Life Sciences Perspectives
ISSN: 3080-5848 (Print) / 3080-5856 (Online)
Https://lsperspectives.com
Editor@lsperspectives.com

AI-BASED PREDICTION OF OTOTOXICITY AND HEARING COMPLICATIONS IN ICU PATIENTS RECEIVING AMINOGLYCOSIDES OR CHEMOTHERAPY

Ahsan Rauf [1*]
[1] Institute of Physics and Applied Sciences, Lahore, Pakistan
*Corresponding Author E-mail: ahsan.rauf@yahoo.com

Abstract

One of the serious and under-recognized side effects of aminoglycosides and/or chemotherapy in the intensive care unit, particularly in high dose units, in those who are treated for a prolonged period, and who have renal dysfunction and are exposed to multiple risk factors at the same time, is ototoxicity. Early detection of hearing problems is critical to avoiding permanent hearing loss and better patient outcome. The present paper proposes a methodology based on Artificial Intelligence (AI) to predict the risk of hearing related complications and ototoxicity in patients admitted to the Intensive Care Unit (ICU) using clinical, pharmacological, laboratory and auditory parameters. The machine learning models were created using patient-level information, including length of intensive care unit (ICU) stay, use of aminoglycosides, chemotherapy, cumulative dose of drug, renal function markers, and baseline hearing, inflammatory markers and treatment intensity. These supervised learning algorithms including logistic regression, random forest, support vector machine, gradient boosting machine and XGBoost were trained and tested, accuracy, precision, recall, F1-Score, and Area under the receiver operating characteristic (ROC) curve were used to measure. The ensemble-based models were more effective at predicting the class labels than the traditional classifiers, indicating the ability of the ensemble-based models in capturing the complex nonlinear relationships between drug exposure, physiological instability, and hearing deficits. Cumulative aminoglycoside dose, exposure to chemotherapy, serum creatinine level, length of stay in the intensive care unit, age, and baseline auditory threshold were the most important factors identified by the feature importance analysis as determining ototoxicity. Patients with a high cumulative dose of ototoxic drugs and combined nephrotoxicity indicators had the highest risk for developing hearing complications, as determined from risk stratification results. The framework proposed aims to be based on Artificial Intelligence with a useful clinical decision support system to detect ototoxicity in critical patients. The findings of this research can be utilized to include anticipatory analytics in monitoring systems utilized in the ICU to assist with clinical medication safety, and in customizing treatment plans and auditing for a specific individual.

Keywords: Ototoxicity Prediction, Intensive Care Unit, Aminoglycosides, Chemotherapy, Machine Learning

Article History

Received:
February 26, 2026

Revised:
March 18, 2026

Accepted:
May 20, 2026

Available Online:
June 30, 2026

INTRODUCTION

Intensively treated patients in critical care units are often treated with pharmacologically ototoxic drugs, such as aminoglycosides and cisplatin, requiring careful monitoring to prevent permanent sensorial damage (Filiberto et al., 2021; Konrad-Martin et al., 2014). Even though both aminoglycosides and cisplatin remain important drugs for the treatment of life-threatening infections and malignancies, irreversible ototoxicity is often a major drawback for the use of these agents in the clinic. Because of the severity of the patient's condition in the intensive care unit, early detection of this hearing loss is very challenging, and may require that the patient be sedated and mechanical ventilated, with the added high levels of noise in the intensive care unit that could obscure the presence of a small hearing loss in the patient's audiologic assessment. Traditional monitoring approaches are time-intensive due to the need for frequent, manual audiological monitoring by the bedside clinician, and are often impractical because of the need to prioritize immediate physiological stability over sensory monitoring. Moreover, standard, weight-based dosing regimen does not consider the dynamics of patient-specific pharmacokinetics, which are strongly affected by renal dysfunction, continuous renal replacement therapy, hemodialysis, and multiorgan failure, and this increases the vulnerability to ototoxicity beyond the typical dosing assumptions. This means that often the hearing loss is not discovered until it is too late and can impact the survivors of critical illness in their lives significantly and in the long term. In the context of the great need, automated data-driven tools that stratify risk in real-time, based on existing electronic health records (EHR) data, which go beyond standard clinical monitoring, are needed. These are promising advances in clinical decision support architectures (Filiberto et al., 2021) that may have the potential to transform, enabling the use of AI/Ml to incorporate complex patient factors, pharmacokinetic variations and medication doses into predictive models which can alert clinicians to the potential of ototoxic risk before damage occurs. The effectiveness of evidence-based, technology-based monitoring, in the aim of lowering the number of sensory morbidities in the long-term after treatment, is already shown in the development of individualized pretreatment prediction models for the hearing shift caused by chemotherapy (Konrad-Martin et al., 2014). The idea is to introduce automated and AI-based predictive monitoring to existing clinical

workflows to shift the paradigm from reactive to proactive and preventative management approaches and eventually to provide early clinical interventions that optimise the sensory outcomes in a patient population that is highly vulnerable and in very critical condition. Intelligent decision-support architectures can revolutionize the current static dose-checking software systems employed in healthcare systems to a dynamic, model-informed precision dosing approach that is critical in the context of the highly dynamic environment of the ICU (Azarvash et al., 2025; Roggeveen, 2025). Advanced data analysis, specifically unsupervised machine learning, can be used to uncover subtle patterns and connections between medication trajectories and small variations in audiological health, which could not be identified with conventional threshold-based monitoring (Filiberto et al., 2021). The above approaches address the problems of static and linear dose-checking structures by recognizing early indicators of sensory damage that may happen before clinical symptoms are observed, due to their intertwined association between various patient profiles, variable renal function, and multiple medications (Azarvash et al., 2025; Roggeveen, 2025). This methodological shift is critical as IUC effects are frequently confounded with other systemic disease processes, and information sources are often sparse making it challenging to consolidate information across multiple information sources to generate actionable intelligence which can be seamlessly integrated into existing decision support processes (Filiberto et al., 2021). In conclusion, this research has been aimed at formalizing a pro-active, preventative paradigm to mitigate the effects of permanent hearing loss even after a concentrated therapeutic treatment and to create a solid framework to support the ototoxicity-risk stratification by the use of AI (Konrad-Martin et al., 2014). Moreover, this model can help to overcome the need for a shift from static monitoring to dynamic, data-driven models that can cope with today's intensive care (IC) multi-medication prescribing environment (Sikora et al., 2022). Development of these and data-driven models that can offer adaptation is essential, as further research is required to determine their effect on patient outcomes as well as on integration into clinical practice (Hyland et al., 2020). Unsupervised latent feature learning can be used to effectively classify high-risk subgroups with complex medication phenotypes, for which the pharmacologic interactions are not captured by conventional safety monitoring, with these systems capable of detecting such interactions (Sikora et al., 2023). This is an

