




# Application of explainable artificial intelligence integrating with electronic health record in oncology

Yuhan Yang<sup>1†\*</sup> , Xici Liu<sup>2†</sup>

<sup>1</sup>West China Hospital, Sichuan University, Chengdu 610000, Sichuan, China

<sup>2</sup>Sichuan Technical Inspection Center for Medical Products (Sichuan Technical Inspection Center for Vaccine), Chengdu 610000, Sichuan, China

<sup>†</sup>These authors contributed equally to this work.

\***Correspondence:** Yuhan Yang, West China Hospital, Sichuan University, Chengdu 610000, Sichuan, China. [yyh\\_1023@163.com](mailto:yyh_1023@163.com)

**Academic Editor:** Gennaro Daniele, Agostino Gemelli University Policlinic, Italy

**Received:** November 3, 2025 **Accepted:** January 11, 2026 **Published:** February 4, 2026

**Cite this article:** Yang Y, Liu X. Application of explainable artificial intelligence integrating with electronic health record in oncology. *Explor Target Antitumor Ther.* 2026;7:1002357. <https://doi.org/10.37349/etat.2026.1002357>

## Abstract

Machine learning (ML) and deep learning (DL) models applied to electronic health records (EHRs) have substantial potential to improve oncology care across diagnosis, prognosis, treatment selection, and trial recruitment. However, opacity of many high-performing models limits clinician trust, regulatory acceptance, and safe deployment. Explainable artificial intelligence (XAI) methods aim to make model behavior understandable and actionable in clinical contexts. The present perspective summarizes current XAI approaches applied to EHR-based oncology tasks, identifies key challenges in evaluation, reproducibility, clinical utility, and equity, and proposes pragmatic recommendations and research directions to accelerate safe adoption in oncology. Common XAI categories used with EHR data include feature importance/interaction methods, intrinsically interpretable models, attention mechanisms, dimensionality reduction, and knowledge distillation or rule extraction. Tree-based models with SHapley Additive exPlanations (SHAP) explanations dominate recent EHR studies. Other interpretable strategies, such as generalized additive models and rule sets, appear in settings where transparency is prioritized. Gaps include inconsistent reporting, scarce formal evaluation of explanations for clinical utility, limited reproducibility for data and code availability, inadequate external validation, and insufficient consideration of fairness and equity that these issues are particularly important in oncology, where heterogeneity and stakes are high. Overall, integrating XAI with EHR-driven oncology models is promising but underdeveloped, which requires further progress by multi-stakeholder evaluation frameworks, reproducible pipelines, prospective and multicenter validations, and equity-aware design. The field should prioritize clinically meaningful explanations beyond ranking features and study how explanations affect clinician decision-making and patient outcomes.



## Keywords

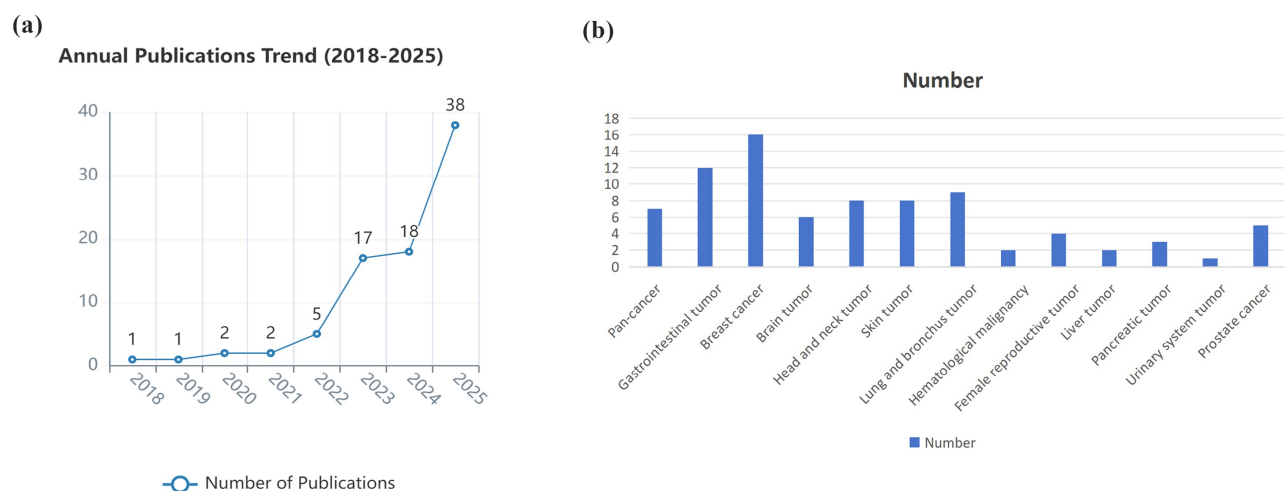
explainable artificial intelligence, electronic health records, oncology, interpretability, model evaluation framework, SHAP, clinical decision support

