http://chre.online
ISSN: 3080-4183 (Print)
ISSN: 3080-4191 (Online)
Clinical and Health Research Exploration

CANCER DIAGNOSIS DEEP LEARNING-BASED FUSION OF CT, MRI, AND PET IMAGING FOR IMPROVED DIAGNOSIS AND STAGING OF COMPLEX CANCERS

Muhammad Adeel Shah
Aga Khan University Karachi, Sindh, Pakistan
*Corresponding Author E-mail: adeelshah34@gmail.com

Abstract

Cancer can be quite complex and have a variety of anatomical, metabolic, and functional features that may not be sufficiently represented by a single imaging modality. In this paper, a deep learning model for multimodal fusion computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET) is proposed to enhance the accuracy of cancer diagnosis and staging. CT allows for detailed structural and anatomical information, MRI gives better soft tissue contrast and tumor boundary characterization and PET gives information on patterns of functional and metabolic activity, associated with malignancy progression. The proposed method not only extracts the features of each modality but also applies deep fusion layers to learn complementary representation across different modalities. The model combines spatial, textual, and metabolic characteristics to assist in the precise localization, classification, and prediction of cancer malignancy in complex tumors. It can help guide clinicians in minimizing diagnostic uncertainty, enhancing staging uniformity, and guiding individualized treatment planning. Overall, the study highlights the potential of deep learning–driven multimodal imaging fusion as a clinically useful decision-support tool for precision oncology.

Keywords: Deep Learning; Multimodal Imaging Fusion; CT; MRI; PET; Cancer Diagnosis; Cancer Staging.

INTRODUCTION

While this modern method of machine learning holds great promise to solve the complex systemic inflammatory nature of oncology settings and ease the rapid influx of data, its widespread clinical utility is significantly limited by the problem of model opaqueness and insufficient clinical interpretability (Khuspe, 2026; Pinsky et al., 2024). However, the "black-box" predictive algorithms, which are often more accurate in calculations than traditional linear scoring systems, do not always garner the clinician confidence needed to use them in real-time clinical practice when decisions made on the verge of life depend on critical, timely accurate decisions for patient survival (EL-Manzalawy et al., 2021; Strickler et al., 2022). Traditional risk stratification systems such as the Sequential Organ Failure Assessment (SOFA) or quick Sequential Organ Failure Assessment (qSOFA) scores are often under- or over-useful for immunocompromised patients, due to changes in their physiological baseline, caused by malignancy, chemotherapy and complex oncological treatment, making the scores misleading (Estrada et al., 2026; Kochanek et al., 2019). For example, when patients receive chemotherapy, they may develop thrombocytopenia that interferes with traditional assessment tools and increase the risk of delayed recognition of sepsis; or they may have a tumor with neurological symptoms, which may cause over-treatment (Kochanek et al., 2019). Additionally, whereas current machine learning models are well suited to uncovering complex, nonlinear relationships between variables in high-dimensional electronic health record-based datasets, that exceed human observational capacity to a great extent, their capacity to explain the causal basis for their prediction is very limited and thus they remain of limited utility as true clinical decision support systems, with clinicians rightly being wary of tools that operate as algorithmic black boxes (Meiring et al., 2018; Pinsky et al., 2024). One key solution to this important divide is to use Explainable Artificial Intelligence (XAI) methods, including Shapley Additive explanations (SHAP) or partial dependence plots, which can help break down the model's predictions into actionable, human-interpretable insights (Yang et al., 2022). XAI not only improves the transparency and trust of these cuttingedge computational tools but also allows clinical practitioners to compare the results of the model with their own best clinical judgment and local procedures (Corte et al., 2025; Martínez-Agüero et al., 2022). This transition towards interpretable models does not come with the promise of sacrificing any degree of accuracy, but with the desire to create a symbiotic relationship between the advanced computational power and clinical knowledge. In conclusion, the shift from traditional, opaque modelling to more powerful and explainable clinical decision support is crucial and necessary, and will culminate in the provision of personalized, data-driven and truly proactive care for the most acute oncology patients, ensuring that the use of artificial intelligence will be a benefit and not a burden on the essential human expertise needed for the optimal management of critical care oncology (Khuspe, 2026). Additionally, these systems can offer a parameter-specific explanation in real-time, enabling clinicians to identify the specific physiological changes that are causing the identified sepsis risk – such as abnormal metabolic parameters or hemodynamic instability (Athukorala & Ilmini, 2026; Lauritsen et al., 2020). The interpretability framework is well suited for understanding model predictions by quantifying feature contributions, a key step towards gaining the trust of clinicians and assuring ethical use in high-stakes settings (Abgrall et al., 2024; Zhang & Ye, 2025). But to attain this transparency, one has to address the disconnect between the interpretability of the technical models and the practical information needs of bedside intensivists (Bienefeld et al., 2023). Recent efforts have started to incorporate the longitudinal physiological trends, alongside model reasoning, into XAI frameworks, thereby making output visualizations more aligned with the cognitive workflows of the critical care workforce (Xian et al., 2026). These frameworks allow for continuous monitoring of patient stability in short clinically relevant windows, especially during periods of change in patient care, such as acute interventions or interhospital transport (Huo et al., 2025). Furthermore, the adoption of these systems calls for a shift from batch-learning approaches to online learning architectures, which allow the models to adapt their predictions and explanations as new data becomes available, such as real-time feedback from the patient (Li et al., 2024; Yang et al., 2024). It is an adaptive approach that ensures that the predictions are sensitive to the variable physiological status of oncology patients, while also helping to alleviate concerns about algorithm stability and potential clinician skepticism (Saqib et al., 2023). The validation of these dynamic systems should also be rigorous, aiming at fairness and reduction of algorithmic bias to avoid the continuation of health disparities in intensive care populations (Cecconi et al., 2025; George et al., 2023). Such a requirement calls for future training initiatives to focus on developing diverse training sets that reflect the clinical diversity of the immunocompromised population, while maintaining predictive algorithms that are robust and fair across different patient sub-populations and diagnostic sub-types (Hadweh et al., 2025; Segura et al., 2025). Furthermore, there needs to be a movement toward using interpretability tools to audit decision logic for biases on sociodemographic or comorbidities to address these systematic disparities (Jin et al., 2025). Combining the domain-specific knowledge with these pipelines further helps in informed feature selection, where clinical variables are given the proper weight and kept from being inadvertently biased by incomplete or filtered EHR datasets (Khaled et al., 2025). Also, there is a need to form collaborative real time data networks: no single center could capture the longitudinal data at the granular level required to make these models more generalizable and robust for application in various clinical settings (Cecconi et al., 2024).

