Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Pathomics-based prediction of metastasis and survival in adenoid cystic carcinoma: a retrospective single-centre study

Jiang Xue
Peking University School and Hospital of Stomatology
Yue Lou
Peking University School and Hospital of Stomatology
Longping Liu
Peking University School and Hospital of Stomatology
Ye Zhang
Beijing Chao-Yang Hospital, Capital Medical University
Lisha Sun
Peking University School and Hospital of Stomatology
Tiejun Li
litiejun22@vip.sina.com
Peking University School and Hospital of Stomatology

Research Article

Keywords: Adenoid cystic carcinoma, Digital pathology, Whole-slide imaging, Weakly supervised learning, Metastasis, Survival prediction

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11697628/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Pathomics-based prediction of metastasis and survival in adenoid cystic carcinoma: a retrospective single-centre study

Jiang Xue, MD, PhD [1, 3, #], Yue Lou, PhD [1, 3, #], Longping Liu, PhD [1, 3, #], Ye Zhang, PhD [4], Lisha Sun, PhD [2, 3, *], Tiejun Li, MD, PhD [1, 3, *]

[1] Department of Oral Pathology, Peking University School and Hospital of Stomatology & National Center of Stomatology & National Clinical Research Center for Oral Diseases &National Engineering Research Center of Oral Biomaterials and Digital Medical Devices, Beijing, China

[2] Central laboratory, Peking University School and Hospital of Stomatology, Beijing, China

[3] Research Unit of Precision Pathologic Diagnosis in Tumors of the Oral and Maxillofacial Regions, Chinese Academy of Medical Sciences, Beijing, China

[4] Department of Stomatology, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China

# Contributed equally.

[*]**Correspondence:**

Tiejun Li

litiejun22@vip.sina.com

Lisha Sun

lisa_sun@bjmu.edu.cn

Abstract

Background: Adenoid cystic carcinoma (ACC) is a heterogeneous salivary gland malignancy with an unpredictable risk of distant metastasis and long-term survival outcomes. We explored whether features derived from routine histopathological whole-slide images could support patient-level prediction of metastasis and survival in a single-centre cohort.

Methods: We retrospectively collected 316 haematoxylin and eosin-stained sections from 71 patients who underwent surgery for ACC at Peking University Hospital of Stomatology between January 2010 and December 2020. Whole-slide images were tiled into 512 × 512-pixel patches. Patients were allocated to training (n = 50) and internal validation (n = 21) cohorts. ResNet50 and CrossFormer were evaluated using patient-level metastasis labels assigned under a weakly supervised framework. Patch-level outputs were aggregated into patient-level features for metastasis classification and Cox proportional-hazards modelling of overall survival (OS) and recurrence-free survival (RFS).

Results: Patch-level discrimination in the internal validation cohort was limited (area under the receiver-operating-characteristic curve [AUC], 0.507 for ResNet50 and 0.524 for CrossFormer). Patient-level validation AUCs for metastasis were 0.818 for the support vector machine (95% confidence interval, 0.627-1.000), 0.709 for random forest, and 0.623 for ExtraTrees. Validation concordance indices for overall and recurrence-free survival were 0.642 and 0.725 for the pathomics model and 0.717 and 0.725 for the combined model, respectively.

Conclusions: In this single-centre exploratory study, pathomics-derived features showed potential for supporting risk modelling in ACC. Larger multi-institutional cohorts and external validation are required before the models can be considered for clinical use.

Keywords: Adenoid cystic carcinoma; Digital pathology; Whole-slide imaging; Weakly supervised learning; Metastasis; Survival prediction

Background

Adenoid cystic carcinoma (ACC) is a salivary gland malignancy with an indolent but potentially prolonged clinical course [1, 2]. Late distant metastasis and heterogeneous biological behaviour complicate management [3, 4]. ACC often grows infiltratively and may recur locally or metastasize, most commonly to the lungs [5, 6, 7]. Survival estimates vary by case mix and follow-up duration [8, 9], supporting the need for better tools to characterize individual risk. Histological features and clinical factors contribute to risk assessment in ACC, although outcomes remain heterogeneous [9, 10]. MYB-NFIB fusions and miRNA signatures have also been investigated [11, 12]. Quantitative image analysis may provide complementary information from routinely available H&E-stained sections.

Recent advances in digital pathology and deep learning enable quantitative analysis of histological images at scale [13, 14]. Whole-slide imaging can capture tumour architecture, stroma, and spatial heterogeneity that may be associated with clinical outcomes [15]. Although deep-learning approaches have been explored for prognostic prediction in head and neck cancers [16, 17], data on ACC remain limited. We therefore explored a weakly supervised pathomics framework for patient-level metastasis and survival prediction in ACC. The study was designed as a single-centre, internally validated feasibility analysis rather than a clinical-decision tool.

