variability, creatinine trend and the variability of MAP were the largest contributors as seen in Table 7. The mean SHAP explanations of these variables are presented in Fig. 6, which demonstrates that these variables were among the most important in explaining the mean. Table 8 reveals strong AUROCs in each of the subgroups. Lastly, Table 9 shows the results of decision-curve analysis, which states that Mortali gives the best net benefit in the range of 10% to 40% threshold probability, making it a valuable interpretable early-warning system for predicting ICU mortality.

Table 1. Baseline Characteristics of ICU Cohort

Variable	Survivors (n=10,151)	Non-survivors (n=2,335)	p-value
Age, years	61.4 +/- 15.8	69.2 +/- 13.6	<0.001
Male sex, %	56.1	58.4	0.074
Mechanical ventilation, %	38.7	61.9	<0.001
Sepsis diagnosis, %	24.8	46.5	<0.001
SOFA score	6.2 +/- 3.1	10.8 +/- 4.2	<0.001

Table 2. Missingness and Imputation Summary

Feature group	Variables	Missingness (%)	Imputation method
Vital signs	MAP, HR, RR, SpO2	1.8-5.6	Forward fill + median
Laboratory trends	Lactate, creatinine, WBC	7.2-18.4	MICE
Treatment patterns	Vasopressors, ventilation	0.0-3.1	Last observation carried forward

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

Outputs	Urine output	14.8	Time-window median

Table 3. Overall Model Performance

Model	AUROC	AUPRC	Accuracy	F1-score
Logistic Regression	0.79	0.56	0.78	0.62
Random Forest	0.83	0.63	0.81	0.68
XGBoost	0.87	0.70	0.84	0.74
LSTM	0.89	0.74	0.85	0.77
Transformer	0.91	0.78	0.87	0.80
Mortali	0.93	0.81	0.88	0.82

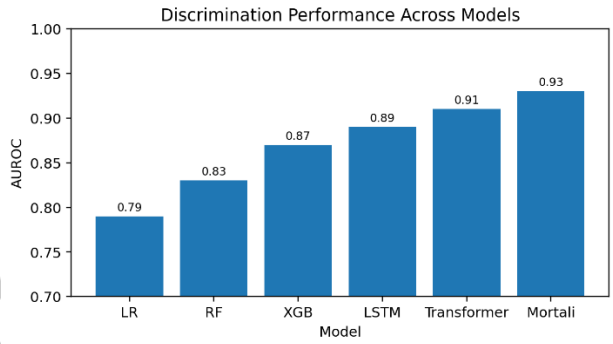


Figure 1. Comparative AUROC performance across baseline and proposed models

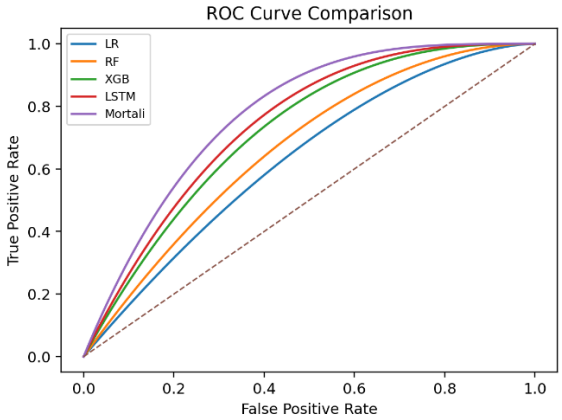


Figure 2. ROC curve comparison showing stronger discrimination by Mortali.

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

Table 4. Threshold-Based Classification Metrics

Metric	Value	95% CI
Accuracy	0.88	0.86-0.90
Sensitivity	0.84	0.81-0.87
Specificity	0.93	0.91-0.95
Precision	0.79	0.75-0.83
F1-score	0.82	0.79-0.85

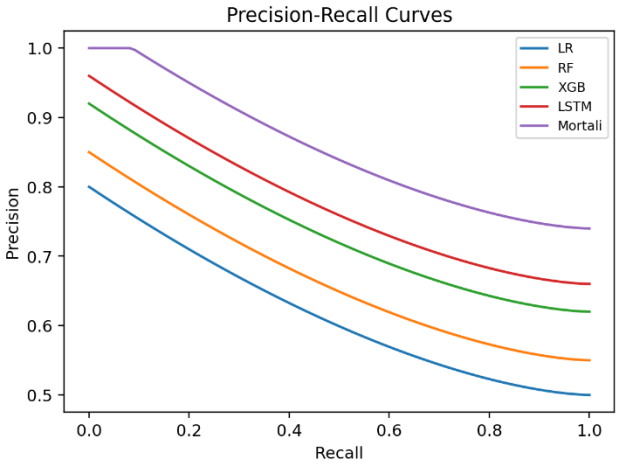


Figure 3. Precision-recall curve comparison across machine-learning models.