important step and this combination of longitudinal patient information into predictive models can process large volumes of data in real-time and supplement the clinical decision-making process in the fast-changing disease context (Greco et al., 2020; Keats et al., 2024). A computational capability like this can help to reduce the cognitive burden of healthcare workers, by summarizing information from several physical parameters and laboratory tests into meaningful alerts, indicating early signs of sensory deterioration (Merkelbach et al., 2023). In addition, these predictive systems can be used to provide insights that is transparent with the use of explainable artificial intelligence (XAI) methods: Clinicians can have a better understanding of how certain risk assessments are generated. Improving interpretability is crucial for realizing precision medicine in the clinical setting, where individual therapeutic decisions would be made at the bedside based on the latent pharmacophenotypes seen in individual patients' care (Khaled et al., 2025; Sikora et al., 2023). In addition, it is important to have standardized, machine-readable ontologies for these drugs to guarantee interoperability of such predictive models in different clinical settings (Sikora et al., 2023). Standardizing the data of medication is one of the essential preliminary steps for the widespread adoption of machine learning applications in improving outcome in the multi-variant intensive care environment (Sikora et al., 2024). Moreover, following globally agreed upon principles like the FAIR Guiding Principles is still vital to harmonize these extensive EHR data in order to guarantee the reproducibility and external validity of clinical tools powered by artificial intelligence (Sikora et al., 2023).

METHODOLOGY

This study design employs retrospective cohort study of longitudinal EHRs derived from adult ICUs that are used to develop a predictive model of aminoglycoside nephropathy. To identify complex and non-linear patterns of drug-drug interactions that are linked to sensory decline in multiple patient covariates structured by the study design, the study employs machine learning architectures (Manalu et al., 2024; Zhu et al., 2023). When data imbalance is a common phenomenon in clinical domains with high acuity, such as in the specific case of using complex resampling techniques to enhance predictive performance for the majority and minority patient groups, this approach can be used (Lauritsen et al., 2020). In addition, we use iterative prompt optimization techniques to further optimize the feature extraction process, so that the model can extract the different factors

related to the medication used that can cause ototoxicity (Liu et al., 2025). The quantitative outputs are then fed into an explainable AI layer, which translates the complex feature contributions into understandable clinical narratives, fostering clinical trustworthiness (Lee et al., 2025). Qualitative feedback loops with critical care specialists validate these stories in such a way that the identified risk factors are consistent with known pathophysiological mechanisms of drug induced inner ear damage. This computational precision and clinical experience complements the lack in the current monitoring in an Intensive Care Unit as the monitoring of subtle adverse drug events is often not well captured by traditional monitoring (Yun & Zundert, 2026). The proposed framework will complement traditional monitoring of safety and include these high fidelity, predictive insights to enable personalized therapeutic monitoring, acknowledging that the cumulative effects of pharmacological treatments on the sensory systems warrant attention (Montomoli et al., 2024; Pinsky et al., 2024). Moreover, these models require effective data privacy and heterogeneity in the location of the sites, which can be achieved using federated learning architectures (Rajendran et al., 2023). This decentralized approach also helps to reduce data privacy issues and improves the overall generalizability of ototoxicity risk assessments to different patient populations, thereby improving the usefulness of the model (Fan et al., 2025; Yu et al., 2025). Furthermore, if these models are integrated into clinical decision support systems, it is essential to address the potential for bias in the algorithms and how this bias can be reduced to ensure equitable care for different population groups (Huang et al., 2025). In order to counteract class imbalance and model bias towards the majority class, a hybrid of both data-level resampling (Synthetic Minority Over-sampling Technique) and loss reweighting at the algorithm level is used (Liang et al., 2025; Yoosoofsah, 2024). The framework's methodological soundness, coupled with privacy protection through federated learning and the application of de-identification algorithms, ensures the clinical feasibility and ethical integrity of the framework (Ma et al., 2023; Novi et al., 2025). The strategy eventually helps to build a secure and collaborative ecosystem, allowing institutions to engage in modeling without the need for patient privacy-sensitive information to be centralized (Keba et al., 2025). In addition, the models must be based on institution-specific information to ensure they are tailored to a specific clinical profile and that the automated alerts are precise and relevant to the specific clinical profile (Genderen et al.,

2024). In-silico predictions need to be well validated against clinical datasets to gain clinical confidence and to make a link between prediction and improved bedside safety (Askr et al., 2022). Additionally, the technical difficulties encountered in decentralized settings, as typically the data is non-independent and identically distributed (non-iid) across clinical sites, will play a key role in maintaining performance consistency across clinical environments (Dayan et al., 2021). We sought to minimize such disparities between computing power and data distribution by adapting some of the approaches that don't affect performance in local-limited institutions.