Methods

Study design and data collection

We developed a pathomics-based model to explore associations between H&E-stained image features and metastasis and survival outcomes in ACC. All cases were diagnosed pathologically and reviewed by two experienced head and neck pathologists according to World Health Organization criteria. We retrospectively collected 71 ACC cases treated at Peking University School and Hospital of Stomatology between January 2010 and December 2020. Patients were allocated at the patient level to a training cohort (50 patients; 221 slides) and an internal validation cohort (21 patients; 95 slides); randomization was repeated until baseline clinical

features showed no statistically significant differences between the cohorts, without reference to outcome status. After excluding faded or unclear H&E sections, 316 sections were included. The median follow-up for the full cohort was 55 months. Age, sex, tumour site, and histological pattern were obtained from medical records, and metastasis and survival outcomes were recorded during follow-up. The study was approved by the Ethics Committee of Peking University School and Hospital of Stomatology (PKUSSIRB-2025119296) and conducted in accordance with the Declaration of Helsinki. Informed consent to participate in the study was obtained from all participants or their legal representatives. All H&E slides were scanned as MRXS whole-slide images using a Pannoramic III scanner (3DHISTECH, Hungary) with a 20× objective lens. Fig. 1 outlines the study workflow.

Pathology procedure

Data processing

Whole-slide images (WSIs) were segmented into 512 × 512-pixel patches at 20× magnification. Patches consisting predominantly of bright background were excluded. Macenko normalization was used to standardize colour representation [18]. This preprocessing generated more than 1.2 million tissue-containing patches. Processing was performed on the OnekeyAI Platform using OKT-crop_WSI2patch, OKT-patch2predict, and OKT-patch_normalize. Additional file 1 describes patch-level training and image preparation.

Patch-level deep learning model training

Our pipeline combined patch-level prediction with slide-level feature aggregation. Because metastasis status was available only at the patient level, that label was assigned to patches from the corresponding patient for weakly supervised training. This approach can introduce label noise, particularly in non-tumour or non-informative tissue regions; background removal and subsequent slide-level aggregation were used to reduce, but not eliminate, this limitation. ResNet50 [19] and the transformer-based CrossFormer [20] were evaluated. Patch-level training is described in Additional file 1.

Patch-level aggregation for wsi fusion

After initial training, predicted labels and probabilities were generated for each patch and aggregated to the WSI and patient levels. Feature aggregation is described in Additional file 2.

Model building

Pathomics model

Using patch-level aggregation, we extracted 206 patient-level features: two sets of 101 probability-histogram features and four additional features derived from aggregated patch labels and probabilities. For survival modelling, feature screening and model tuning were performed within the training cohort. In each outer cross-validation fold, candidate features were screened using univariable Cox regression ($p < 0.05$); LASSO-Cox regression was then fitted with the regularization parameter tuned by 10-fold cross-validation on partial-likelihood deviance. Retained features entered multivariable Cox models for OS and RFS. The fitted models were evaluated in the internal validation cohort. Given the limited sample size, all estimates are exploratory.

Clinical model

Using the training cohort, the clinical model was developed with Cox proportional-hazards regression based on selected clinical features.

Combined model

To explore whether clinical variables added prognostic information to pathomics features, selected clinical and pathomics features were entered into multivariable Cox proportional-hazards models.

Survival analysis

X-tile software [21] was used to derive candidate cut-points for stratifying model-derived risk scores into high- and low-risk groups. Kaplan-Meier curves were compared with the log-rank test. These analyses are descriptive and should be interpreted in the context of the small cohort.

Statistical analysis

The Shapiro-Wilk test assessed normality of continuous clinical variables. Normally distributed variables were compared with t-tests and non-normally distributed variables with Mann-Whitney U tests. Categorical variables were compared using chi-square tests. Baseline characteristics are presented in Table 1. Analyses and machine-learning development were conducted using Python 3.7.12, Onekey 3.3.5, and scikit-learn 1.0.2. Model training used an NVIDIA GeForce RTX 4090 GPU and PyTorch 1.8.1.

Use of generative artificial intelligence

OpenAI ChatGPT/Codex was used during manuscript preparation to assist with English-language editing, organization, and identification of reporting inconsistencies. This assistance did not generate patient data or model performance estimates. The authors are responsible for verifying the final text and its consistency with the original analyses.

Results

This study included ACC tissue samples and clinicopathological data from 71 patients treated at Peking University Hospital of Stomatology. Fifty patients were allocated to the training cohort and 21 to the internal validation cohort. Baseline characteristics are summarized in Table 1; the allocation procedure sought baseline balance, so nonsignificant differences should not be taken as independent evidence of comparability. The study workflow is shown in Fig. 1.

Pathomics signature based on patch-level aggregation

Patch-level performance

We evaluated ResNet50 and CrossFormer using patch-level outputs generated under weak supervision. ResNet50 is a convolutional neural network, whereas CrossFormer is a vision transformer with cross-scale attention [20]. Patch-level ROC curves and WSI-level visualizations were generated to characterize model behaviour. As shown in Table 2 and Fig. 2, validation performance was limited for both models (validation AUC, 0.507 for ResNet50 and 0.524 for CrossFormer). CrossFormer was used for subsequent feature aggregation because its

validation AUC was marginally higher; this selection should not be interpreted as evidence of robust generalizability or accurate tumour-region identification. The probability-level histogram and term frequency-inverse document frequency (TF-IDF)-weighted bag-of-words pipelines therefore represent exploratory feature-engineering approaches based on CrossFormer outputs.