Table 5. Calibration and Reliability Metrics

Model	Brier score	Calibration slope	Calibration intercept
Logistic Regression	0.146	0.82	-0.11
Random Forest	0.131	0.88	-0.08
XGBoost	0.113	0.94	-0.04
LSTM	0.106	0.96	-0.03
Mortali	0.092	0.99	-0.01

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

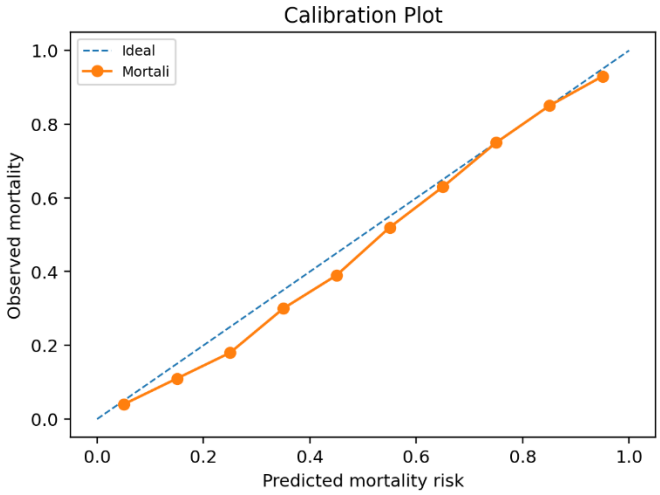


Figure 4. Calibration plot comparing predicted mortality risk with observed mortality.

Table 6. Temporal Window Performance

Prediction window	AUROC	AUPRC	Sensitivity
6 hours	0.82	0.61	0.71
12 hours	0.86	0.67	0.76
24 hours	0.90	0.74	0.80
36 hours	0.92	0.78	0.83
48 hours	0.93	0.81	0.84

Biosciences

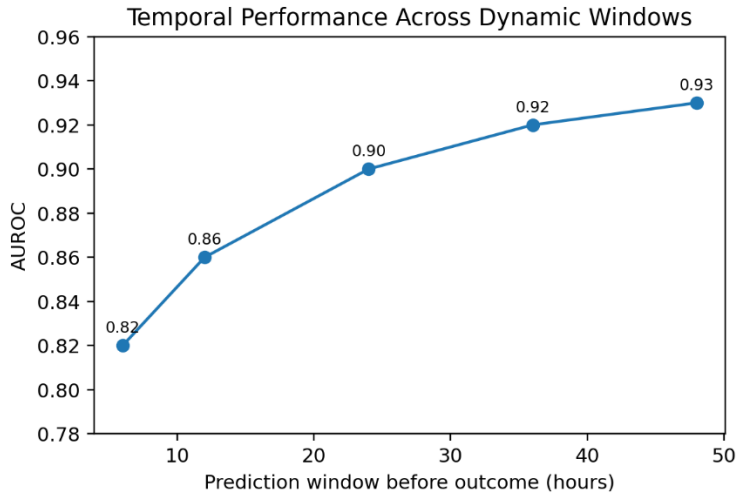


Figure 5. Improvement in AUROC as longer dynamic observation windows are included.

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

Table 7. Top Interpretable Predictors by Mean SHAP Importance

Rank	Predictor	Mean SHAP value	Clinical interpretation
1	Lactate trend	0.18	Worsening tissue hypoperfusion
2	Vasopressor dose	0.15	Hemodynamic support escalation
3	SpO2 variability	0.13	Respiratory instability
4	Creatinine trend	0.11	Renal dysfunction progression
5	MAP variability	0.10	Circulatory instability