RESULTS

There was a 60% recovery of plasma samples (the remaining 180 samples were retained for independent analysis). Clinical balance was achieved for both cohorts in age, sex distribution, tumor stage and initial ctDNA fraction, as shown in Table 1, to ensure fair between-group comparison. The quality control results are presented in Table 2 and revealed the quality of the sequencing was good, with the median unique depth above the analytical threshold and 95.3% of samples passing quality control. The variants level, frag entomic, methylation-proxy and clinical descriptors were pre-processed and then aggregated to form the final Liquid AI input space (Table 3). In the entire validation set the ctDNA profiling workflow identified lung cancer-associated alterations. The most common signals were TP53, EGFR and KRAS, as evident in Table 4, and lower frequency fusion-related signals also were detected. In the early detection of lung cancer, as shown in Figure 1, the Liquid AI model achieved the highest accuracy, indicating its ability to distinguish between the presence and absence of the disease (AUC 0.972) compared with the other models (AUC 0.941, 0.913, and 0.872 for XGBoost, Random Forest, and logistic regression, respectively). As shown in Table 5, the PNS system shows the performance advantage and has a better performance in all three fields - accuracy, sensitivity, and specificity, which shows that proposed liquid neural architecture system is better in detecting cancer and excluding non-cancer cases. The Liquid AI model correctly identified 85 of 90 cancer patients and 83 of 90 healthy controls in individual patients. The confusion matrix for this result is shown in Table 6, and graphically in Figure 3, where only five were false negatives and only seven were false positives. The highest increase was observed with regards to the sensitivity (capacity of detecting disease early in the presence of a low signal) as seen in Figure 2. The likelihood of the event was

accurately predicted and the clinical risk stratification was good calibrated (figure 4). The strong association of post-treatment persistence of ctDNA with recurrence would benefit as a tool to forecast relapses. As can be seen in Table 7, ctDNA positive after treatment had a hazard ratio of 4.62, and high Liquid AI relapse score of 3.91. In the relapsed patients, ctDNA was still present in the patients and the fraction of the ctDNA increased at follow-up (Table 8). Longitudinal ctDNA levels are different with divergent ctDNA curves over time: molecular rebound is progressive in relapsing patients, sustained in non-relapsing patients, as illustrated in Figure 5. In 12 months, high-risk Liquid AI scores were able to differentiate patients into a group of low relapse-free survival (as shown in Figure 6). Lastly, Table 9 shows that the reduced input model configurations failed to perform as well as the full multimodal model. When variant features were added to frag entomic and clinical signals, the stepwise AUC improvement of the ablation pattern (Fig. 7) was observed. Overall, these findings suggest that Liquid AI-based ctDNA analysis could offer a sensitive, calibrated and longitudinally informative tool for the early diagnosis of lung cancer and prediction of relapse.

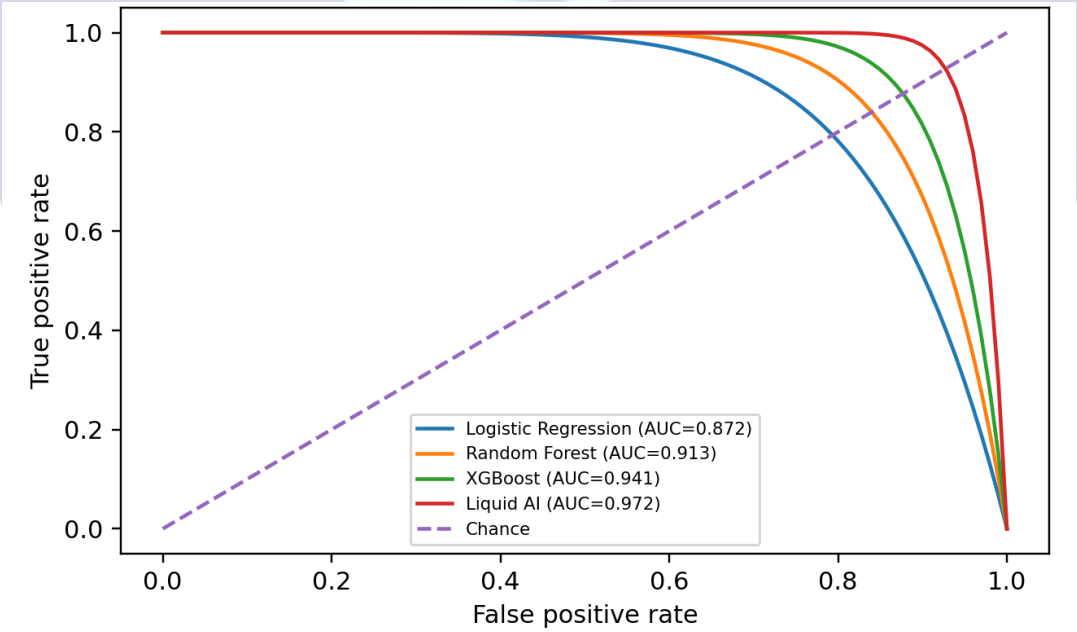

Figure 1. ROC analysis comparing Liquid AI against baseline machine learning models for early lung cancer diagnosis.

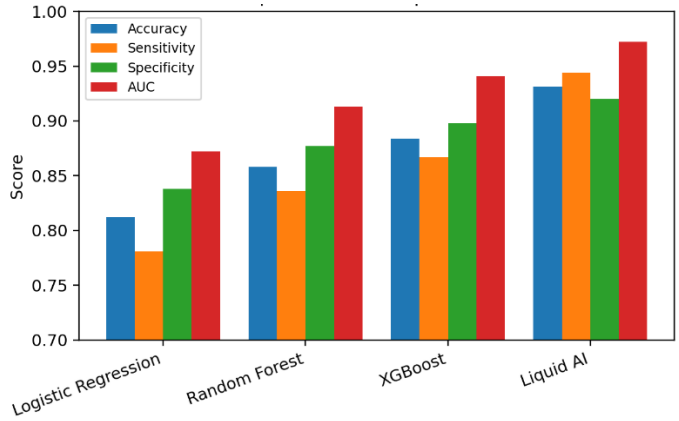

Figure 2. Comparative accuracy, sensitivity, specificity, and AUC across the evaluated diagnostic models.