Model attribution and spatial visualization

To explore model attribution, we generated class-activation maps for representative samples using gradient-weighted class activation mapping (Grad-CAM) for the convolutional model [22]. For ResNet50, maps were derived from the final convolutional layer; for CrossFormer, they were derived from self-attention maps. The maps summarize spatial variation in model output after aggregation of tile-level scores (Fig. 3). Blue-to-red colour gradients indicate increasing model-assigned weight for metastasis-associated image patterns. They do not constitute manual tumour annotations or establish histological specificity. A total of 344,527 patches were represented in these visualizations.

Fig. 4 presents the t-distributed stochastic neighbor embedding (t-SNE) visualization of the aggregated image features (A), the regularization coefficient profiles (B), and the cross-validation curve (C).

Metastasis analysis

For exploratory patient-level metastasis prediction, support vector machine (SVM), random forest, and ExtraTrees classifiers were trained using aggregated features from the 50-patient training cohort. Internal validation AUCs were 0.818 (95% CI, 0.627-1.000) for SVM, 0.709 (95% CI, 0.498-0.921) for random forest, and 0.623 (95% CI, 0.376-0.870) for ExtraTrees (Table 3). SVM had the highest observed validation AUC among the evaluated classifiers. The substantial training-to-validation decrease for ExtraTrees (0.926 to 0.623) suggests overfitting. The wide confidence intervals preclude firm conclusions about comparative performance (Fig. 5).

Survival analysis

Survival prediction models were developed in the training cohort and evaluated in the internal validation cohort. In the training cohort, C-indices were 0.682 and 0.729 for the pathomics model for OS and RFS, respectively, and 0.570 and 0.612 for the clinical model. In the internal validation cohort, the corresponding C-indices were 0.642 and 0.725 for the pathomics model and 0.632 and 0.560 for the clinical model (Table 4). These values describe internal performance only.

The combined model yielded C-indices of 0.701 for OS and 0.729 for RFS in the training cohort and 0.717 for OS and 0.725 for RFS in the internal validation cohort. Thus, the combined model showed a higher C-index than the pathomics model for OS in the internal validation cohort, whereas no increase was observed for RFS. Formal comparison and external validation are required before concluding that clinical variables improve model performance.

We also performed exploratory 3-year risk classification analyses. AUC estimates are shown in Tables 5 and 6. In the internal validation cohort, the AUC was 0.641 for pathomics-only OS prediction and 0.719 for combined OS prediction; for RFS prediction, pathomics-only and combined models both yielded an AUC of 0.850. Owing to wide confidence intervals and the small validation cohort, these estimates should be viewed as hypothesis-generating. Fig. 6 shows Kaplan-Meier curves, and Fig. 7 shows ROC curves for the internal analyses.

We constructed exploratory nomograms integrating selected clinical features and the pathomics signature using Cox proportional-hazards models. Figures 8 and 9 illustrate the nomograms for OS and RFS. They are intended to display the fitted models and are not ready for clinical implementation without external validation and calibration assessment.

Discussion

This exploratory single-centre study assessed whether deep-learning-derived pathomics features from routine H&E whole-slide images could support metastasis and survival modelling in ACC. The results provide a feasibility signal rather than a validated clinical prediction tool. Patch-level

discrimination was limited in the internal validation cohort, and the patient-level metastasis and survival estimates were based on a small number of patients.

Weakly supervised approaches can reduce the need for pixel-level annotations, which is particularly attractive in rare tumours such as ACC [23]. In this study, patient-level metastasis labels were assigned to image patches and then aggregated at the slide and patient levels. This design reduces annotation burden but also introduces label noise and prevents direct claims about tumour localisation. The pathomics and clinical features were therefore examined as complementary exploratory predictors, not as independently established biomarkers.

CrossFormer was selected for feature aggregation because it had a marginally higher validation AUC than ResNet50. However, the validation AUC of 0.524 indicates limited discrimination. It is therefore not possible to infer that the model captured biologically meaningful global context or molecular mechanisms from the present data. Future work should test these hypotheses with independently annotated slides and molecularly characterised cohorts.

Several limitations should be emphasized. First, the retrospective, single-centre design and small cohort (n = 71) limit precision and generalizability. Second, the validation cohort was internal and contained only 21 patients. Third, weak supervision assigned patient-level labels to all patches, so attribution maps cannot be interpreted as tumour annotations. Fourth, calibration, clinical utility, and transportability require assessment in independent multi-institutional cohorts. Finally, the biological correlates of the derived image features remain unknown. Future studies should incorporate external validation, adequately reported event counts, calibration analyses, and molecular or microenvironmental correlates. In addition, CrossFormer was selected using validation performance, and the patient split was repeated to seek baseline balance. These choices limit the independence of the internal validation estimates. Patch-level confidence intervals also need to account for clustering within patients. Binary metastasis classification does not explicitly account for differing follow-up durations or censoring.