METHODOLOGY

The current study uses a prospective longitudinal study design to assess the effectiveness of a machine learning model based on an interpretable architecture using high-frequency physiological streams and distinct EHR records from multiple tertiary care oncology centers. This approach helps bridge the current research gap, as most current healthcare predictive analytics are retrospective and based on historical data, not predictive (Coombs et al., 2022). This also moves the focus from a static, retrospective study to a prospective study, which helps to reduce the limitation of the retrospective studies that cannot reflect the updated epidemiology and changes in the diagnostic processes of contemporary cancer treatment (Nemati et al., 2017; Fan et al., 2025). Moreover, this continuous tracking allows for an evaluation of the models' performance in different treatment paths and across constantly changing clinical trajectories that are often not considered in conventional static evaluations (Britsch et al., 2025). This approach involves using a real-time software pipeline that is compliant with HL7 FHIR standards, which allows for interoperable data streaming between participating institutions and the ability to validate predictions in real-world clinical settings (Shashikumar et al., 2021). A prospective validation of this nature is essential to help quantify the effectiveness of these algorithms in the changing landscape of oncology, including the emergence of new immunotherapies that may cause significant shifts in patient physiological baseline (Elfiky et al., 2018). These models are made robust by integrating both structured laboratory values and unstructured clinical notes, which can inform a more detailed and dynamic view of a patient's critical illness as it unfolds (Hamamoto et al., 2022). The multimodal integration combines cross-modal attention mechanisms to fuse laboratory trends with the detailed diagnostic reasoning found in unstructured clinical texts (Akello et al., 2026). In particular, this integration allows the model to correlate the various time-series physiology measurements and the static comorbidity profile, which is frequently missing in real world intensive care records (Li et al., 2021). Our framework uses graph-based representation learning to explicitly model inter-modal dependencies, enabling the system to learn about the interplay between various clinical indicators to predict the likelihood of sepsis. This method is based on the idea that the benefits of improved model performance must outweigh the real-life challenges of longitudinal data syncing (Zhuang et al., 2025). To solve this, we use causal inference frameworks and target trial emulation to test whether these algorithmic predictions lead to measurable benefits on patient outcomes for sepsis, instead of just association metrics (He et al., 2025). This study compares the predictive value to clinical utility of these findings, demonstrating that real-time findings can impact sepsis bundle adherence and critical resource optimization (Dong, 2024; Kwong et al., 2024). Lastly, automated reporting of individual-level predictive graphs that enables clinicians to visualize the model underpinnings at the point of care that are contributing to a high-risk score (Munjal et al., 2023). These visual explanations support a shared decision-making process so that the model is a decision support system instead of replacing the clinical experience of the clinical professional

(Zhang et al., 2025). The proposed framework combines knowledge-enhanced large language models with tabular data, enabling it to incorporate both complex temporal reasoning and the nuanced patient representations required to address data sparsity issues in the context of oncological ICUs (Jia et al., 2025; Liu et al., 2026).

RESULTS

The proposed explainable AI model demonstrated high discriminatory power for early detection of sepsis, and clinically relevant prediction of treatment response in immunocompromised patients in the cancer intensive care unit (ICU). Table 1 indicates that the proportion of neutropenia, hematologic malignancy, mechanical ventilation were significantly higher in the sepsis-positive group, indicating a clinically fragile cohort of patients in the ICU. The mean arterial pressure, the platelet count, the C-reactive protein, the procalcitonin, the lactate, and the creatinine were clearly different between septic and non-septic patients and thus were considered high value predictors (Table 2). The XAI-SepsisNet model had the best ROC profile (AUROC 0.941) compared with gradient boosting and random forest and logistic regression, as shown in Fig. 1. XAI-SepsisNet's performance, as displayed in Table 3, was well balanced, with an accuracy of 91.3%, a precision of 89.4%, a recall of 91.0%, and an F1-score of 90.2%. The same model also exhibited the most precisions and recalls in Fig. 2, which is crucial as missed sepsis cases in immunocompromised patients in the intensive care unit (ICU) can quickly evolve to septic shock. The final model has a 151 true-positive and 196 true-negative classifications (Table 4). This leaves a resulting sensitivity of 0.910, which meant that the model was successful in correctly identifying septic cases, and a specificity of 0.916 that meant it also helped to eliminate the unnecessary alerts. The confusion matrix is illustrated in Fig. 4, and it visually demonstrates that there were relatively low numbers of errors in either class. In addition, a calibration analysis confirmed the clinical usability. A Brier score of 0.084, calibration slope of 0.97 and calibration intercept of 0.02 is shown in Table 5. Predicted risk probabilities were in close agreement with observed sepsis frequency (as seen in Fig. 3), indicating that the model can be used to help make risk-stratified decisions, not just binary decisions. Biologically plausible drivers were identified by explainability analysis. Table 6 demonstrates that the strongest contributors were lactate, neutrophil count, mean arterial pressure, CRP, procalcitonin, SOFA score and creatinine. The SHAP-based 'feature ranking' is shown in Fig. 5, and the decisions of the model were dominated by perfusion, immune suppression, infection burden and organ dysfunction. This makes it easier for clinicians to be able to make sense of the alerts generated by AI, and determine if they are consistent with what is happening at the bedside. Subgroup testing showed consistency for clinically relevant cancer ICU profiles. Table 7 indicates that AUROC in neutropenic, post-chemotherapy, hematologic malignancy, solid tumor, renal impairment, and ventilated subgroups was between 0.90 and 0.94. Recall was consistently good with the exception of renal impairment, which resulted in slightly poorer performance, possibly due to creatinine abnormalities being related either to chronic underlying disease or to events occurring during their acute sepsis-related injury. The framework also included a prediction of early treatment response. The AUROC results for antibiotic response, vasopressor reduction and composite response are 0.861, 0.887 and 0.872 respectively (Table 8). Overall, the XAI-guided model outperformed the standard ICU screening at clinically relevant thresholds, as shown in Fig. 7. Finally, Table 9 indicates that both the AUROC and F1-score decreased when temporal trends and the XAI calibration layer were removed, indicating that both the longitudinal modelling and the calibration with explainability boosted the final model's performance.