## Introduction

Electronic health records (EHRs) provide rich, longitudinal, multimodal clinical data that can power machine learning (ML)/deep learning (DL) models, tasks including early detection, risk stratification, prognosis prediction, and treatment selection for oncology patients [1, 2], and optimizing clinical trial registration and enrollment for medical professionals [3]. High-performing “black box” models, such as ensembles and neural networks, frequently outperform traditional approaches on complex prediction tasks, yet their opacity impedes clinician trust, regulatory approval, and safe deployment in high-stakes oncology care [4–6]. Explainable artificial intelligence (XAI) techniques seek to open the black box by producing human-interpretable descriptions of model behavior and rationale [7–9]. Recent systematic and scoping reviews of XAI applied to EHR data emphasize a surge in XAI use but highlight notable heterogeneity and a lack of rigorous evaluation and reproducibility [10–12]. This perspective aims to synthesize those insights and apply them specifically to oncology, as well as outlining current practices, limitations, and actionable recommendations.

## XAI categorization relevant to EHR-based oncology

Building on prior taxonomies, XAI approaches used with EHR data can be organized as follows. Each has different implications for clinical adoption in oncology. The publication trend related to XAI approaches in the field of oncology using EHR data in recent years is shown in Figure 1.



**Figure 1. Publication summary of XAI approaches used with EHR data in oncology.** (a) Publication trend of XAI studies in the field of oncology using EHR data between January 1st, 2018, and Nov 1st, 2025. (b) Oncology categorization of XAI studies using EHR data in recent publications. Panel (a) shows the number of publications per year meeting the inclusion; panel (b) categorizes included studies by oncology subdomain and XAI method. We searched PubMed, Web of Science, and arXiv for articles published Jan 1, 2018, to Nov 1, 2025, using terms (“explainable AI” OR “XAI”) AND (“electronic health records” OR “EHR”) AND (“oncology” OR “cancer”). Included studies applied XAI methods to EHR-derived oncology tasks. Exclusions: non-EHR imaging-only studies, purely technical XAI methods without EHR use.

## Feature importance and interaction methods

The post-hoc, model-agnostic, or model-specific methods have been found to quantify the contribution of input features to model predictions, for instance, SHapley Additive exPlanations (SHAP), permutation importance, and partial dependence plots (PDPs) [13, 14]. Recent studies have highlighted SHAP’s widespread use in survival or stage prediction using ensemble/tree-based models on cancer cohorts, such as The Cancer Genome Atlas (TCGA) [15] that the application of SHAP could produce global feature

rankings in which genes, mutations, or clinical covariates are most predictive of survival and local SHAP values to explain an individual patient's risk. These kinds of XAIs have the potential to identify prognostic features like lab values, genomic markers, and comorbidities, explain individual risk predictions, and surface potential feature interactions, such as drug-comorbidity interactions.

### **Intrinsically interpretable models**

The models designed for transparency consist of generalized additive models (GAMs), decision trees, rule lists, and sparse linear models, which provide inherently understandable decision logic but may trade off predictive performance. The approaches like decision trees, logistic regression variants, and cluster-based methods have been used for cancer diagnosis and risk prediction with interpretable survival scores from nomograms construction [16, 17]. The boosted/ensemble decision tree adaptations and microarray datasets have also been applied to produce interpretable models for cancer staging or subtype classification [18, 19]. It notes that the model development could be improved from an interpretable design, such as shallow decision trees, logistic regression with limited features, and Bayesian rule lists, rather than explaining a black box post-hoc. The optimized interpretability of XAI on EHRs would be used in treatment-decision aids with a clear rationale to generate triage tools guiding therapy escalation, as well as regulatory scenarios with required traceability.