At present, the nomograms should be regarded as research visualizations of internally fitted models. They are not sufficiently validated to guide surveillance intensity, adjuvant treatment, or systemic therapy decisions. Their potential clinical value should be assessed prospectively only after external validation and calibration.

Conclusions

In conclusion, this single-centre feasibility study suggests that pathomics-derived features from H&E whole-slide images warrant further investigation for risk modelling in ACC. Independent external validation, transparent model reporting, and assessment of calibration and clinical utility are required before any clinical application.

Abbreviations

ACC, adenoid cystic carcinoma; AUC, area under the receiver-operating-characteristic curve; BoW, bag of words; CI, confidence interval; CNN, convolutional neural network; Grad-CAM, gradient-weighted class activation mapping; H&E, haematoxylin and eosin; KM, Kaplan-Meier; LASSO, least absolute shrinkage and selection operator; NPV, negative predictive value; OS, overall survival; PLH, probability-level histogram; PPV, positive predictive value; ResNet50, Residual Network 50; RFS, recurrence-free survival; ROC, receiver-operating characteristic; SVM, support vector machine; TF-IDF, term frequency-inverse document frequency; WSI, whole-slide image.

Ethics approval and consent to participate

Human Ethics and Consent to Participate declarations: This study was approved by the Ethics Committee of Peking University School and Hospital of Stomatology (approval number: PKUSSIRB-2025119296) and was conducted in accordance with the Declaration of Helsinki. Informed consent to participate in the study was obtained from all participants or their legal representatives.

Clinical trial number

Clinical trial number: not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Research Unit of Precision Pathologic Diagnosis in Tumors of the Oral and Maxillofacial Regions, Chinese Academy of Medical Sciences (2019RU034), and the National Natural Science Foundation of China (81671006, 81300894, 30901680).

Authors' contributions

LS and TL contributed to conceptualization and study design. LL, JX, and YL contributed to methodology and manuscript review and revision. LL, JX, YZ, and YL contributed to data acquisition, analysis, and interpretation. LL and JX provided technical and material support.

Acknowledgements

Not applicable.

Additional files

Additional file 1. Additional_file_1.docx. Patch-level model training. Image preparation, augmentation, architectures, and optimization procedures.

Additional file 2. Additional_file_2.docx. Multiple-instance feature aggregation for whole-slide images. Probability histograms, bag-of-words representations, and feature concatenation.

References

1. Bradley PJ. Adenoid cystic carcinoma of the head and neck: a review. Curr Opin Otolaryngol Head Neck Surg. 2004;12(2):127-132.

2. Renehan A, Gleave EN, Hancock BD, Smith P, McGurk M. Long-term follow-up of over 1000 patients with salivary gland tumours treated in a single centre. Br J Surg. 1996;83(12):1750-1754.

3. Zhou J, Zhao G, Wang S, Li N. Systemic therapy in the management of metastatic or locally recurrent adenoid cystic carcinoma of the salivary glands: a systematic review of the last decade. Br J Cancer. 2024;131(6):1021-1031.

4. Lorini L, Ardighieri L, Bozzola A, Romani C, Bignotti E, Buglione M, et al. Prognosis and management of recurrent and/or metastatic head and neck adenoid cystic carcinoma. Oral Oncol. 2021;115:105213.

5. Oplatek A, Ozer E, Agrawal A, Bapna S, Schuller DE. Patterns of recurrence and survival of head and neck adenoid cystic carcinoma after definitive resection. Laryngoscope. 2010;120(1):65-70.

6. He S, Li P, Zhong Q, Hou L, Yu Z, Huang Z, et al. Clinicopathologic and prognostic factors in adenoid cystic carcinoma of head and neck minor salivary glands: A clinical analysis of 130 cases. Am J Otolaryngol. 2017;38(2):157-162.

7. Bhayani MK, Yener M, El-Naggar A, Garden A, Hanna EY, Weber RS, et al. Prognosis and risk factors for early-stage adenoid cystic carcinoma of the major salivary glands. Cancer. 2012;118(11):2872-2878.

8. Stawarz K, Durzynska M, Gałązka A, Gorzelnik A, Zwolinski J, Paszkowska M, et al. Current landscape and future directions of therapeutic approaches for adenoid cystic carcinoma of the salivary glands (Review). Oncol Lett. 2025;29(3):153.

9. Xu B, Drill E, Ho A, Ho A, Dunn L, Prieto-Granada CN, et al. Predictors of Outcome in Adenoid Cystic Carcinoma of Salivary Glands: A Clinicopathologic Study With Correlation Between MYB Fusion and Protein Expression. Am J Surg Pathol. 2017;41(10):1422-1432.

10. Ikawa H, Koto M, Takagi R, Ebner DK, Hasegawa A, Naganawa K, et al. Prognostic factors of adenoid cystic carcinoma of the head and neck in carbon-ion radiotherapy: The impact of histological subtypes. Radiother Oncol. 2017;123(3):387-393.