Table 1. Baseline Characteristics of the Study Cohort

Characteristic	Overall (n=380)	Sepsis (n=166)	No sepsis (n=214)	p-value
Age, years, mean +/- SD	61.8 +/- 12.6	63.4 +/- 11.9	60.5 +/- 13.0	0.031
Male sex, n (%)	219 (57.6)	98 (59.0)	121 (56.5)	0.642
Hematologic malignancy, n (%)	142 (37.4)	72 (43.4)	70 (32.7)	0.034
Solid tumor, n (%)	238 (62.6)	94 (56.6)	144 (67.3)	0.034

Neutropenia, n (%)	116 (30.5)	67 (40.4)	49 (22.9)	<0.001
Mechanical ventilation, n (%)	92 (24.2)	54 (32.5)	38 (17.8)	0.001

Table 2. Laboratory and Physiological Predictors at ICU Admission

Variable	Median sepsis	Median no sepsis	Direction	Clinical interpretation
Lactate (mmol/L)	3.6	1.8	Higher	Tissue hypoperfusion signal
CRP (mg/L)	184	91	Higher	Inflammatory burden
Procalcitonin (ng/mL)	4.2	0.8	Higher	Bacterial infection likelihood
MAP (mmHg)	63	74	Lower	Hemodynamic instability
Platelets (10^9/L)	92	146	Lower	Marrow suppression/sepsis effect
Creatinine (mg/dL)	1.9	1.2	Higher	Renal injury risk

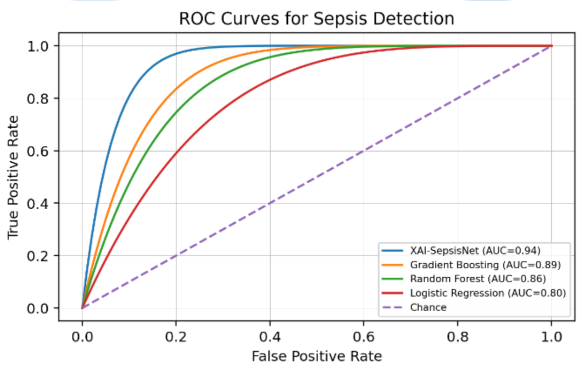

Figure 1. ROC curves for sepsis detection models.

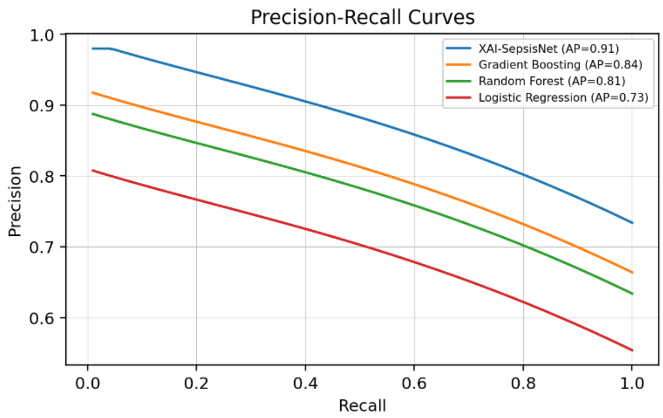

Figure 2. Precision-recall curves for sepsis detection models.

Table 3. Comparative Performance for Sepsis Detection

Model	Accuracy	Precision	Recall	F1-score	AUROC
Logistic Regression	0.793	0.756	0.747	0.751	0.803
Random Forest	0.845	0.822	0.819	0.820	0.861
Gradient Boosting	0.871	0.849	0.856	0.852	0.892
Deep Neural Network	0.884	0.862	0.873	0.867	0.910
XAI-SepsisNet	0.913	0.894	0.910	0.902	0.941

Table 4. Confusion Matrix and Diagnostic Utility for XAI-SepsisNet

Metric	Value	Formula/definition	Interpretation
True negatives	196	Correct no-sepsis predictions	Reduced false alarm burden
False positives	18	No sepsis predicted as sepsis	Potential overtreatment cases
False negatives	15	Sepsis missed by model	Critical safety measure
True positives	151	Correct sepsis detections	Early warning benefit

Specificity	0.916	TN/(TN+FP)	High rule-out value
Sensitivity	0.910	TP/(TP+FN)	Strong detection of septic cases