### **Attention mechanisms and representation-focused explanations**

The sequence/time-series models have usually been used for longitudinal EHR data to highlight temporally important events and attention weights are sometimes interpreted as importance scores, although their reliability as explanations is debated. For clinical sequences, Reverse Time Attention model (RETAIN)-style reverse attention designed for interpretability and self-attention variants have been adapted for clinical sequences, that RETAIN/RetainVis-inspired models were applied to EHRs [20], which might have transferable potential to oncology longitudinal HER. Attention highlights prior imaging diagnoses, progressive lab trends, or particular chemotherapy cycles driving a prediction of near-term deterioration or recurrence [15, 21, 22], which implies potential application on oncology treatment timelines to sequential HER by addressing critical time windows or events of rapid biomarker decline, driving a prognosis prediction.

### **Dimensionality reduction and visualization**

The methods, including t-SNE, UMAP, and concept bottlenecks, project high-dimensional features into lower-dimensional interpretable representations, which are useful for cancer cohort stratification and exploratory analyses, especially sparse DL, Least Absolute Shrinkage and Selection Operator (LASSO), clustering to identify exemplar patients, selecting gene pathways or features most informative for cancer subtype or survival [23–25]. The applied sparse DL and pathway-level selection have been applied in glioblastoma survival prediction using TCGA-like data, while dimensionality reduction has also found for gene and pathway discovery in cancer datasets [26]. The clustering or affiliation analysis on EHRs shows abilities in stratifying patients with stage 1 lung cancer for further definition of clinically meaningful subgroups [27].

### **Knowledge distillation and rule extraction**

There have been techniques for distilling complex model behavior into simpler surrogate models or rule sets to approximate decision boundaries, including surrogate decision trees, decision-sets, rule extraction, and mimic learning. The mimic learning and distillation for clinical re-admission and mortality tasks have been discussed with applicable approaches to oncology models by distilling a neural survival model into a set of rules that estimate recurrence risk based on a small set of features [28]. The decision-sets design using metrics like rule count and rule length can be used to produce compact rule lists in order to classify tumor subtype or indicate eligibility for a clinical pathway [29].

## SHAP: benefits, known limitations, and recommended mitigations for oncology EHRs

SHAP has been widely used for feature-level explanations because of its theoretically grounded attribution and local explanations. However, several practical issues are critical in oncology applications:

- **Instability and sampling variability:** For complex, high-dimensional EHRs or rare endpoints (common in oncology), SHAP values can vary substantially with bootstrap resampling or small data shifts. Recommendation: report confidence intervals/variance for local SHAP attributions obtained via bootstrapping and quantify explanation reproducibility across folds.
- **Feature correlation and attribution ambiguity:** Correlated clinical or genomic features can produce misleading attributions since Shapley values distribute shared contribution among correlated features arbitrarily [30]. Recommendation: pre-process by grouping highly correlated features (e.g., pathway-level aggregation for genomics, summary lab trends), use conditional SHAP approximations when feasible, and complement with interaction analyses (SHAP interaction values) and causal or counterfactual checks.
- **Computational burden:** Exact SHAP for complex models and large feature sets is expensive. Recommendation: use approximations (TreeSHAP for tree ensembles), feature grouping, or mimic/surrogate models for explanations at scale; report compute resources and latency relevant for EHR integration.
- **Interpretability in temporality/multimodality:** Standard SHAP applied to tabular snapshots may miss sequential or imaging features' contributions. Recommendation: apply SHAP to modality-specific components (e.g., per-visit features, embedding dimensions for images) and combine with temporal attention visualizations and counterfactual narratives.