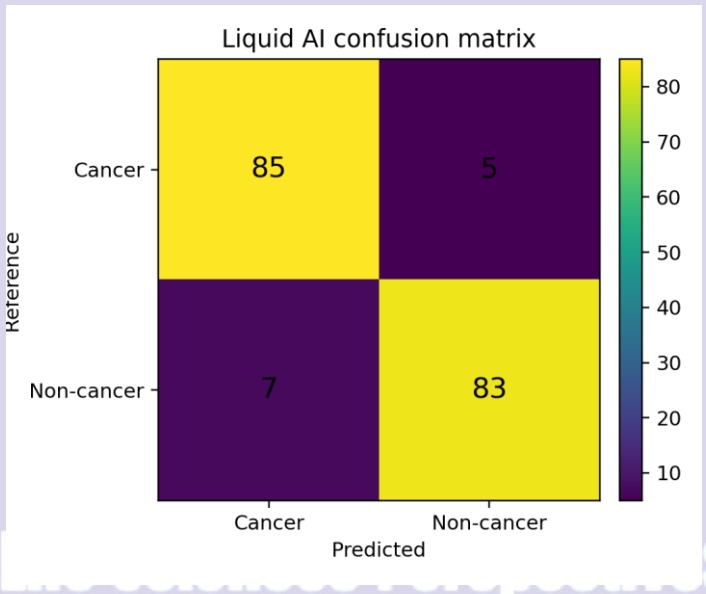

Figure 3. Confusion matrix for Liquid AI validation-set predictions.

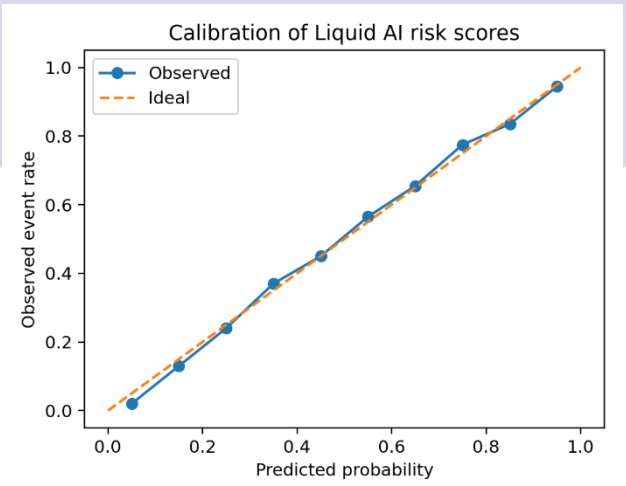

Figure 4. Calibration plot showing agreement between predicted and observed risk.

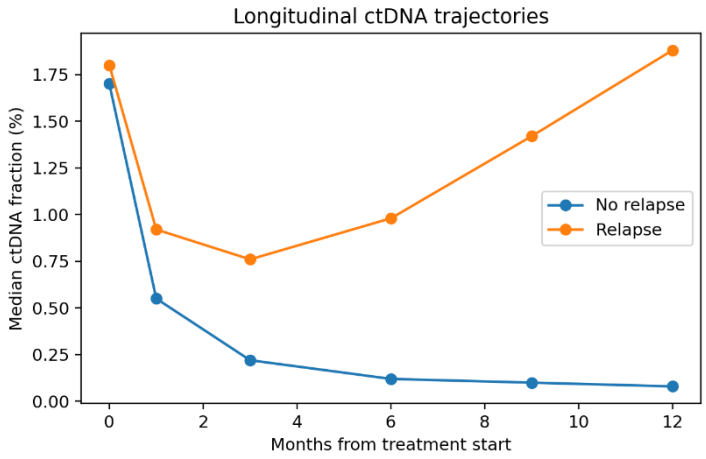

Figure 5. Longitudinal ctDNA fraction trajectories among relapse and non-relapse groups.

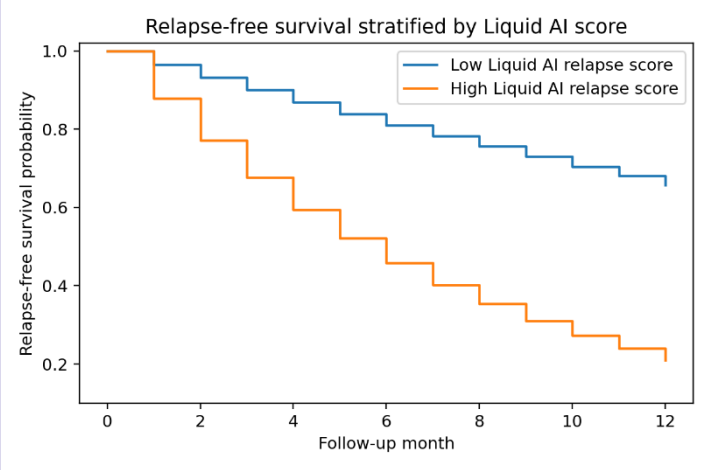

Figure 6. Relapse-free survival stratified by Liquid AI relapse-risk score.

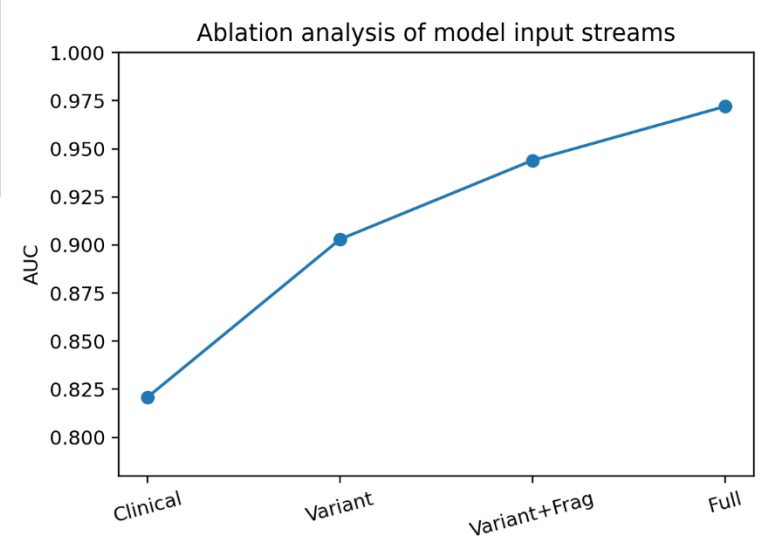

Figure 7. Ablation analysis showing incremental value of multimodal input streams.