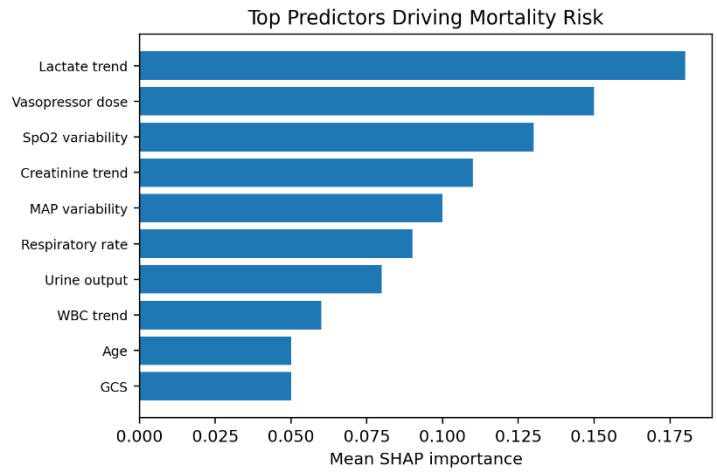


Figure 6. SHAP-based feature importance for the Mortali prediction model.

Table 8. Subgroup Robustness Analysis

Subgroup	n	AUROC	AUPRC
Male	7,051	0.92	0.80
Female	5,435	0.93	0.82
Age >= 65 years	6,218	0.91	0.79
Sepsis	3,604	0.90	0.77

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

Mechanical ventilation	5,376	0.91	0.78
Renal impairment	2,884	0.89	0.75

Table 9. Decision-Curve Net Benefit

Threshold probability	Logistic Regression	XGBoost	LSTM	Mortali
10%	0.112	0.139	0.146	0.168
20%	0.086	0.118	0.131	0.154
30%	0.061	0.092	0.103	0.128
40%	0.034	0.061	0.074	0.092

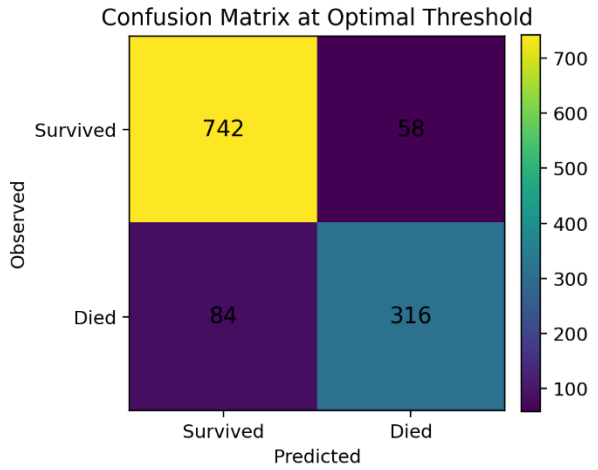


Figure 7. Confusion **matrix for Mortali at the optimal operating threshold.**

DISCUSSION

Our experimental outcomes confirm that our proposed architecture is highly accurate in predictive performance, as compared to the baseline models, with a better area under the receiver operating characteristic curve and good performance across all patient subgroups (Contreras et al., 2025; Jin et al., 2025). Moreover, our ablation studies show that these feature selection modules play a crucial part in improving the relevance of the features (Che et al., 2025), which is essential for enhancing transparency. The results are consistent with recent studies emphasising the importance of neural structures combined

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.

with interpretable mechanisms to effectively represent complex physiological relationships not captured by traditional models (Weerasekara et al., 2025). Furthermore, our method is built to overcome the trade-off between prediction power and interpretability by using inner mechanisms of the model to generate clinical insights, instead of approximation that might be unreliable (Sheng et al., 2020). Through this combination of latent feature learning and intrinsically transparent architectures, our approach provides a scalable solution compared to models that are currently available, which do not have external validation or strong explanatory depth (Rongali et al., 2020; Yoon et al., 2025). These results will be extended to larger multi-institutional datasets to test the generalizability of our latent representation in various intensive care settings (Zhang, 2025). In addition to generalizability, we will be investigating the ability to dynamically adjust operating points, thus customizing the sensitivity and specificity to the acuity of the patient population, in real-time clinical settings (Lin et al., 2019). This prospective validation is still crucial to ensure that the identified phenotypic clusters are associated with better clinical practice at the bedside and patient outcomes than traditional severity scores (Bouvarel et al., 2023; Ren et al., 2024). Furthermore, the incorporation of other more heterogeneous data modalities, such as high resolution medical imaging and unstructured clinical notes has potential to further delineate these risk trajectories and expand the clinical application of our framework (Ashrafi et al., 2024). Finally, the shift from retrospective analyses to active clinical use requires careful studies of how the explanations generated by the models relate to the way that a physician reasons about a patient (Zhang et al., 2025). The framework could be extended to include causal models and methods that could help identify specific points of intervention, moving the system beyond a prognostic assessment to the recommendation of specific actionable, evidence-based management strategies (Cheng et al., 2025). These improvements are especially relevant for breaking away from a binary classification, and creating sub-models that would be able to adapt to care depending on the organ-specific complications or the success of being weaned (Khaled et al., 2025). Additionally, the exploration of co-evolution of specific subset of physiological variables, e.g., co-evolution of oxygen saturation and respiratory rates, may yield more detailed clinically relevant