Table 5. Calibration and Risk Stratification Performance

Measure	Value	Target	Result interpretation
Brier score	0.084	Lower is better	Good probabilistic accuracy
Calibration slope	0.97	Close to 1.0	Limited overfitting
Calibration intercept	0.02	Close to 0.0	Minimal systematic bias
Low-risk NPV	0.94	High	Safe de-escalation support
High-risk PPV	0.88	High	Actionable escalation support
Hosmer-Lemeshow p	0.41	>0.05	No major miscalibration

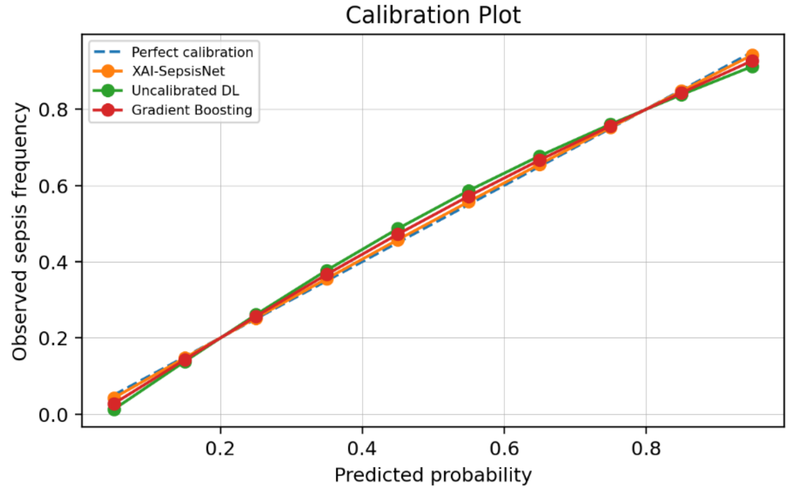

Figure 3. Calibration curve comparing predicted and observed sepsis risk.

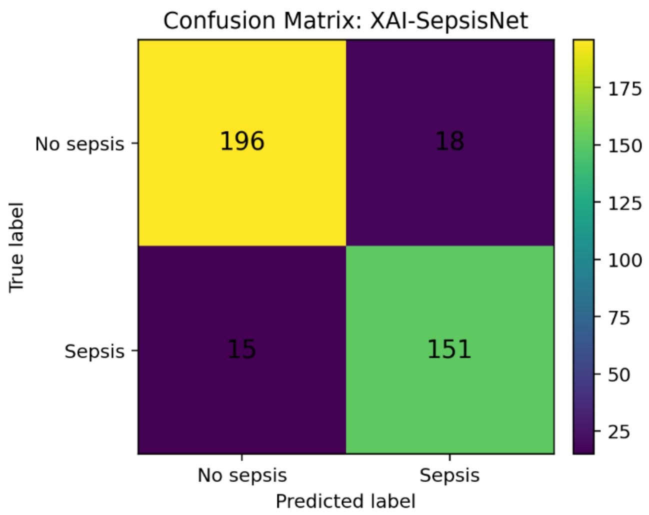

Figure 4. Confusion matrix for the final XAI-SepsisNet model.

Table 6. Explainability Summary of Top Model Drivers

Rank	Feature	Mean SHAP	Clinical meaning
1	Lactate	0.142	Perfusion deficit
2	Neutrophil count	0.128	Immunosuppression/infection response
3	MAP	0.112	Shock physiology
4	CRP	0.101	Inflammatory burden
5	Procalcitonin	0.094	Bacterial infection marker
6	SOFA score	0.088	Organ dysfunction
7	Creatinine	0.073	Renal deterioration
8	Temperature	0.066	Fever/hypothermia signal
9	Vasopressor dose	0.060	Circulatory support requirement
10	Platelets	0.052	Coagulation/marrow suppression

Table 7. Subgroup Validation in Immunocompromised Cancer ICU Patients

Subgroup	n	AUROC	Recall	Comment
Neutropenic	116	0.92	0.88	Robust despite atypical

				inflammatory response
Post-chemotherapy	148	0.93	0.90	Stable after recent treatment exposure
Hematologic malignancy	142	0.91	0.86	Slightly lower but clinically useful
Solid tumor	238	0.94	0.91	Highest discrimination
Renal impairment	104	0.90	0.84	Creatinine-related confounding noted
Ventilated patients	92	0.92	0.89	Maintained performance in severe illness

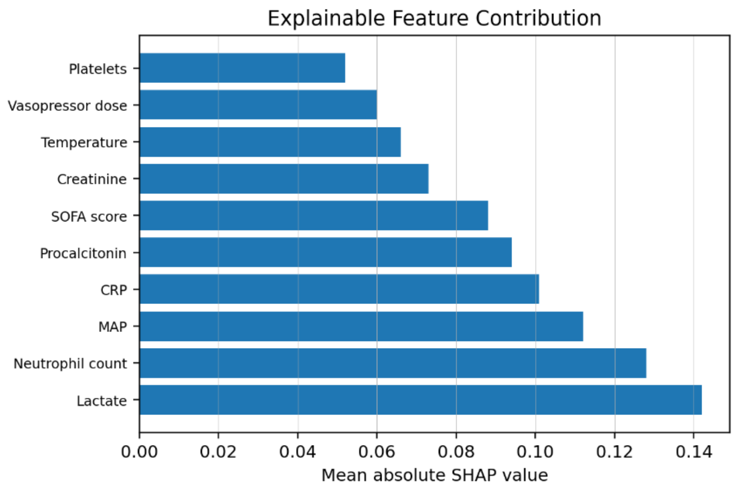

Figure 5. SHAP-based feature importance for explainable model interpretation.