11. Mitani Y, Li J, Rao PH, Zhao YJ, Bell D, Lippman SM, et al. Comprehensive analysis of the MYB-NFIB gene fusion in salivary adenoid cystic carcinoma: Incidence, variability, and clinicopathologic significance. Clin Cancer Res. 2010;16(19):4722-4731.

12. Li Y, Gu F, He Y, Xiang J, Huang X, Wang Y, et al. MicroRNA in adenoid cystic carcinoma (Review). Int J Oncol. 2023;62(1):17.

13. Prasad P, Khair AMB, Saeed M, Shetty N. Artificial intelligence in histopathology. J Pharm Bioallied Sci. 2024;16(Suppl 5):S4226-S4229.

14. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88.

15. Dang C, Qi Z, Xu T, Gu M, Chen J, Wu J, et al. Deep Learning-Powered Whole Slide Image Analysis in Cancer Pathology. Lab Invest. 2025;105(7):104186.

16. Bitto V, Jiang X, Baumann M, Kather JN, Kurth I. Deep Learning Predicts Survival Across Squamous Tumor Entities From Routine Pathology: Insights From Head and Neck, Esophagus, Lung, and Cervical Cancer. Mod Pathol. 2025;38(12):100845.

17. Peng J, Lu Y, Chen L, Qiu K, Chen F, Liu J, et al. The prognostic value of machine learning techniques versus Cox regression model for head and neck cancer. Methods. 2022;205:123-132.