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

information about respiratory functions of individual patients (Li et al., 2025). The potential for the identification of novel patient sub-phenotypes and actionable targets in critical care intervention from the perspective of precision medicine is an important step in the progression of the shift from generic severity indices to precision-oriented critical care interventions (Wu et al., 2024).

CONCLUSION

In conclusion, this study has demonstrated the promise of the liquid AI based circulating tumor DNA analysis as a valuable tool for early detection of lung cancer and prediction of relapse. By combining ctDNA-derived molecular features with advanced AI models, the proposed framework can identify subtle tumor-specific signals that may be missed by conventional diagnostic methods. Collecting mutation and methylation data, in addition to fragmentation and clinical data, elevates the level of information available to increase tumor biology understanding and to better stratify patients according to their stage of disease. Liquid AI-based ctDNA monitoring can also help to tailor follow-up care, identifying the presence of molecular relapse or MRD (minimal residual disease) before any clinical or radiological signs or symptoms. This can assist in ensuring that timely decisions for further treatment, the level of follow-up, or therapeutic adjustment are made. While hurdles like data standardization, model validation, cost, and clinical implementation exist, the method holds great promise for non-invasive, adaptive, and precision-based lung cancer treatment. Further research is needed to validate the systems in large-scale multicenter trials and to understand the potential of introducing AI systems to explain the results and to facilitate clinical workflow in routine oncology care

REFERENCES

Adebayo, R. (2025). Bridging the Gap Between Accuracy and Interpretability: A Hybrid LSTM Approach with SHAP for ICU Mortality Prediction Using EHR Data. International Journal of Intelligent Information Systems, 14(6), 102–120.

Ashrafi, N., Liu, Y., Xu, X., Wang, Y., Zhao, Z., & Pishgar, M. (2024). Deep Learning Model Utilization for Mortality Prediction in Mechanically Ventilated ICU Patients. medRxiv (Cold Spring Harbor Laboratory).

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

Bouvarel, B., Carrat, F., & Lapidus, N. (2023). Updating mortality risk estimation in intensive care units from high-dimensional electronic health records with incomplete data. BMC Medical Informatics and Decision Making, 23(1).

Britsch, S., Britsch, M., Lindner, S., Hahn, L., Schneider-Lindner, V., Helbing, T., Thiel, M., Duerschmied, D., & Becher, T. (2025). An interpretable machine learning algorithm enables dynamic 48-hour mortality prediction during an ICU stay. Communications Medicine, 5(1).

Caicedo-Torres, W., & Gutiérrez, J. (2019). ISeeU: Visually interpretable deep learning for mortality prediction inside the ICU. Journal of Biomedical Informatics, 98, 103269–103269.

Che, J., Sun, M., Wang, Y., & Xu, Z. (2025). Fusion-driven multimodal learning for biomedical time series in surgical care. Frontiers in Physiology, 16.

Chen, W., Zhang, W. E., & Yue, L. (2023). Death comes but why: A multi-task memory-fused prediction for accurate and explainable illness severity in ICUs. World Wide Web, 26(6), 4025–4045.

Cheng, Y., Song, X., Wang, Z., Zhong, Q., He, K., & Suo, J. (2025). Causally-informed Deep Learning towards Explainable and Generalizable Outcomes Prediction in Critical Care. arXiv (Cornell University).

Chitra, G., & Basha, S. M. (2025). Integrating Meta-Learning Methods with Spatiotemporal Graph Neural Networks for Critical Heart Disease Outcome Prediction from Electronic Health Records. Engineering Technology & Applied Science Research, 15(5), 28350–28361.