## Applications and further considerations of XAI integrating EHR

For the diagnosis of oncology patients, the EHR-based predictors for early cancer detection benefit from explanations in flagging at-risk patients from routine labs and symptoms by justifying screening or referral recommendations and avoiding alarm fatigue [31, 32]. Explaining survival or recurrence risk allows oncologists to contextualize model outputs with patient-specific drivers on tumor markers and performance status, which enables shared decision-making for prognosis and survival prediction [33]. For patients who need precise treatment, XAI reveals clues about which clinical, genomic, or comorbidity features could drive predictions favoring a given therapy, potentially supporting treatment personalization by offering interpretability when model recommendations diverge from standard of care [34]. To optimize clinical trial matching and operational workflows, transparent models can justify eligibility flags and improve trial recruitment by allowing trial teams to understand why patients are prioritized [35, 36]. Otherwise, explanations could strengthen treatment selection by helping clinicians assess whether a predicted adverse event is plausible and what modifiable factors contributed to safety monitoring [37].

It has been found a lack of rigorous evaluation of XAI methods for clinical relevance that typical assessments often stop at qualitative plausibility or visualization [38]. For further oncology deployment, it requires comprehensive evaluations by addressing fidelity about how well explanations reflect true model computations, clinical usefulness of whether explanations change clinician decisions appropriately, robustness in stability of explanations across input perturbations and cohorts, and human-centered evaluation concerning usability studies with oncologists and multidisciplinary teams. Otherwise, reproducibility remains limited that many studies have used proprietary EHRs without sharing code, independent verification, or external validation [33, 39, 40]. The XAI models in area of oncology must prioritize data-sharing strategies, code release, and detailed reporting of preprocessing and model pipelines. Moreover, clinician trust depends not only on explanation correctness but also on relevance, presentation, and fit within clinical workflows. Therefore, explainability should be tailored that simple

feature lists might suffice for some tasks with binary classes while causal or counterfactual explanations might be more persuasive for treatment decisions. To identify legal and regulatory factors, regulatory frameworks increase the imperative for meaningful explanations. For oncology patients with life-altering interventions, regulators and institutions will expect rigorous validation and traceability.

## Challenges and future directions of XAI in oncology

There have been several challenges influencing future application of XAI on EHRs in oncology, including heterogeneity, multimodal data, class imbalance, and equity and representation. Heterogeneity is attributed to tumor biology diversity, treatment modalities, and patient comorbidities, which increase model complexity, as well as the challenge of producing generalizable explanations. For multimodal data, oncology relies on imaging, pathology, genomics, and EHR-derived clinical features that integrating and explaining multimodal models remains technically and conceptually difficult. Otherwise, there exists class imbalance. Many oncology endpoints are rare, which make model explanations sensitive to sampling variability. Concerning healthcare equity, under-representation of low-resource settings and minority populations risks biased models and explanations without generalization, which would be raised as a major equity gap in AI-in-healthcare reviews and domain-specific syntheses.

There still be several oncology-specific challenges and implications:

- Heterogeneity: inter-tumor and intra-patient heterogeneity mean features predictive in one cohort may not generalize. Action: require external validation and subgroup explanation audits; present explanation differences by tumor subtype/stage.
- Sparsity and rare events: many oncology endpoints (e.g., rare toxicities) lead to class imbalance, making explanation reliability low. Action: bootstrap estimation of explanation variance; use data augmentation or synthetic controls to stabilize estimates.
- Multimodality: EHRs combined with imaging, pathology, and genomics complicate explanation attributions. Action: use modality-specific explainers, hierarchical explanations (first indicate which modality, then which features), and design visualization for tumor boards.
- Temporal complexity: treatments and biomarker trajectories matter. Action: adopt sequence-aware explainers (RETAIN-like models with clear attention maps, counterfactual time-window analyses) and present temporal narratives.
- Equity/regulatory stakes: life-altering decisions require rigorous fairness auditing and traceability. Action: add fairness metrics and transparent documentation of datasets and preprocessing.