18. Macenko M, Niethammer M, Marron JS, Borland D, Woosley JT, Guan X, et al. A method for normalizing histology slides for quantitative analysis. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. 2009. p. 1107-1110.

19. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. p. 770-778.

20. Wang W, Yao L, Chen L, Lin B, Cai D, He X, et al. CrossFormer: a versatile vision transformer hinging on cross-scale attention. In: International Conference on Learning Representations. 2022. https://openreview.net/forum?id=_PHymLIxuI.

21. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res. 2004;10(21):7252-7259.

22. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. Int J Comput Vis. 2020;128(2):336-359.

23. Li B, Li Y, Eliceiri KW. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021. p. 14318-14328.

Figures

Fig. 1. Study workflow. Workflow in developing metastasis and survival prediction model of adenoid cystic carcinoma based on deep learning.

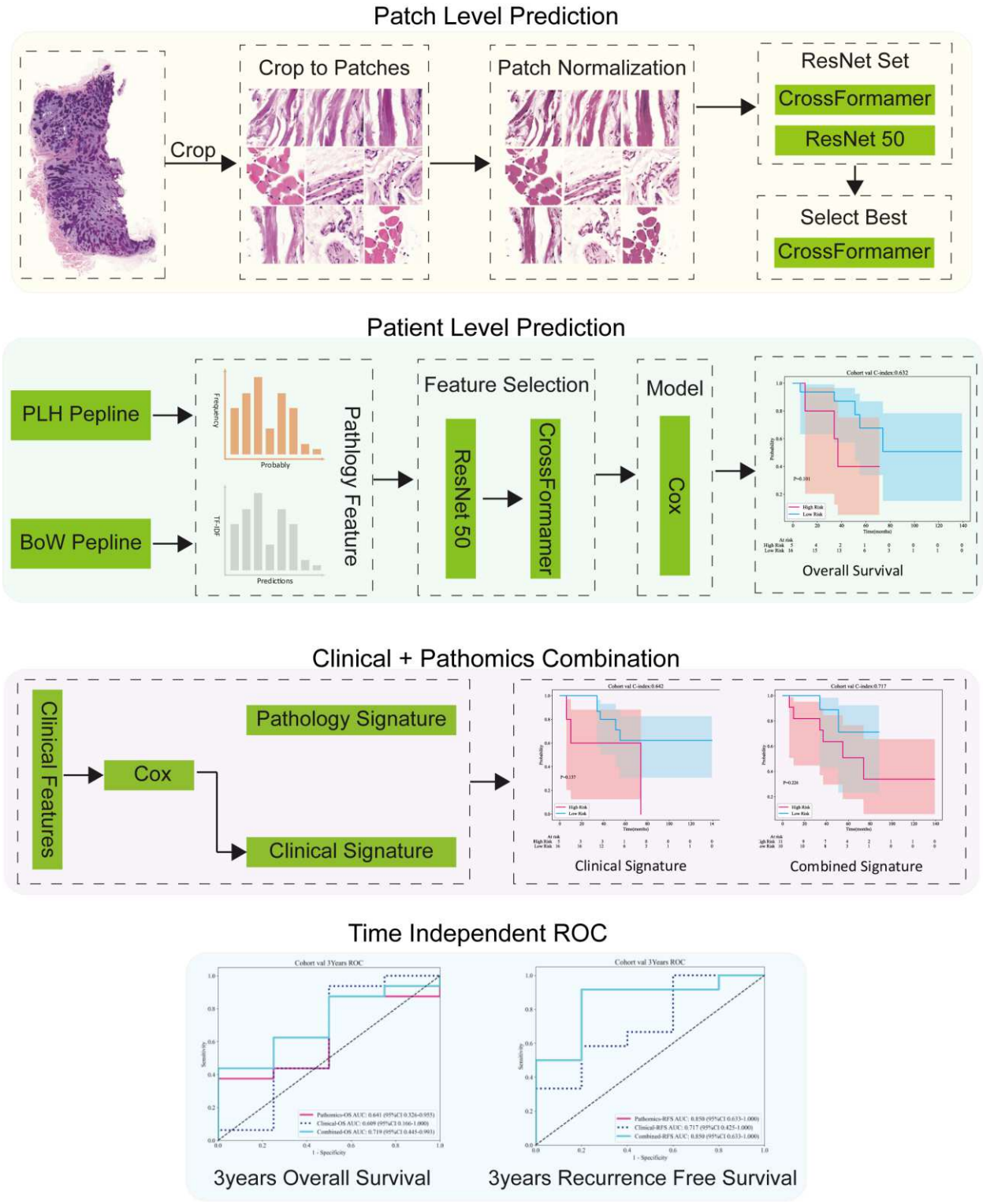
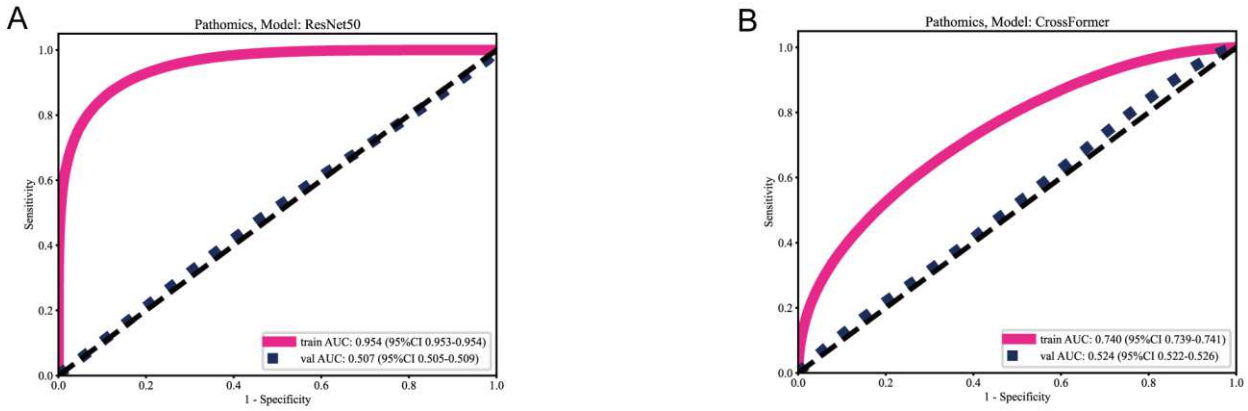