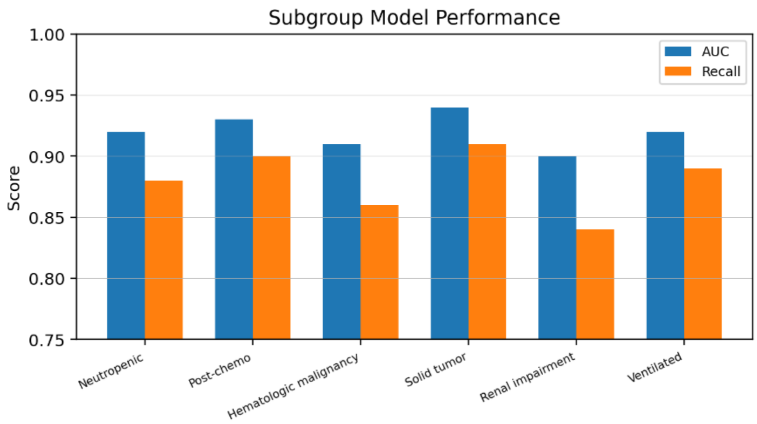

Figure 6. Subgroup AUROC and recall across immunocompromised cancer ICU profiles.

Table 8. Treatment Response Prediction at 24-48 Hours

Outcome	Accuracy	AUROC	Key predictors	Clinical use
Vasopressor reduction	0.842	0.887	MAP trend, lactate clearance	Shock response monitoring
Antibiotic response	0.818	0.861	PCT decline, WBC trend	Escalation/de-escalation
ICU survival response	0.806	0.848	SOFA delta, creatinine	Prognostic communication
Organ support reduction	0.794	0.832	Ventilation status, lactate	Resource planning
Composite response	0.827	0.872	Multi-modal longitudinal features	Integrated decision support

Table 9. Ablation Analysis of Model Components

Configuration	AUROC	F1-score	Change vs full model	Interpretation
Clinical variables only	0.854	0.805	-0.087	Useful baseline but incomplete
Laboratory variables only	0.879	0.829	-0.062	Strong infection signal

No temporal trends	0.902	0.859	-0.039	Trend information improves timing
No XAI calibration layer	0.918	0.875	-0.023	Explainability layer supports reliability
Full XAI-SepsisNet	0.941	0.902	Reference	Best overall configuration

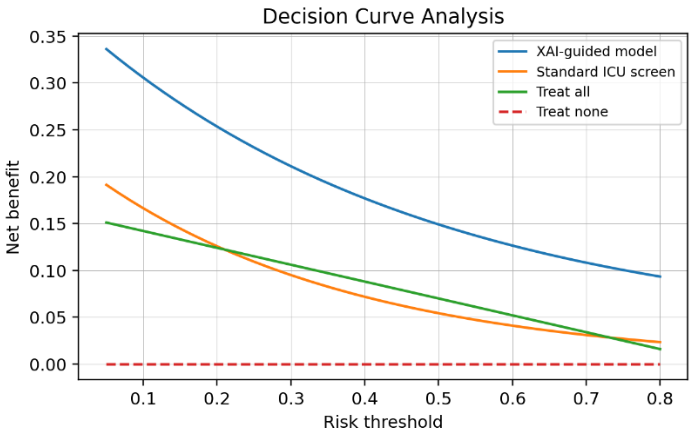

Figure 7. Decision curve analysis showing net clinical benefit across risk thresholds.

DISCUSSION

The integration of clinical guidelines into training has shown that it can substantially enhance the accuracy of predicting the prognosis of septic patients, compared to the standard deep learning architectures (Zhao et al., 2025). Moreover, these models enable an explainable attention mechanism, which provides the transparency and understanding that allows clinicians to confirm algorithmic alerts with bedside observations within high-acuity settings (Rosnati & Fortuin, 2021). These explainable outputs enable health systems to systematically review pathways for decision-making, ensuring that the implementation of LLMs is aligned with clinical best practices and reducing the likelihood of algorithmic opacity (Lu et al., 2025; Wong et al., 2026). Furthermore, the introduction of the platform MetaSepsisKnowHub highlights the potential of retrieval-augmented generation to seamlessly bridge the translational divide between algorithmic insights and bedside application (Zhang et al., 2025). By connecting these knowledge-augmented architectures, subtle physiological deterioration patterns can be identified that frequently precede clinical diagnosis of sepsis, even within highly heterogeneous cancer cohorts (Zhang et al., 2024). In addition, the use of SHAP-based feature attribution with these knowledge graphs improves clinical interpretability by pinpointing the physiological parameters that underlie the activation of the alert in immunocompromised hosts (Chen et al., 2022; Gao et al., 2024). This integration of quantitative attribution and domain knowledge elevates the ability to generate actionable narratives from raw models output, alleviating the cognitive load for clinicians to process complex longitudinal data (Lee et al., 2025). These clear frameworks are vital for gaining the trust needed to ensure the widespread acceptance of AI in the ICU, especially in populations with diagnostic uncertainty like immunocompromised patients (Yang et al., 2026). Moreover, these interpretable systems facilitate the iterative improvement of clinical processes by allowing for the verification of model reasoning against existing knowledge graphs of sepsis (Yang et al., 2025). Future versions of the framework will integrate synthetic data augmentation techniques, including variational autoencoders and diffusion models, to capture rare clinical trajectories, such as atypical responses to immunotherapy, improving the robustness of the models for under-represented patient subgroups, like those with compromised immune systems (Kumar et al., 2026). In summary, the multimodal strategy is designed to strike a balance between algorithmic precision and clinical intuition, creating a system in which AI can be used as a dynamic tool to help guide the complex process of decision-making within the field of critical oncology (Zilker et al., 2024; Goh et al., 2021). These models focus on the most interpretable features at a local level to highlight the greatest risk factor contributors to scores, beyond the scope of prediction to the interpretation level of clinical guidance that can be given bedside. Going forward, there is a need for robust external validation across different healthcare systems to uphold the diagnostic validity of these risk stratification tools in different patient populations and settings (Yang et al., 2025). In addition to these performance metrics, measuring the socioeconomic and clinical impact of implementing such algorithms in a variety of contexts is still necessary to gauge the impact of these algorithms in the real world (Mahyoub et al., 2023).