The further application of XAI on EHRs in oncology needs to prioritizing clinically meaningful explanation types by moving beyond feature importance ranks to context-rich explanations, for example, counterfactuals about what would change the prediction, causal insights where feasible, and patient-level narratives connecting model drivers to actionable clinical considerations. The standardized evaluation framework is also warranted by adopting batteries to assess fidelity, robustness, clinical utility, and user acceptance and incorporating prospective human-in-the-loop studies to measure decision changes and patient-level outcomes. The enhancement of reproducibility and reporting is also required by sharing code, model checkpoints, synthetic or de-identified datasets when possible as well as adhering to reporting checklists including data preprocessing, feature construction, model selection, and XAI settings. Otherwise, the requirement of external validation across diverse oncology centers with prospective pilot deployments need to be implemented to assess transportability and real-world impact. With equity-centered design and fairness auditing, systematical evaluation of subgroup performance and explanation differences consists of fairness metrics and qualitative assessments with clinicians caring for under-represented populations. Furthermore, it's necessary to develop interactive explanation interfaces integrated into oncology workflows, such as EHR dashboards and tumor boards co-designed with clinicians to present explanations at the appropriate level of detail.



## Multi-dimensional evaluation and integration framework for XAI in oncology EHRs

A practical framework has been proposed to evaluate and integrate XAI into oncology workflows. The framework explicitly evaluates explanations across five dimensions and prescribes integration steps:

- Fidelity: quantitative checks that explanations reflect model computation (e.g., agreement between surrogate and original model, perturbation tests).
- Robustness and stability: sensitivity analyses across input perturbations, bootstrapping to measure explanation variance, and tests across cohorts [41].
- Clinical utility: prospective human-in-the-loop studies assessing whether explanations change clinician decisions in intended directions; pre/post comparison of decision concordance and downstream patient-relevant outcomes.
- Fairness and transportability: subgroup performance and explanation disparity audits across demographic, socioeconomic, and clinical strata; external validation across centers.
- Usability and workflow fit: qualitative usability testing with oncologists, nurses, and tumor board members, integrated User experience (UX) for EHR dashboards, and contextualized explanation level (summary vs. detail).

### Integration steps

- Pre-deployment: internal fidelity and stability tests + fairness audits on retrospective data.
- Pilot deployment: prospective small-scale human-in-the-loop pilot within tumor boards; measure decision changes and feasibility.
- Multicenter validation: external validation at heterogeneous centers, with standardized reporting and code release.
- Deployment & monitoring: continuous post-deployment monitoring of model and explanation drift and periodic re-audits.

### Use-case: explainable multimodal model for recurrence risk in stage II–III colorectal cancer

- Dataset: de-identified EHR (demographics, labs, medications, comorbidities), pathology reports (NLP-extracted features), targeted genomic panel (mutations), and structured treatment records (retrospective).
- Task: 2-year recurrence risk prediction after definitive surgery (binary), with class imbalance.
- Model architecture: multimodal fusion model using gradient boosted trees (LightGBM) on tabular EHR plus logistic regression on aggregated genomic pathway scores; output fused via stacking ensemble to predict recurrence probability. Temporal features encoded as summary trends (slope, variance) over pre-surgery period.
- XAI pipeline: TreeSHAP for global and local attribution on ensemble predictions; bootstrap SHAP variance estimates; grouped features for genomic pathways and lab trends; counterfactual generation (sparse counterfactuals) for actionable narratives (“If neutrophil-to-lymphocyte ratio had been X lower, predicted risk would fall by Y%”).
- Evaluation: (1) Fidelity: compare surrogate decision tree accuracy with original model ( $\Delta$ AUC), (2) Robustness: bootstrapped SHAP confidence intervals and perturbation tests, (3) Clinical utility: within a tumor board pilot, randomized case vignettes with vs. without explanations to measure change in treatment recommendations and confidence, (4) Fairness: subgroup AUC & explanation disparity by race/insurance status, (5) Reproducibility: share code and synthetic dataset; provide preprocessing scripts.

## Conclusions

In conclusion, explainable AI methods hold promise to unlock EHR-driven advances in oncology by improving interpretability, trust, and the safe translation of predictive models into clinical practice. The current practices have been characterized by rapid uptake of post-hoc explanation tools notably SHAP but insufficient evaluation, reproducibility, and equity assessment. For application of XAI on EHRs in oncology, future work must emphasize rigorous evaluation of explanations, multicenter validation, transparency, and human-centered design to ensure the clinically meaningful, equitable, and trustworthy support delivered by XAI.