Fig. 2. Patch-level discrimination. Patch-level ROC curves for ResNet50 (A) and CrossFormer (B) in the training and internal validation cohorts. These curves reflect weakly supervised patient-level metastasis labels assigned to patches.



352

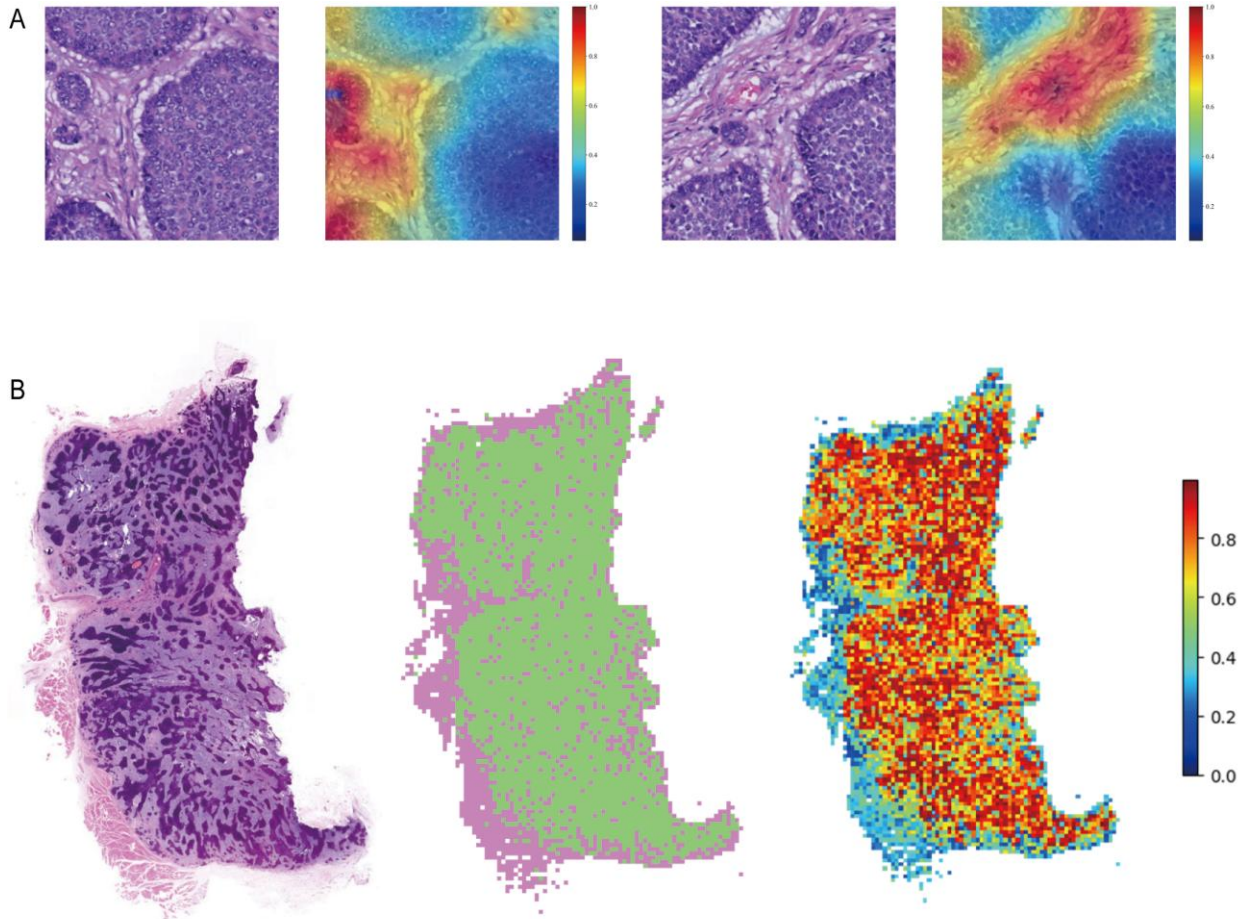
353

354 Fig. 3. Spatial visualization of model output. Representative spatial visualizations of model

355 output for samples 100794-3 (A) and 144383-1 (B). Blue-to-red gradients indicate increasing

356 model-assigned weight for metastasis-associated image patterns; the maps are not manual

357 tumour annotations. Specimen identifiers are de-identified.



358

359

360

Fig. 4. Feature visualization and LASSO-Cox selection. Visualization of patch-level aggregation features using t-SNE dimensionality reduction (A). Coefficients from 10-fold cross-validation (B). Partial likelihood deviance from 10-fold cross-validation in the LASSO-Cox model (C).

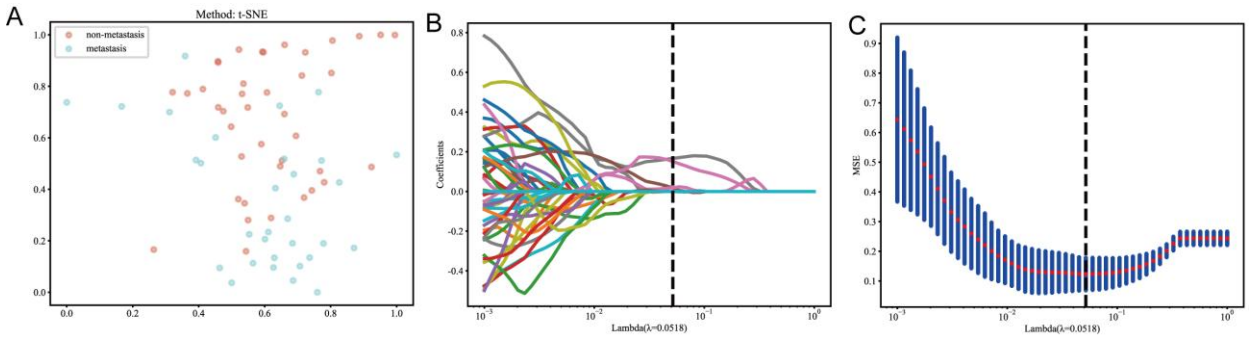


Fig. 5. Patient-level metastasis discrimination. ROC curves for exploratory patient-level metastasis prediction using ExtraTrees and SVM classifiers in the training (A) and internal validation (B) cohorts. Validation performance should be interpreted cautiously because of the small cohor