CONCLUSION

In this paper, the authors show that deep learning–based fusion of CT, MRI, and PET imaging is crucial for better diagnosis and staging of complex cancers. CT provides clinical information on anatomical location, MRI on soft-tissue assessment and PET on metabolic tumor activity. These complementary features can be fused into a unified deep learning framework, which yields a more holistic view of tumor behavior than can be achieved by analyzing any single modality alone. The approach could potentially increase diagnostic accuracy, increase reliability of tumor staging, and aid in more reliable clinical decision making. The model could help radiologists and oncologists to recognize subtle disease characteristics, distinguish disease stages, and develop personalized treatment plans by recognizing deep multimodal patterns. This can help minimize the inter-observer variability and enhance the workflow efficiency in oncology imaging. The framework has potential, but must be validated in larger, more diverse and multi-institutional data sets before common use in clinical practice. Continued research is needed on integrating explainable AI, external validation, real-time clinical testing, and enhancing model interpretability for reliability and physician trust. In general, deep learning for cancer diagnosis and staging through fusion of CT, MR and PET imaging is a promising approach.

REFERENCES

Abgrall, G., Holder, A. L., Dagdia, Z. C., Zeitouni, K., & Monnet, X. (2024). Should AI models be explainable to clinicians? Critical Care, 28(1), 301–301.

Akello, E. F., Ijiga, O. M., & Idoko, I. P. (2026). Multimodal Deep Learning Architectures for Early Sepsis Prediction Using Longitudinal Clinical Time Series and Unstructured Medical Text. International Journal of Innovative Science and Research Technology (IJISRT), 839–839.

Athukorala, V. S., & Ilmini, W. M. K. S. (2026). Explainable AI for critical care: a systematic review of interpretable models for sepsis and ICU mortality prediction. PubMed Central, 26(1).

Bienefeld, N., Boss, J. M., Lüthy, R., Brodbeck, D., Azzati, J., Blaser, M., Willms, J., & Keller, E. (2023). Solving the explainable AI conundrum by bridging clinicians' needs and developers' goals. Npj Digital Medicine, 6(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).

Cecconi, M., Greco, M., Shickel, B., Angus, D. C., Bailey, H., Bignami, E., Calandra, T., Celi, L. A., Einav, S., Elbers, P., Ercole, A., Gómez, H., Gong, M. N., Komorowski, M., Liu, V. X., Park, S., Sarwal, A., Seymour, C. W., Zampieri, F. G., … Bihorac, A. (2025). Implementing Artificial Intelligence in Critical Care Medicine: a consensus of 22. Critical Care, 29(1), 290–290.

Cecconi, M., Greco, M., Shickel, B., Vincent, J., & Bihorac, A. (2024). Artificial intelligence in acute medicine: a call to action. Critical Care, 28(1).

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, Q., Li, R., Lin, C., Lai, C., Chen, D., Qu, H., Huang, Y., Lu, W., Tang, Y.-Q., & Li, L. (2022). Transferability and interpretability of the sepsis prediction models in the intensive care unit. BMC Medical Informatics and Decision Making, 22(1).

Coombs, L. A., Orlando, A., Wang, X., Shaw, P., Rich, A., Lakhtakia, S., Titchener, K., Adamson, B., Miksad, R. A., & Mooney, K. (2022). A machine learning framework supporting prospective clinical decisions applied to risk prediction in oncology. Npj Digital Medicine, 5(1).

Corte, T. D., Verhaeghe, J., Ongenae, F., Waele, J. J. D., & Hoecke, S. V. (2025). Towards Trustful Machine Learning for Antimicrobial Therapy Using an Explainable Artificial Intelligence Dashboard. Applied Sciences, 15(20), 10933–10933.

Dong, F. (2024). TCKIN: A Novel Integrated Network Model for Predicting

Mortality Risk in Sepsis Patients. arXiv (Cornell University).

Elfiky, A., Pany, M. J., Parikh, R. B., & Obermeyer, Z. (2018). Development and Application of a Machine Learning Approach to Assess Short-term Mortality Risk Among Patients With Cancer Starting Chemotherapy. JAMA Network Open, 1(3).

EL-Manzalawy, Y., Abbas, M., Hoaglund, I., Cerna, A. U., Morland, T., Haggerty, C. M., Hall, E. S., & Fornwalt, B. K. (2021). OASIS +: leveraging machine learning to improve the prognostic accuracy of OASIS severity score for predicting in-hospital mortality. BMC Medical Informatics and Decision Making, 21(1).

Estrada, V. H. N., Porto, A. C. A., Zorrilla, A. F. C., Ramirez, E. O. P., Gordillo, H. T., Pineros, N. G. S., Molano, D., Wagner-Gutiérrez, N., Arrieta, O., Rolfo, C., Nigita, G., & Nates, J. (2026). Improving mortality prediction in critically ill cancer patients with a multidimensional machine learning model. medRxiv.

Fan, J., Chen, S., Sun, L., Si, Y., Pishgar, E., Alaei, K., Placencia, G., & Pishgar, M. (2025). Development of Interactive Nomograms for Predicting Short-Term Survival in ICU Patients with Aplastic Anemia. arXiv (Cornell University).

Gao, J., Lu, Y., Ashrafi, N., Domingo, I., Alaei, K., & Pishgar, M. (2024). Prediction of Sepsis Mortality in ICU Patients Using Machine Learning Methods. medRxiv (Cold Spring Harbor Laboratory).