## Abbreviations

DL: deep learning

EHRs: electronic health records

PDPs: partial dependence plots

RETAIN: Reverse Time Attention model

SHAP: SHapley Additive exPlanations

TCGA: The Cancer Genome Atlas

XAI: explainable artificial intelligence

## Declarations

### Author contributions

YY: Conceptualization, Investigation, Writing—original draft, Writing—review & editing. XL: Conceptualization, Investigation, Writing—original draft, Writing—review & editing. Both authors read and approved the submitted version.

### Conflicts of interest

The authors declare that they have no conflicts of interest.

### Ethical approval

Not applicable.

### Consent to participate

Not applicable.

### Consent to publication

Not applicable.

### Availability of data and materials

Not applicable.

### Funding

Not applicable.

### Copyright

© The Author(s) 2026.

## Publisher's note

Open Exploration maintains a neutral stance on jurisdictional claims in published institutional affiliations

and maps. All opinions expressed in this article are the personal views of the author(s) and do not represent the stance of the editorial team or the publisher.

## References

1. Hossain E, Rana R, Higgins N, Soar J, Barua PD, Pisani AR, et al. Natural Language Processing in Electronic Health Records in relation to healthcare decision-making: A systematic review. *Comput Biol Med.* 2023;155:106649. [DOI] [PubMed]
2. Yang S, Varghese P, Stephenson E, Tu K, Gronsbell J. Machine learning approaches for electronic health records phenotyping: a methodical review. *J Am Med Inform Assoc.* 2023;30:367–81. [DOI] [PubMed] [PMC]
3. Jain NM, Culley A, Knoop T, Micheel C, Osterman T, Levy M. Conceptual Framework to Support Clinical Trial Optimization and End-to-End Enrollment Workflow. *JCO Clin Cancer Inform.* 2019;3:1–10. [DOI] [PubMed] [PMC]
4. Xu Q, Xie W, Liao B, Hu C, Qin L, Yang Z, et al. Interpretability of Clinical Decision Support Systems Based on Artificial Intelligence from Technological and Medical Perspective: A Systematic Review. *J Healthc Eng.* 2023;2023:9919269. [DOI] [PubMed] [PMC]
5. Jee J, Fong C, Pichotta K, Tran TN, Luthra A, Waters M, et al.; MSK Cancer Data Science Initiative Group. Automated real-world data integration improves cancer outcome prediction. *Nature.* 2024;636:728–36. [DOI] [PubMed] [PMC]
6. Onitiu D, Wachter S, Mittelstadt B. How AI challenges the medical device regulation: patient safety, benefits, and intended uses. *J Law Biosci.* 2024;lsae007. [DOI]
7. Keyl J, Keyl P, Montavon G, Hosch R, Brehmer A, Mochmann L, et al. Decoding pan-cancer treatment outcomes using multimodal real-world data and explainable artificial intelligence. *Nat Cancer.* 2025;6:307–22. [DOI] [PubMed] [PMC]
8. Wesołowski S, Lemmon G, Hernandez EJ, Henrie A, Miller TA, Weyhrauch D, et al. An explainable artificial intelligence approach for predicting cardiovascular outcomes using electronic health records. *PLOS Digit Health.* 2022;1:e0000004. [DOI] [PubMed] [PMC]
9. Lauritsen SM, Kristensen M, Olsen MV, Larsen MS, Lauritsen KM, Jørgensen MJ, et al. Explainable artificial intelligence model to predict acute critical illness from electronic health records. *Nat Commun.* 2020;11:3852. [DOI] [PubMed] [PMC]
10. Caterson J, Lewin A, Williamson E. The application of explainable artificial intelligence (XAI) in electronic health record research: A scoping review. *Digit Health.* 2024;10:20552076241272657. [DOI] [PubMed] [PMC]
11. Payrovnaziri SN, Chen Z, Rengifo-Moreno P, Miller T, Bian J, Chen JH, et al. Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. *J Am Med Inform Assoc.* 2020;27:1173–85. [DOI] [PubMed] [PMC]
12. Borys K, Schmitt YA, Nauta M, Seifert C, Krämer N, Friedrich CM, et al. Explainable AI in medical imaging: An overview for clinical practitioners - Saliency-based XAI approaches. *Eur J Radiol.* 2023;162:110787. [DOI] [PubMed]
13. Kim YK, Koo JH, Lee SJ, Song HS, Lee M. Explainable Artificial Intelligence Warning Model Using an Ensemble Approach for In-Hospital Cardiac Arrest Prediction: Retrospective Cohort Study. *J Med Internet Res.* 2023;25:e48244. [DOI] [PubMed] [PMC]
14. Mujawar S, Sayed M, Chaware A. Empowering Decision Making in Healthcare-A Comparative Analysis of eXplainable AI Techniques Using SHAP and LIME for Cancer Patients Data. *Proceedings of 3rd Edition of IEEE Delhi Section Flagship Conference (DELCON); 2024 Nov 21–23; New Delhi, India. Piscataway: IEEE; 2025. pp. 1–5. [DOI]*