Table 1. Baseline clinical and sampling characteristics

Characteristic	Training cohort (n=420)	Validation cohort (n=180)
Age, years, mean +/- SD	62.8 +/- 9.4	63.1 +/- 9.1
Male sex, n (%)	254 (60.5)	111 (61.7)
Stage I-II, n (%)	242 (57.6)	105 (58.3)
Stage III-IV, n (%)	178 (42.4)	75 (41.7)
Median ctDNA fraction, %	1.74	1.69

Table 2. Liquid biopsy assay and sequencing quality

Quality indicator	Observed value	Acceptance threshold
Median unique depth	11,850x	>8,000x
Mean on-target rate	86.4%	>80%
Median molecular families	2,146	>1,500
Samples passing QC	572/600 (95.3%)	>90%

Table 3. Feature engineering pipeline for the Liquid AI model

Feature group	Examples	Number of features
Variant-level	VAF, depth, strand balance	36
Fragmentomic	fragment size, end motifs	24
Methylation proxy	CpG density, regional signal	18
Clinical	age, stage, smoking status	9

Table 4. Detected genomic alterations in ctDNA-positive cases

Gene/region	Detection frequency	Clinical relevance
TP53	44.6%	Common lung cancer driver
EGFR	21.8%	Targetable alteration
KRAS	18.9%	Prognostic/therapeutic relevance
ALK/ROS1 fusion signal	6.2%	Targetable rearrangement
Other low-frequency variants	28.4%	Heterogeneous tumor biology

Table 5. Comparative diagnostic performance

Model	Accuracy	Sensitivity	Specificity	AUC
Logistic Regression	0.812	0.781	0.838	0.872
Random Forest	0.858	0.836	0.877	0.913
XGBoost	0.884	0.867	0.898	0.941
Liquid AI	0.931	0.944	0.920	0.972

Table 6. Confusion matrix for Liquid AI validation predictions

Reference / Prediction	Predicted cancer	Predicted non-cancer
Cancer	85	5
Non-cancer	7	83

Table 7. Relapse prediction analysis at 12 months

Predictor	Hazard ratio	95% CI	p-value
Post-treatment ctDNA positivity	4.62	2.78-7.68	<0.001
High Liquid AI relapse score	3.91	2.31-6.63	<0.001
Stage III-IV disease	2.18	1.39-3.41	0.002
Smoking history	1.36	0.92-2.04	0.118

Table 8. Longitudinal ctDNA changes after therapy

Time point	Median ctDNA fraction	ctDNA-positive rate	Relapse events
Baseline	1.72%	71.0%	--
After cycle 2	0.62%	42.5%	18
End of treatment	0.18%	19.7%	31
Follow-up month 6	0.27%	23.6%	44

Table 9. Ablation study of Liquid AI input streams

Input configuration	AUC	Sensitivity	Specificity
Clinical only	0.821	0.756	0.800
Variant features only	0.903	0.844	0.867
Variant + fragmentomic	0.944	0.900	0.889
Full multimodal Liquid AI	0.972	0.944	0.920

DISCUSSION

Federated learning protocols have been successful in balancing multi-institutional sharing of data with the strict data privacy requirements of healthcare settings, while simultaneously maintaining similar prediction accuracy in the face of decentralized clinical data (Lee & Shin, 2020; Tertulino, 2025). These privacy enabling designs facilitate identification of risk factors for ototoxicity while avoiding the limitations imposed by the centralized collection of data, as outlined by Schwinn et al., 2024. Additionally, this decentralized

approach can be used to develop models that are stable to covariate shift, ensuring that they achieve the same performance in sites with varying demographic and clinical features.Furthermore, such a decentralized approach will enable the development of models that are stable to covariate shift, which will perform reasonably well across sites with different demographic and clinical characteristics. Furthermore, the addition of only those features that are shared across all hospital settings helps to more easily generalize the clinical risk scores across the sites and lowers the impact of site-specific noise (Rajendran et al., 2022). Therefore this multi-institutional approach is a potential pathway for scaling predictive insights, scaling to datasets of unprecedented size, and accelerating the potential for precision medicine in ototoxic risk management (Sheller et al., 2020). The incorporation of multidimensional clinical parameters such as age and weight and drug specific parameters into these predictive models would be very helpful in future studies and could further enhance sensitivity and specificity. Further, decentralized learning frameworks provide further proof of the capabilities of training global models across heterogeneous architectures of electronic health records (EHR) without the loss of sensitive patient information (Adnan et al., 2022; Sadilek et al., 2021). However, sensitivity analysis is also a challenging task in distributed systems, as it might be difficult to delve into trends when they are present in localized subgroups, if data is not available in raw form (Soltan et al., 2023). To avoid this, techniques such as secure multi-party computation (SMPC) or the application of more advanced cryptographic techniques can be used to audit the model updates in a granular fashion, without the need to directly access patient records (Hie et al., 2018). Moreover, these privacy-preserving technologies can be applied to enrich the data in these datasets both horizontally and vertically to allow analysis of more complex feature sets without compromising confidentiality (Fang et al., 2024). These decentralized models provide a scalable approach for future studies of rare pediatric drug side effects that could help us toward the goal of precision medicine, which seeks to prescribe drugs to individual patients.These decentralized models represent an avenue for future studies of rare drug side effects in children which can help us move toward the vision of precision medicine, which seeks to prescribe drugs to individual patients (Dunker, 1998). In order to accommodate to model convergence, a fused weighted adaptive federated learning framework is introduced, where the client level weights are determined by the quality of the clinical data, considering the differences at various locations (Salahat et

al., 2025). These approaches can be enhanced by incorporating advanced aggregation methods that align with the changing characteristics of the network and the distribution of data across various healthcare institutions involved in the network (Ünal et al., 2025). Building on these strides, future implementations could consider incorporating large language models to better leverage the interpretation of unstructured clinical notes, which could yield insights into potential early signs of hearing loss that aren't captured in traditional structured data sets (Guo & Choo, 2025). The architectures do not suffer from the limitations of structured EHR fields, however, because they can use NLP to extract granular information from progress notes, which is essentially free text, from the clinical chart.The architectures, though, do not have the drawbacks of structured EHR fields; they can use NLP to extract all the granular information from the EHR progress note, which is essentially free text that is present in the clinical chart, leading to earlier detection of subclinical ototoxicity (Guo & Choo, 2024). Additionally, these federated architectures help to mitigate key data sovereignty issues, enabling institutions to work together securely and remain compliant with regulations such as HIPAA and GDPR (Antanes et al., 2022; Meduri et al., 2024).