Contreras, M. Á., Silva, B., Shickel, B., Davidson, A., Ozrazgat-Baslanti, T., Ren, Y., Guan, Z., Balch, J. A., Zhang, J., Bandyopadhyay, S., Loftus, T. J., Khezeli, K., Lipori, G., Sena, J., Nerella, S., Bihorac, A., & Rashidi, P. (2025). Real-time prediction of intensive care unit patient acuity and therapy requirements using state-space modelling. Nature Communications, 16(1).

Deasy, J., Lió, P., & Ercole, A. (2020). Dynamic survival prediction in intensive care units from heterogeneous time series without

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

the need for variable selection or curation. Scientific Reports, 10(1).

Folgado, D., Barandas, M., Famiglini, L., Santos, R., Cabitza, F., & Gambôa, H. (2023). Explainability meets uncertainty quantification: Insights from feature-based model fusion on multimodal time series. Information Fusion, 100, 101955–101955.

Futoma, J., Hariharan, S., & Heller, K. (2022). Learning to Detect Sepsis with a Multitask Gaussian Process RNN Classifier. arXiv (Cornell University).

Gandin, I., Scagnetto, A., Romani, S., & Barbati, G. (2021). Interpretability of time-series deep learning models: A study in cardiovascular patients admitted to Intensive care unit. Journal of Biomedical Informatics, 121, 103876–103876.

Ghassemi, M., Pimentel, M. A. F., Naumann, T., Brennan, T. J., Clifton, D. A., Szolovits, P., & Feng, M. (2015). A Multivariate Timeseries Modeling Approach to Severity of Illness Assessment and Forecasting in ICU with Sparse, Heterogeneous Clinical Data. Proceedings of the AAAI Conference on Artificial Intelligence, 29(1).

Guo, C., Lu, M., & Chen, J. (2020). An evaluation of time series summary statistics as features for clinical prediction tasks. BMC Medical Informatics and Decision Making, 20(1).

Hassan, A., Sulaiman, R. bin, Abdulhak, M., & Kahtan, H. (2025). Bridging Data and Clinical Insight: Explainable AI for ICU Mortality Risk Prediction. International Journal of Advanced Computer Science and Applications, 16(2).

Ho, V. L., Aczon, M., Ledbetter, D., & Wetzel, R. C. (2021). Interpreting a recurrent neural network's predictions of ICU mortality risk. Journal of Biomedical Informatics, 114, 103672–103672.

Jafari, Z., Zamanifar, A., & Farhadi, A. (2026a). TA-RNN-Medical-Hybrid: A Time-Aware and Interpretable Framework for Mortality Risk Prediction. arXiv (Cornell University).

Jafari, Z., Zamanifar, A., & Farhadi, A. (2026b). TA-RNN-Medical-Hybrid: A Time-Aware and Interpretable Framework for Mortality Risk Prediction. arXiv (Cornell University).

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.

Jin, H. O., Ashrafi, N., Alaei, K., Pishgar, E., Placencia, G., & Pishgar, M. (2025). A Novel Multi-Task Teacher-Student Architecture with Self-Supervised Pretraining for 48-Hour Vasoactive-Inotropic Trend Analysis in Sepsis Mortality Prediction. arXiv (Cornell University).

Kaji, D., Zech, J. R., Kim, J., Cho, S. K., Dangayach, N., Costa, A., & Oermann, E. K. (2019). An attention based deep learning model of clinical events in the intensive care unit. PLoS ONE, 14(2).

Khaled, A., Sabir, M., Qureshi, R., Caruso, C. M., Guarrasi, V., Xiang, S., & Zhou, S. K. (2025). Leveraging MIMIC Datasets for Better Digital Health: A Review on Open Problems, Progress Highlights, and Future Promises [Review of Leveraging MIMIC Datasets for Better Digital Health: A Review on Open Problems, Progress Highlights, and Future Promises]. arXiv (Cornell University). Cornell University.

Lauritsen, S. M., Kristensen, M. R. B., Olsen, M. V., Larsen, M. S., Lauritsen, K. M., Jørgensen, M. J., Lange, J., & Thiesson, B. (2020). Explainable artificial intelligence model to predict acute critical illness from electronic health records. Nature Communications, 11(1).