15. Hong JH, Ouh YT, Jeong S, Oh Y, Cho HW, Lee JK, et al. Integrative analysis of DNA methylation, RNA sequencing, and genomic variants in the cancer genome atlas (TCGA) to predict endometrial cancer recurrence. *Front Genet.* 2025;16:1569122. [DOI] [PubMed] [PMC]
16. Alabi RO, Mäkitie AA, Pirinen M, Elmusrati M, Leivo I, Almangush A. Comparison of nomogram with machine learning techniques for prediction of overall survival in patients with tongue cancer. *Int J Med Inform.* 2021;145:104313. [DOI] [PubMed]
17. Liu YM, Du YY, Song Y, Xiong HT, Yu HB, Li BH, et al. Predicting chemotherapy-induced myelosuppression in colorectal cancer: An interpretable, machine learning-based nomogram. *World J Gastroenterol.* 2025;31:112180. [DOI] [PubMed] [PMC]
18. Ghiasi MM, Zendehboudi S. Application of decision tree-based ensemble learning in the classification of breast cancer. *Comput Biol Med.* 2021;128:104089. [DOI] [PubMed]
19. Amiri Souri E, Chenoweth A, Cheung A, Karagiannis SN, Tsoka S. Cancer Grade Model: a multi-gene machine learning-based risk classification for improving prognosis in breast cancer. *Br J Cancer.* 2021;125:748–58. [DOI] [PubMed] [PMC]
20. Kwon BC, Choi MJ, Kim JT, Choi E, Kim YB, Kwon S, et al. RetainVis: Visual Analytics with Interpretable and Interactive Recurrent Neural Networks on Electronic Medical Records. *IEEE Trans Vis Comput Graph.* 2019;25:299–309. [DOI] [PubMed]
21. Finkelstein J, Smiley A, Echeverria C, Mooney K. Deep Learning Approaches to Forecast Physical and Mental Deterioration During Chemotherapy in Patients with Cancer. *Diagnostics (Basel).* 2025;15:956. [DOI] [PubMed]
22. Wang J, Qi M, Xiang Z, Tian Y, Tong D. SaraNet: Semantic aggregation reverse attention network for pulmonary nodule segmentation. *Comput Biol Med.* 2024;177:108674. [DOI] [PubMed]
23. Mayoral-Peña K, González Peña OI, Artzi N, de Donato M. Biomarker discovery for early breast cancer diagnosis using machine learning on transcriptomic data for biosensor development. *Comput Biol Med.* 2025;196:110584. [DOI] [PubMed]
24. Bouazza SH. A Novel Adaptive Sparse Deep Feature Selection Method for Enhanced Gene-Based Cancer Classification. *Eng Technol Appl Sci Res.* 2025;15:22494–9. [DOI]
25. Lee TF, Chao PJ, Ting HM, Chang L, Huang YJ, Wu JM, et al. Using multivariate regression model with least absolute shrinkage and selection operator (LASSO) to predict the incidence of Xerostomia after intensity-modulated radiotherapy for head and neck cancer. *PLoS One.* 2014;9:e89700. [DOI] [PubMed] [PMC]
26. Hao J, Kim Y, Kim TK, Kang M. PASNet: pathway-associated sparse deep neural network for prognosis prediction from high-throughput data. *BMC Bioinformatics.* 2018;19:510. [DOI] [PubMed] [PMC]
27. Aljohani A. Optimizing Patient Stratification in Healthcare: A Comparative Analysis of Clustering Algorithms for EHR Data. *Int J Comput Intell Syst.* 2024;17:173. [DOI]
28. Han S, Zhang T, Deng W, et al. Deep learning progressive distill for predicting clinical response to conversion therapy from preoperative CT images of advanced gastric cancer patients. *Sci Rep.* 2025;15:17092. [DOI] [PubMed] [PMC]
29. Tan AC, Naiman DQ, Xu L, Winslow RL, Geman D. Simple decision rules for classifying human cancers from gene expression profiles. *Bioinformatics.* 2005;21:3896–904. [DOI] [PubMed] [PMC]
30. Moldovanu S, Munteanu D, Biswas KC, Moraru L. Breast Lesion Detection Using Weakly Dependent Customized Features and Machine Learning Models with Explainable Artificial Intelligence. *J Imaging.* 2025;11:135. [DOI] [PubMed] [PMC]
31. Murphy DR, Laxmisan A, Reis BA, Thomas EJ, Esquivel A, Forjuoh SN, et al. Electronic health record-based triggers to detect potential delays in cancer diagnosis. *BMJ Qual Saf.* 2014;23:8–16. [DOI] [PubMed]
32. Zhuang Q, Zhang AY, Cong RSTY, Yang GM, Neo PSH, Tan DS, et al. Towards proactive palliative care in oncology: developing an explainable EHR-based machine learning model for mortality risk prediction. *BMC Palliat Care.* 2024;23:124. [DOI] [PubMed] [PMC]