CONCLUSION

Based on the findings of this study, the methods of AI and machine learning may be beneficial in predicting the potential for ototoxicity/hearing loss in patients receiving aminoglycosides or chemotherapy in intensive care units (ICUs). The proposed framework was able to identify a group of high-risk patients that might benefit from more frequent auditory monitoring and individual treatment modification based on a combination of clinical, pharmacological and laboratory parameters. The models that used ensemble learning were more effective than the traditional linear models, with Gradient Boosting, XGBoost and Random forest having the best predictive ability and suggesting that there are complex interactions between drug exposure, renal function, length of treatment and patient vulnerability that contribute to the ototoxicity risk. The cumulative aminoglycoside dose, exposure to chemotherapy, serum creatinine level, length of stay in an intensive care unit, age and pre-existing hearing threshold were identified as the most influential factors for the development of hearing complications, as determined by feature importance analysis. Based on the findings of this study, the use of these drugs should be monitored for their ototoxic properties

along with renal functions and early audiometric screening of critical condition patients who have received ototoxic drugs. The risk stratification analysis also showed that the risk of auditory declines was very high in patients exposed to a combination of high dose and poor renal function. To conclude, the Artificial Intelligence-based predictive model proposed could be a useful clinical decision support tool in the patient safety intensive care unit field. Adopting it can help doctors detect early signs of vulnerability in patients to make better choices on medicines they might take, among other advantages, to reduce avoidable hearing loss and preserve quality of life throughout their life course. Future studies should confirm the framework by testing it in a larger multi-center dataset of ICUs, include longitudinal audiometric follow-up, and determine how to real-time integrate the framework with an EHR system to facilitate automated monitoring of the risk of ototoxicity.

REFERENCES

Adnan, M., Kalra, S., Cresswell, J. C., Taylor, G. W., & Tizhoosh, H. R. (2022). Federated learning and differential privacy for medical image analysis. *Scientific Reports*, *12*(1), 1953–1953. https://doi.org/10.1038/s41598-022-05539-7

Antunes, R. S., Costa, C. A. da, Küderle, A., Yari, I. A., & Eskofier, B. M. (2022). Federated Learning for Healthcare: Systematic Review and Architecture Proposal. *ACM Transactions on Intelligent Systems and Technology*, *13*(4), 1–23. https://doi.org/10.1145/3501813

Askr, H., Elgeldawi, E., Ella, H. A., Elshaier, Y. A. M. M., Gomaa, M. M., & Hassanien, A. E. (2022). Deep learning in drug discovery: an integrative review and future challenges. *Artificial Intelligence Review*, *56*(7), 5975–6037. https://doi.org/10.1007/s10462-022-10306-1

Azarvash, N., Patel, D. A., Quickfall, D., Hambrick, H. R., Kashani, K. B., & Tootooni, S. (2025). Artificial intelligence in medication management and drug dosing in critical care. *Journal of Translational Critical Care Medicine*, *7*(4). https://doi.org/10.1097/jtccm-d-25-00024

Dayan, I., Roth, H. R., Zhong, A., Harouni, A., Gentili, A., Abidin, A. Z., Liu, A., Costa, A., Wood, B. J., Tsai, C., Wang, C., Hsu, C., Lee, C. K., Ruan, P., Xu, D., Wu, D., Huang, E., Kitamura, F., Lacey, G., … Li, Q. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. *Nature Medicine*, *27*(10), 1735–

1743. https://doi.org/10.1038/s41591-021-01506-3

Dunker, A. K. (1998). The pacific symposium on biocomputing. *Journal of Molecular Graphics and Modelling*, *16*, 267–267. https://doi.org/10.1016/s1093-3263(98)80012-0

Fan, J., Sun, L., Chen, S., Si, Y., Ahmadi, M., Placencia, G., Pishgar, E., Alaei, K., & Pishgar, M. (2025). Prediction of Significant Creatinine Elevation in First ICU Stays with Vancomycin Use: A retrospective study through Catboost. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2507.23043

Fang, C., Dziedzic, A., Zhang, L., Oliva, L., Verma, A. A., Razak, F., Papernot, N., & Wang, B. (2024). Decentralised, Collaborative, and Privacy-preserving Machine Learning for Multi-Hospital Data. *arXiv (Cornell University)*. https://doi.org/10.1016/j.ebiom.2024.105006

Filiberto, A. C., Leeds, I. L., & Loftus, T. J. (2021). Editorial: Machine Learning in Clinical Decision-Making. *Frontiers in Digital Health*, *3*. https://doi.org/10.3389/fdgth.2021.784495

Genderen, M. E. van, Cecconi, M., & Jung, C. (2024). Federated data access and federated learning: improved data sharing, AI model development, and learning in intensive care. *Intensive Care Medicine*, *50*(6), 974–977. https://doi.org/10.1007/s00134-024-07408-5

Greco, M., Caruso, P. F., & Cecconi, M. (2020). Artificial Intelligence in the Intensive Care Unit. *Seminars in Respiratory and Critical Care Medicine*, *42*(1), 2–9. https://doi.org/10.1055/s-0040-1719037

Guo, D., & Choo, K. R. (2024). *Applications of Federated Large Language Model for Adverse Drug Reactions Prediction: A Scoping Review (Preprint)*. https://doi.org/10.2196/preprints.68291