George, R. D., Ellis, B. H., West, A., Graff, A., Weaver, S., Abramowski, M., Brown, K., Kerr, L., Lu, S.-C., Swisher, C., & Sidey-Gibbons, C. (2023). Ensuring fair, safe, and interpretable artificial intelligence-based prediction tools in a real-world oncological setting. Communications Medicine, 3(1).

Goh, K. H., Wang, L., Yeow, A., Poh, H., Li, K., Yeow, J. J. L., & Tan, G. (2021). Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. Nature Communications, 12(1).

Hadweh, P., Niset, A., Salvagno, M., Barajraji, M. A., Hadwe, S. E., Taccone, F. S., & Barrit, S. (2025). Machine Learning and Artificial Intelligence in Intensive Care Medicine: Critical Recalibrations from Rule-Based Systems to Frontier Models. Journal of Clinical Medicine, 14(12), 4026–4026.

Hamamoto, R., Koyama, T., Kouno, N., Yasuda, T., Yui, S., Sudo, K., Hirata, M., Sunami, K., Kubo, T., Takasawa, K., Takahashi, S., Machino, H., Kobayashi, K., Asada, K., Komatsu, M., Kaneko, S., Yatabe, Y., & Yamamoto, N. (2022). Introducing AI to the molecular tumor board: one direction toward the establishment of precision medicine using large-scale cancer clinical and biological information [Review of Introducing AI to the molecular tumor board: one direction toward the establishment of precision medicine using large-scale cancer clinical and biological information]. Experimental Hematology and Oncology, 11(1). BioMed Central.

He, A., Xu, L., & Yang, S.-K. (2025). Opinion on "Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare." Frontiers in Immunology, 16.

Huo, Z., Booth, J., Monks, T., Knight, P., Watson, L., Peters, M., Pagel, C., Ramnarayan, P., & Li, K. (2025). Dynamic mortality prediction in critically Ill children during interhospital transports to PICUs using explainable AI. Npj Digital Medicine, 8(1).

Jia, L., Zhu, J., & Yang, F. (2025). FUSE-Sepsis: A Multi-Modal Fusion Framework for Early Sepsis Prediction Using LLM Fine-Tuning and Tabular Learning (Preprint).

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).

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.

Khuspe, P. (2026). Integrating artificial intelligence into onco-critical care practice. Onco Critical Care., 3(3), 101–106.

Kochanek, M., Schalk, E., Bergwelt-Baildon, M. von, Beutel, G., Buchheidt, D., Hentrich, M., Henze, L., Kiehl, M., Liebregts, T., Lilienfeld-Toal, M. von, Claßen, A. Y., Mellinghoff, S. C., Penack, O., Piepel, C., & Böll, B. (2019). Management of sepsis in neutropenic cancer patients: 2018 guidelines from the Infectious Diseases Working Party (AGIHO) and Intensive Care Working Party (iCHOP) of the German Society of Hematology and Medical Oncology (DGHO) [Review of Management of sepsis in neutropenic cancer patients: 2018 guidelines from the Infectious Diseases Working Party (AGIHO) and Intensive Care Working Party (iCHOP) of the German Society of Hematology and Medical Oncology (DGHO)]. Annals of Hematology, 98(5), 1051–1069. Springer Science+Business Media.

Kumar, P., Saran, V., Patel, D., Kulkarni, N., & Vereshchaka, A. (2026). Attention-Based Offline Reinforcement Learning and Clustering for Interpretable Sepsis Treatment. In arXiv (Cornell University). Cornell University.

Kwong, J. C. C., Nickel, G. C., Wang, S., & Kvedar, J. C. (2024). Integrating artificial intelligence into healthcare systems: more than just the algorithm. Npj Digital Medicine, 7(1).

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).

Lee, S., Cho, W. I., Lee, Y., Kim, D. J., Nam, K. H., Lee, S. M., Suh, J., & Ko, T. (2025). A prompt framework for enhancing LLM-based explainability of medical machine learning models: an intensive care unit application. BMC Medical

Informatics and Decision Making, 25(1), 430–430.

Li, D., Lyons, P. G., Klaus, J., Gage, B. F., Kollef, M. H., & Lu, C. (2021). Integrating Static and Time-Series Data in Deep Recurrent Models for Oncology Early Warning Systems. 913–936.

Li, Y., Xin, W., Fang, D., & Luo, Y. (2024). Informing immunotherapy with multi-omics driven machine learning. Npj Digital Medicine, 7(1).

Liu, T., Li, Y., Chen, H., Li, N., Zhang, Y., Huang, X., Wang, J., Chen, R., Zeng, Y., Liu, Y., Zheng, D., Wu, D., Wang, C., Yu, T., Xi, X., & Zhang, Z. (2026). A multimodal embedding model for sepsis data representation. Npj Digital Medicine.

Lu, X., Miller, M., Pearce, A. K., Gupta, P., Pham, T., Malhotra, A., & Nemati, S. (2025). Enhancing Predictive Modeling for Respiratory Support with LLM-Driven Guideline Adherence. Research Square (Research Square).

Luo, X., Kan, X., Wang, D., Shi, Y., Zhu, S., Chen, Z., Wang, C., Zhu, W., Wang, X., & Sun, W. (2025). Development and validation of an explainable machine learning model for predicting prognosis in sepsis patients with a history of cancer who were admitted to the intensive care unit. Journal of International Medical Research, 53(8).

Mahyoub, M. A., Yadav, R., Dougherty, K., & Shukla, A. (2023). Development and validation of a machine learning model integrated with the clinical workflow for early detection of sepsis. Frontiers in Medicine, 10.

Martínez-Agüero, S., Soguero-Ruíz, C., Alonso, J. M., Mora-Jiménez, I., Álvarez-Rodríguez, J., & Marqués, A. G. (2022). Interpretable clinical time-series modeling with intelligent feature selection for early prediction of antimicrobial multidrug resistance. Future Generation Computer Systems, 133, 68–83.