33. Pan W, Hathi D, Xu Z, Zhang Q, Li Y, Wang F. Identification of predictive subphenotypes for clinical outcomes using real world data and machine learning. *Nat Commun.* 2025;16:3797. [DOI] [PubMed] [PMC]
34. Saleem M, Watson AE, Anwaar A, Jasser AO, Yusuf N. Optimizing Immunotherapy: The Synergy of Immune Checkpoint Inhibitors with Artificial Intelligence in Melanoma Treatment. *Biomolecules.* 2025;15:589. [DOI] [PubMed] [PMC]
35. Uppuluri V. AI-Driven Patient Recruitment Strategies for Healthcare Clinical Trials. *J Artif Intell Mach Learn Data Sci.* 2024;2:3038–44. [DOI]
36. Ni Y, Wright J, Perentesis J, Lingren T, Deleger L, Kaiser M, et al. Increasing the efficiency of trial-patient matching: automated clinical trial eligibility pre-screening for pediatric oncology patients. *BMC Med Inform Decis Mak.* 2015;15:28. [DOI] [PubMed] [PMC]
37. Denck J, Ozkirimli E, Wang K. Machine-learning-based adverse drug event prediction from observational health data: A review. *Drug Discov Today.* 2023;28:103715. [DOI] [PubMed]
38. Abbas Q, Jeong W, Lee SW. Explainable AI in Clinical Decision Support Systems: A Meta-Analysis of Methods, Applications, and Usability Challenges. *Healthcare (Basel).* 2025;13:2154. [DOI] [PubMed] [PMC]
39. Denaxas S, Direk K, Gonzalez-Izquierdo A, Pikoula M, Cakiroglu A, Moore J, et al. Methods for enhancing the reproducibility of biomedical research findings using electronic health records. *BioData Min.* 2017;10:31. [DOI] [PubMed]
40. Springate DA, Kontopantelis E, Ashcroft DM, Olier I, Parisi R, Chamapiwa E, et al. ClinicalCodes: an online clinical codes repository to improve the validity and reproducibility of research using electronic medical records. *PLoS One.* 2014;9:e99825. [DOI] [PubMed] [PMC]
41. Munteanu D, Moldovanu S, Miron M. The Explanation and Sensitivity of AI Algorithms Supplied with Synthetic Medical Data. *Electronics.* 2025;14:1270. [DOI]