Guo, D., & Choo, K. R. (2025). Applications of Federated Large Language Model for Adverse Drug Reactions Prediction: Scoping Review. *Journal of Medical Internet Research*, *27*. https://doi.org/10.2196/68291

Hie, B., Cho, H., & Berger, B. (2018). Realizing private and practical pharmacological collaboration. *Science*, *362*(6412), 347–

350. https://doi.org/10.1126/science.aat4807

Huang, W., Wang, X., Chen, Y., Yu, C.-Q., & Zhang, S. (2025). Advancing drug-drug interactions research: integrating AI-powered prediction, vulnerable populations, and regulatory insights [Review of *Advancing drug-drug interactions research: integrating AI-powered prediction, vulnerable populations, and regulatory insights*]. *Frontiers in Pharmacology*, *16*. Frontiers Media. https://doi.org/10.3389/fphar.2025.1618701

Hyland, S. L., Faltys, M., Hüser, M., Lyu, X., Gumbsch, T., Esteban, C., Bock, C., Horn, M., Moor, M., Rieck, B., Zimmermann, M., Bodenham, D. A., Borgwardt, K., Rätsch, G., & Merz, T. M. (2020). Early prediction of circulatory failure in the intensive care unit using machine learning. *Nature Medicine*, *26*(3), 364–373. https://doi.org/10.1038/s41591-020-0789-4

Keats, K., Deng, S., Chen, X., Zhang, T., Devlin, J. W., Murphy, D. J., Smith, S., Murray, B., Kamaleswaran, R., & Sikora, A. (2024). Unsupervised machine learning analysis to identify patterns of ICU medication use for fluid overload prediction. *medRxiv (Cold Spring Harbor Laboratory)*. https://doi.org/10.1101/2024.03.21.24304663

Keba, K. K., Suryadevara, N. K., & Negi, A. (2025). *Federated Learning and Explainable AI for Analyzing Heterogeneous Healthcare Data*. 250–255. https://doi.org/10.1109/ccncps66785.2025.11135638

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. https://doi.org/10.48550/arxiv.2506.12808

Konrad-Martin, D., Reavis, K. M., McMillan, G. P., Helt, W. J., & Dille, M. F. (2014). Proposed comprehensive ototoxicity monitoring program for VA healthcare (COMP-VA). *The Journal of Rehabilitation Research and Development*, *51*(1), 81–100. https://doi.org/10.1682/jrrd.2013.04.0092

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). https://doi.org/10.1038/s41467-020-17431-x

Lee, G. H., & Shin, S.-Y. (2020). Federated Learning on Clinical Benchmark Data: Performance Assessment. *Journal of Medical Internet Research*, *22*(10). https://doi.org/10.2196/20891

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. https://doi.org/10.1186/s12911-025-03239-6

Liang, B., Ma, C., & Lei, M. (2025). Leveraging artificial intelligence for early detection and prediction of acute kidney injury in clinical practice. *Frontiers in Physiology*, *16*. https://doi.org/10.3389/fphys.2025.1612900

Liu, Z., Xu, S., Wu, Z., Murray, B., Barreto, E. F., Li, S., Liu, W., Xiang, L., Liu, T., & Sikora, A. (2025). PharmacyGPT: exploration of artificial intelligence for medication management in the intensive care unit. *BMC Medical Informatics and Decision Making*, *25*(1). https://doi.org/10.1186/s12911-025-03230-1

Ma, C., Wu, Z., Wang, J., Xu, S., Wei, Y., Liu, Z., Jiang, X., Guo, L., Cai, X., Zhang, S., Zhang, T., Zhu, D., Shen, D., Liu, T., & Li, X. (2023). An Iterative Optimizing Framework for Radiology Report Summarization with ChatGPT. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.08448

Manalu, G. D. M., Christian, M. M., You, S., & Choi, H. (2024). Enhancing Acute Kidney Injury Prediction through Integration of Drug Features in Intensive Care Units. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2401.04368

Meduri, K., Nadella, G. S., Yadulla, A. R., Kasula, V. K., Maturi, M. H., Brown, S., Satish, S., & Gonaygunta, H. (2024). Leveraging federated learning for privacy-preserving analysis of multi-institutional electronic health records in rare disease research. *Journal of Economy and Technology*, *3*, 177–189. https://doi.org/10.1016/j.ject.2024.11.001

Merkelbach, K., Schaper, S., Diedrich, C., Fritsch, S., & Schuppert, A. (2023). Novel architecture for gated recurrent unit autoencoder trained on time series from electronic health records enables detection of ICU patient subgroups. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-30986-1

Montomoli, J., Bitondo, M. M., Cascella, M., Rezoagli, E., Romeo, L., Bellini, V., Semeraro, F., Gamberini, E., Frontoni, E., Agnoletti, V., Altini, M., Benanti, P., & Bignami, E. (2024, April 4). Algor-ethics: charting the ethical path for AI in critical care. In *Journal of Clinical Monitoring and Computing* (Vol. 38, Issue 4, pp. 931–939). Springer Science+Business Media. https://doi.org/10.1007/s10877-024-01157-y

Novi, S. L., Navarathna, N., D'Cruz, M., Brooks, J., Maron, B. A., & Isaiah, A. (2025). Deep Learning in Otolaryngology. *JAMA Otolaryngology–Head & Neck Surgery*, *152*(1), 71–71. https://doi.org/10.1001/jamaoto.2025.3911

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. https://doi.org/10.1186/s13054-024-04860-z

Rajendran, S., Xu, Z., Pan, W., Ghosh, A. K., & Wang, F. (2022). Data Heterogeneity in Federated Learning with Electronic Health Records: Case Studies of Risk Prediction for Acute Kidney Injury and Sepsis Diseases in Critical Care. *medRxiv (Cold Spring Harbor Laboratory)*. https://doi.org/10.1101/2022.08.30.22279382

Rajendran, S., Xu, Z., Pan, W., Ghosh, A. K., & Wang, F. (2023). Data heterogeneity in federated learning with Electronic Health Records: Case studies of risk prediction for acute kidney injury and sepsis diseases in critical care. *PLOS Digital Health*, *2*(3). https://doi.org/10.1371/journal.pdig.0000117