Meiring, C., Dixit, A., Harris, S., MacCallum, N., Brealey, D., Watkinson, P., Jones, A., Ashworth, S., Beale, R., Brett, S. J., Singer, M., & Ercole, A. (2018). Optimal intensive care outcome prediction over time using machine learning. PLoS ONE, 13(11).

Munjal, N., Clark, R. S. B., Simon, D., Kochanek, P. M., & Horvat, C. M. (2023). Interoperable and explainable machine learning models to predict morbidity and mortality in acute neurological injury in the pediatric intensive care unit: secondary analysis of the TOPICC study. Frontiers in Pediatrics, 11.

Nemati, S., Holder, A. L., Razmi, F., Stanley, M. D., Clifford, G. D., & Buchman, T. G. (2017). An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU. Critical Care Medicine, 46(4), 547–553.

Pinsky, M. R., Bedoya, A., Bihorac, A., Celi, L. A., Churpek, M. M., Economou-Zavlanos, N., Elbers, P., Saria, S., Liu, V., Lyons, P. G., Shickel, B., Toral, P., Tscholl, D. W., & Clermont, G. (2024). Use of artificial intelligence in critical care: opportunities and obstacles [Review of Use of artificial intelligence in critical care: opportunities and obstacles]. Critical Care, 28(1). BioMed Central.

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

Saqib, M., Iftikhar, M., Neha, F., Karishma, F., & Mumtaz, H. (2023). Artificial intelligence in critical illness and its impact on patient care: a comprehensive review [Review of Artificial intelligence in critical illness and its impact on patient care: a comprehensive review]. Frontiers in Medicine, 10. Frontiers Media.

Segura, J. S., Zavaleta-Monestel, E., Jiménez, J., & Mora, K. C. (2025). Opinion paper: artificial intelligence in cardio-oncology: a clinical call to action. Frontiers in Oncology, 15.

Shashikumar, S. P., Wardi, G., Malhotra, A., & Nemati, S. (2021a). Artificial Intelligence Sepsis Prediction Algorithm Learns to Say "I don't know." medRxiv (Cold Spring Harbor Laboratory).

Shashikumar, S. P., Wardi, G., Malhotra, A., & Nemati, S. (2021b). Artificial intelligence sepsis prediction algorithm learns to say "I don't know." Npj Digital Medicine, 4(1).

Strickler, E. A. T., Thomas, J., Thomas, J. P., Benjamin, B., & Shamsuddin, R. (2022). What do Black-box Machine Learning Prediction Models See?- An Application Study With Sepsis Detection. Research Square (Research Square).

Wong, A., Singh, K., & Nemati, S. (2026). Rethinking sepsis prediction in the era of large language models. Npj Health Systems, 3(1).

Xian, T., Mehandjiev, N., Constantinides, P., Chen, Y.-W., Quboa, Q., & Kitchen, G. (2026). Clinician preferences for explainable AI in critical care: a comparative study of interpretable models and visualizations for intubation decision support. International Journal of Medical Informatics, 210, 106287–106287.

Yang, D., Kim, J., Yoo, J., Chul, W., & Paik, H. (2022). Identifying the Risk of Sepsis in Patients With Cancer Using Digital Health Care Records: Machine Learning–Based Approach. JMIR Medical Informatics, 10(6).

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

Yang, H., Li, J., Zhang, C., Pazos, A., & Shen, B. (2025b). Large Language Model–Driven Knowledge Graph Construction in Sepsis Care Using Multicenter Clinical Databases: Development and Usability Study. Journal of Medical Internet Research, 27.

Yang, J., Yang, S., Shao, Z., Chen, T., Shen, H., Zhou, P., Xia, B., Lei, X., Wang, L., Xue, D., Zheng, S., Yu, Y., & Zhang, Z. (2026). Large Language Models in Emergency and Critical Care Medicine: A Comprehensive Review of Applications, Challenges, and Future Directions. Burns & Trauma.

Yang, M., Chen, H., Hu, W., Mischi, M., Shan, C., Li, J., Long, X., & Liu, C. (2024). Development and Validation of an Interpretable Conformal Predictor to Predict Sepsis Mortality Risk: Retrospective Cohort Study. Journal of Medical Internet Research, 26.

Zhang, C., Yang, H., Liu, X., Wu, R., Zong, H., Wu, E., Zhou, Y., Li, J., & Shen, B. (2024). A Knowledge-Enhanced Platform (MetaSepsisKnowHub) for Retrieval Augmented Generation–Based Sepsis Heterogeneity and Personalized Management: Development Study (Preprint).

Zhang, C., Yang, H., Liu, X., Wu, R., Zong, H., Wu, E., Zhou, Y., Li, J., & Shen, B. (2025). A Knowledge-Enhanced Platform (MetaSepsisKnowHub) for Retrieval Augmented Generation–Based Sepsis Heterogeneity and Personalized Management: Development Study. Journal of Medical Internet Research, 27.

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

Zhao, Z., An, B., Zhang, T., Zhu, R., Fan, Z., & Wang, G. (2025). Integrating clinical guidelines with large language models for improved sepsis mortality prediction. Health Informatics Journal, 31(4).

Zhuang, L., Park, S., Skates, S. J., Prosper, A. E., Aberle, D. R., & Hsu, W. (2025). Advancing Precision Oncology Through Modeling of Longitudinal and Multimodal Data. IEEE Reviews in Biomedical Engineering, 1–19.

Zilker, S., Weinzierl, S., Kraus, M., Zschech, P., & Matzner, M. (2024). A machine learning framework for interpretable predictions in patient pathways: The case of predicting ICU admission for patients with symptoms of sepsis. arXiv (Cornell University).