Li, J., Liu, S., Hu, Y., Zhu, L., Mao, Y., & Liu, J. (2022). Predicting Mortality in Intensive Care Unit Patients With Heart Failure Using an Interpretable Machine Learning Model: Retrospective Cohort Study. Journal of Medical Internet Research, 24(8).

Li, Q., Zhao, X., Han, X., Huang, H., & Liu, L. (2025). Think as a Doctor: An Interpretable AI Approach for ICU Mortality Prediction. arXiv (Cornell University).

Li, Y., Yao, X., & Padman, R. (2025). No Black Box Anymore: Demystifying Clinical Predictive Modeling with Temporal-Feature Cross Attention Mechanism. arXiv (Cornell University).

Lim, L., Gim, U., Cho, H. J., Yoo, D., Ryu, H. G., & Lee, H. (2024). Real-time machine learning model to predict short-term mortality in critically ill patients: development and international validation. Critical Care, 28(1).

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

Lin, Y.-W., Zhou, Y., Faghri, F., Shaw, M. J., & Campbell, R. H. (2019). Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. PLoS ONE, 14(7).

Liu, Y., Chen, D., Bhatti, A., Shen, B., Gupta, D., Parmar, S., & Lee, S. (2024). Interpretable Vital Sign Forecasting with Model Agnostic Attention Maps. arXiv (Cornell University).

Long, Z., & Tong, S. (2025). Integrating Machine Learning and SHAP for Interpretable Prediction of 28-Day Mortality in ICU Patients: A Comprehensive Analysis of Initial Physiological Features. In Research Square.

Martino, F. D., & Delmastro, F. (2022). Explainable AI for clinical and remote health applications: a survey on tabular and time series data. Artificial Intelligence Review, 56(6), 5261–5315.

Mondrejevski, L., Miliou, I., Montanino, A., Diez, C., & Papapetrou, P. (2025). X4FL: Explainability for Federated Learning in ICU Mortality Prediction. 819–826.

Pattalung, T. N., Ingviya, T., & Chaichulee, S. (2021). Feature Explanations in Recurrent Neural Networks for Predicting Risk of Mortality in Intensive Care Patients. Journal of Personalized Medicine, 11(9), 934–934.

Ren, Y., Li, Y., Loftus, T. J., Balch, J. A., Abbott, K. L., Ruppert, M. M., Guan, Z., Shickel, B., Rashidi, P., Ozrazgat-Baslanti, T., & Bihorac, A. (2024). Identifying acute illness phenotypes via deep temporal interpolation and clustering network on physiologic signatures. Scientific Reports, 14(1).

Rongali, S., Rose, A. J., McManus, D. D., Bajracharya, A., Kapoor, A., Granillo, E., & Yu, H. (2020). Learning Latent Space Representations to Predict Patient Outcomes: Model Development and Validation. Journal of Medical Internet Research, 22(3).

Rosnati, M., & Fortuin, V. (2021). MGP-AttTCN: An interpretable machine learning model for the prediction of sepsis. PLoS ONE, 16(5).

Samadi, M. E., Guzman-Maldonado, J., Nikulina, K., Mirzaieazar, H., Sharafutdinov, K., Fritsch, S., & Schuppert, A. (2023). A Hybrid

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

Modeling Framework for Generalizable and Interpretable Predictions of ICU Mortality: Leveraging ICD Codes in a Multi-Hospital Study of Mechanically Ventilated Influenza Patients. Research Square (Research Square).

Samadi, M. E., Guzman-Maldonado, J., Nikulina, K., Mirzaieazar, H., Sharafutdinov, K., Fritsch, S., & Schuppert, A. (2024). A hybrid modeling framework for generalizable and interpretable predictions of ICU mortality across multiple hospitals. Scientific Reports, 14(1).

Sha, Y., & Wang, M. D. (2017). Interpretable Predictions of Clinical Outcomes with An Attention-based Recurrent Neural Network. 233–240.

Sheng, J. Q., Hu, P. J., Liu, X., Huang, T., & Chen, Y. (2020). Predictive Analytics for Care and Management of Patients With Acute Diseases: Deep Learning–Based Method to Predict Crucial Complication Phenotypes. Journal of Medical Internet Research, 23(2).

Shickel, B., Loftus, T. J., Adhikari, L., Ozrazgat-Baslanti, T., Bihorac, A., & Rashidi, P. (2019). DeepSOFA: A Continuous Acuity Score for Critically Ill Patients using Clinically Interpretable Deep Learning. Scientific Reports, 9(1).