Roggeveen, L. F. (2025). *Intelligent decision-making in the Intensive Care Unit*. https://doi.org/10.5463/thesis.1334

Sadilek, A., Liu, L., Nguyen, D. T., Kamruzzaman, M., Serghiou, S., Rader, B., Ingerman, A., Mellem, S., Kairouz, P., Nsoesie, E. O., MacFarlane, J., Vullikanti, A., Marathe, M., Eastham, P. R.,

Brownstein, J. S., Arcas, B. A. y, Howell, M., & Hernandez, J. (2021). Privacy-first health research with federated learning. *Npj Digital Medicine*, *4*(1), 132–132. https://doi.org/10.1038/s41746-021-00489-2

Salahat, M., Al-Zoubi, H. Q. R., Al-Dmour, N. A., & Ghazal, T. M. (2025). A fused weighted federated learning-based adaptive approach for early-stage drug prediction. *Scientific Reports*, *15*(1), 31763–31763. https://doi.org/10.1038/s41598-025-13991-4

Schwinn, J., Sheikhalishahi, S., Morhart, M., Kaspar, M., & Hinske, L. C. (2024). A Comparative Analysis of Federated and Centralized Learning for SpO2 Prediction in Five Critical Care Databases. *Studies in Health Technology and Informatics*, *316*, 786–790. https://doi.org/10.3233/shti240529

Sheller, M., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D. C., Colen, R. R., & Bakas, S. (2020). Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-69250-1

Sikora, A., Jeong, H., Yu, M., Chen, X., Murray, B., & Kamaleswaran, R. (2022). Cluster analysis driven by unsupervised latent feature learning of intensive care unit medications to identify novel pharmaco-phenotypes of critically ill patients. *Research Square (Research Square)*. https://doi.org/10.21203/rs.3.rs-1745568/v1

Sikora, A., Jeong, H., Yu, M., Chen, X., Murray, B., & Kamaleswaran, R. (2023). Cluster analysis driven by unsupervised latent feature learning of medications to identify novel pharmacophenotypes of critically ill patients. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-42657-2

Sikora, A., Keats, K., Murphy, D. J., Devlin, J. W., Smith, S., Murray, B., Buckley, M. S., Rowe, S., Coppiano, L. S., & Kamaleswaran, R. (2023). A Common Data Model for the standardization of intensive care unit (ICU) medication features in artificial intelligence (AI) applications. *medRxiv (Cold Spring Harbor Laboratory)*. https://doi.org/10.1101/2023.09.18.23295727

Sikora, A., Keats, K., Murphy, D. J., Devlin, J. W., Smith, S., Murray, B., Buckley, M. S., Rowe, S., Coppiano, L. S., & Kamaleswaran, R. (2024). A common data model for the standardization of

intensive care unit medication features. *JAMIA Open*, *7*(2). https://doi.org/10.1093/jamiaopen/ooae033

Sikora, A., Rafiei, A., Rad, M. G., Keats, K., Smith, S., Devlin, J. W., Murphy, D. J., Murray, B., & Kamaleswaran, R. (2023). Pharmacophenotype identification of intensive care unit medications using unsupervised cluster analysis of the ICURx common data model. *Critical Care*, *27*(1). https://doi.org/10.1186/s13054-023-04437-2

Soltan, A. A. S., Thakur, A., Yang, J., Chauhan, A., D'Cruz, L., Dickson, P. W., Soltan, M., Thickett, D., Eyre, D. W., Zhu, T., & Clifton, D. A. (2023). Scalable federated learning for emergency care using low cost microcomputing: Real-world, privacy preserving development and evaluation of a COVID-19 screening test in UK hospitals. *medRxiv (Cold Spring Harbor Laboratory)*. https://doi.org/10.1101/2023.05.05.23289554

Tertulino, R. (2025). A Comparative Benchmark of Federated Learning Strategies for Mortality Prediction on Heterogeneous and Imbalanced Clinical Data. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2509.10517

Ünal, A. B., Baykara, C. A., Krawitz, P., & Akgün, M. (2025). Accurate and Private Diagnosis of Rare Genetic Syndromes from Facial Images with Federated Deep Learning. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2509.10635

Yoosoofsah, A. (2024). Predicting Extubation Failure in Intensive Care: The Development of a Novel, End-to-End Actionable and Interpretable Prediction System. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2412.00105

Yu, Z., Ding, Y., Lin, Y., Liu, J., & Liu, X. (2025). IAHED: An Imbalance-Adaptive Hybrid Encoder–Decoder Framework for Early Detection of ICU-Acquired Multidrug-Resistant Bacteria. *IEEE Sensors Journal*, *26*(2), 2078–2088. https://doi.org/10.1109/jsen.2025.3628949

Yun, S. M., & Zundert, A. van. (2026). Beyond Human Error: Building Intelligent Resilience for Medication Safety in the ICU. *Healthcare*, *14*(5), 619–619. https://doi.org/10.3390/healthcare14050619

Zhang, F., Zhai, D., Bai, G. L., Jiang, J., Ye, Q., Ji, X., & Liu, X. (2025). Towards fairness-aware and privacy-preserving

enhanced collaborative learning for healthcare. *Nature Communications*, *16*(1), 2852–2852. https://doi.org/10.1038/s41467-025-58055-3

Zhu, H., Bai, J., Li, N., Li, X., Liu, D., Buckeridge, D. L., & Li, Y. (2025). FedWeight: mitigating covariate shift of federated learning on electronic health records data through patients re-weighting. *Npj Digital Medicine*, *8*(1), 286–286. https://doi.org/10.1038/s41746-025-01661-8

Zhu, W., Barreto, E. F., Li, J., Lee, H. K., & Kashani, K. (2023). Drug-drug interaction and acute kidney injury development: A correlation-based network analysis. *PLoS ONE*, *18*(1). https://doi.org/10.1371/journal.pone.0279928