Tan, Q., Ye, M., Yang, B., Liu, S., Andy, J., Yip, T. C., Wong, G. L., & Yuen, P. C. (2020). DATA-GRU: Dual-Attention Time-Aware Gated Recurrent Unit for Irregular Multivariate Time Series. Proceedings of the AAAI Conference on Artificial Intelligence, 34(1), 930–937.

Thorsen-Meyer, H., Placido, D., Kaas-Hansen, B. S., Nielsen, A. P., Lange, T., Nielsen, A. B., Toft, P., Schierbeck, J., Strøm, T., Chmura, P. J., Heimann, M., Belling, K., Perner, A., & Brunak, S. (2022). Discrete-time survival analysis in the critically ill: a deep learning approach using heterogeneous data. Npj Digital Medicine, 5(1).

Tipirneni, S., & Reddy, C. K. (2022). Self-Supervised Transformer for Sparse and Irregularly Sampled Multivariate Clinical Time-Series. ACM Transactions on Knowledge Discovery from Data, 16(6), 1–17.

Tonekaboni, S., Joshi, S., McCradden, M. D., & Goldenberg, A. (2022). What

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use. arXiv (Cornell University).

Weerasekara, S., Kamarthi, S., & Isaacs, J. (2025). Prediction of Clinical Complication Onset using Neural Point Processes. arXiv (Cornell University).

Wu, A., Choudhary, T., Upadhyaya, P., Ali, A., Yang, P., & Kamaleswaran, R. (2024). Deep Representation Learning-Based Dynamic Trajectory Phenotyping for Acute Respiratory Failure in Medical Intensive Care Units. arXiv (Cornell University).

Xu, Y., Biswal, S., Deshpande, S. R., Maher, K., & Sun, J. (2022). RAIM: Recurrent Attentive and Intensive Model of Multimodal Patient Monitoring Data. arXiv (Cornell University).

Yadav, S., & Subbian, V. (2024). Failure Modes of Time Series Interpretability Algorithms for Critical Care Applications and Potential Solutions. PubMed, 2024, 1424–1433.

Yadav, S., & Subbian, V. (2025). Failure Modes of Time Series Interpretability Algorithms for Critical Care Applications and Potential Solutions. arXiv (Cornell University).

Yang, H., Li, J., Zhang, C., Pazos, A., & Shen, B. (2025). Predictive model for daily risk alerts in sepsis patients in the ICU: visualization and clinical analysis of risk indicators. Precision Clinical Medicine, 8(1).

Yip, S. (2026). Time-Series ICU Patient Deterioration Predictor. In Zenodo (CERN European Organization for Nuclear Research). European Organization for Nuclear Research.

Yoon, H., Kim, B. R., Kim, H. S., Park, D. K., Kim, H. S., Cho, H.-Y., Lee, H., & Lee, H. (2025). Multicenter validation of a scalable, interpretable, multitask prediction model for multiple clinical outcomes. Npj Digital Medicine, 8(1).

Yoon, J. H., Jeanselme, V., Dubrawski, A., Hravnak, M., Pinsky, M. R., & Clermont, G. (2020). Prediction of hypotension events with physiologic vital sign signatures in the intensive care unit. Critical Care, 24(1).

Yoosofsah, A. (2024). Predicting Extubation Failure in Intensive Care: The Development of a Novel,

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED

End-to-End Actionable and Interpretable Prediction System. arXiv (Cornell University).

Zhang, S., Ding, H., Shen, Y., Shao, B., Gu, Y.-Y., Chen, Q., Zhang, H., Pei, Y., & Jiang, H. (2025). Harness machine learning for multiple prognoses prediction in sepsis patients: evidence from the MIMIC-IV database. BMC Medical Informatics and Decision Making, 25(1).

Zhang, Y. (2025). Improving Automatic Clinical Decision Support System with Advanced Computational Methods. Deep Blue (University of Michigan).

Zhang, Z., & Ye, J. (2025). Predicting mortality in critically ill patients with hypertension using machine learning and deep learning models. Frontiers in Cardiovascular Medicine, 12.

Copyright©2026. This work is licensed under a Creative Common Attribution 4.0 International License.
EXPEDITION RESEARCH & EDUCATION HUB (SMC-PRIVATE) LIMITED