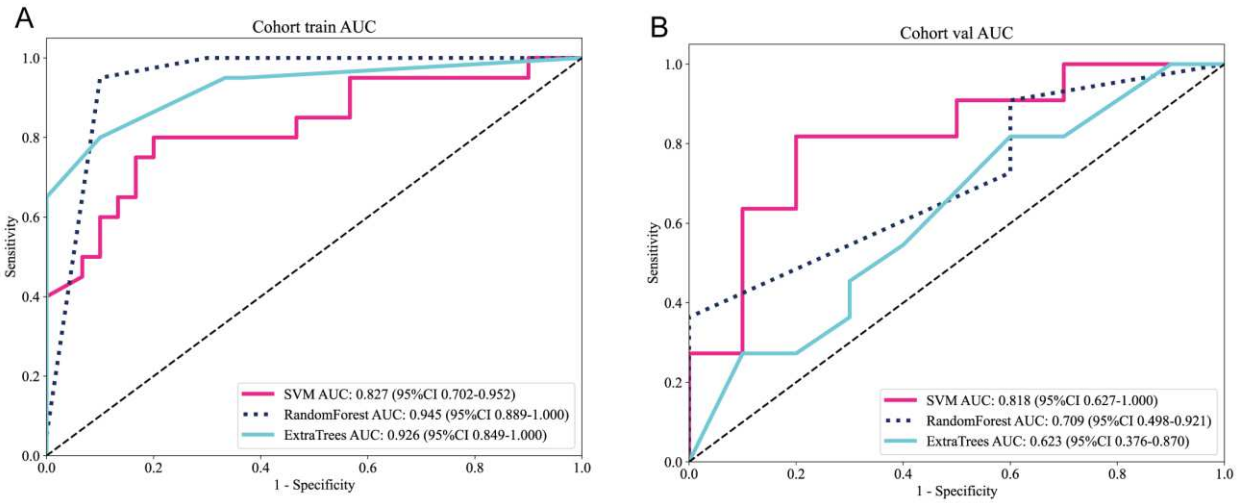


Fig. 6. Kaplan-Meier analyses. Kaplan-Meier analysis for the training and internal validation cohorts. From left to right, the panels show curves for the clinical, pathomics, and combined

models for OS (A-F) and RFS (G-L). These internally derived risk-group comparisons are exploratory.

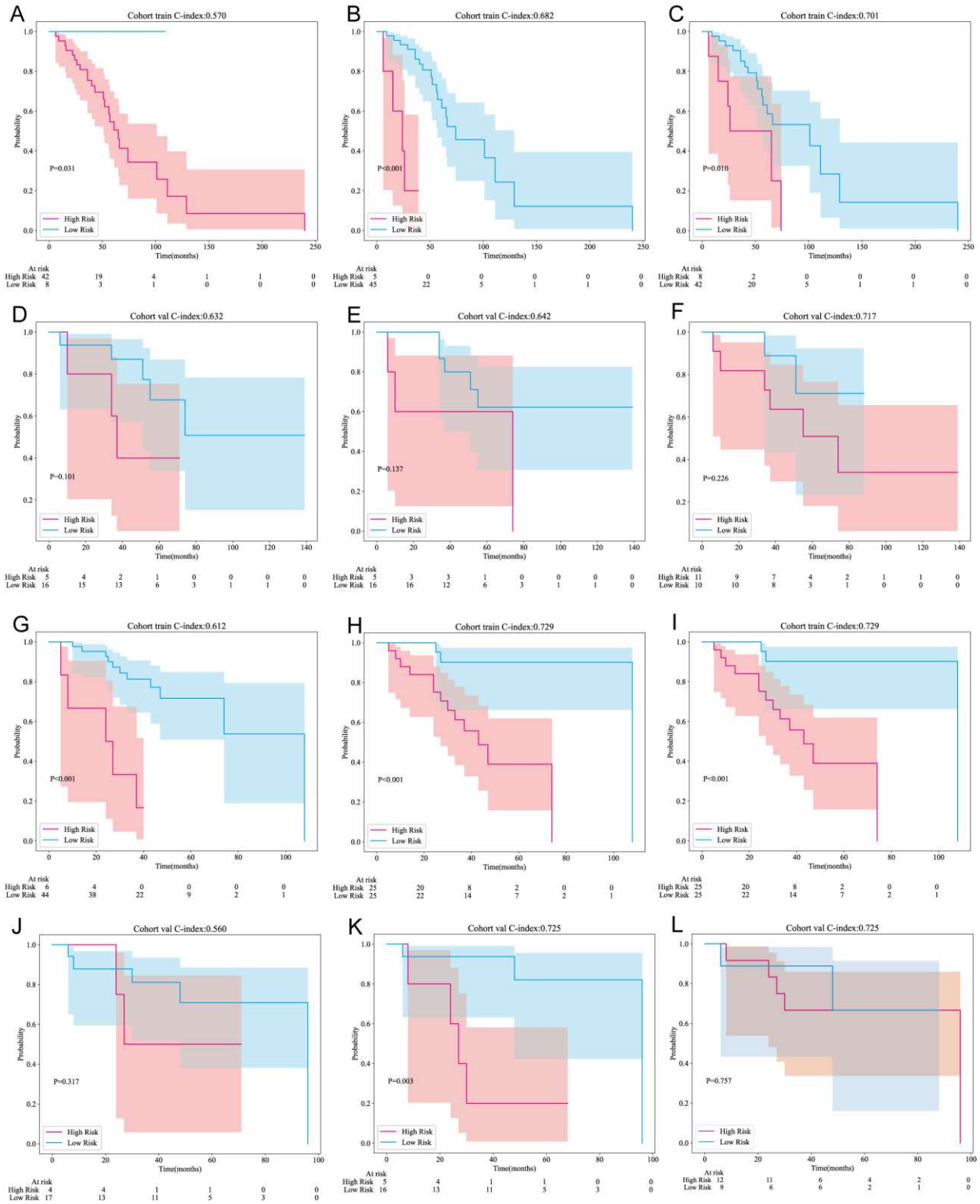
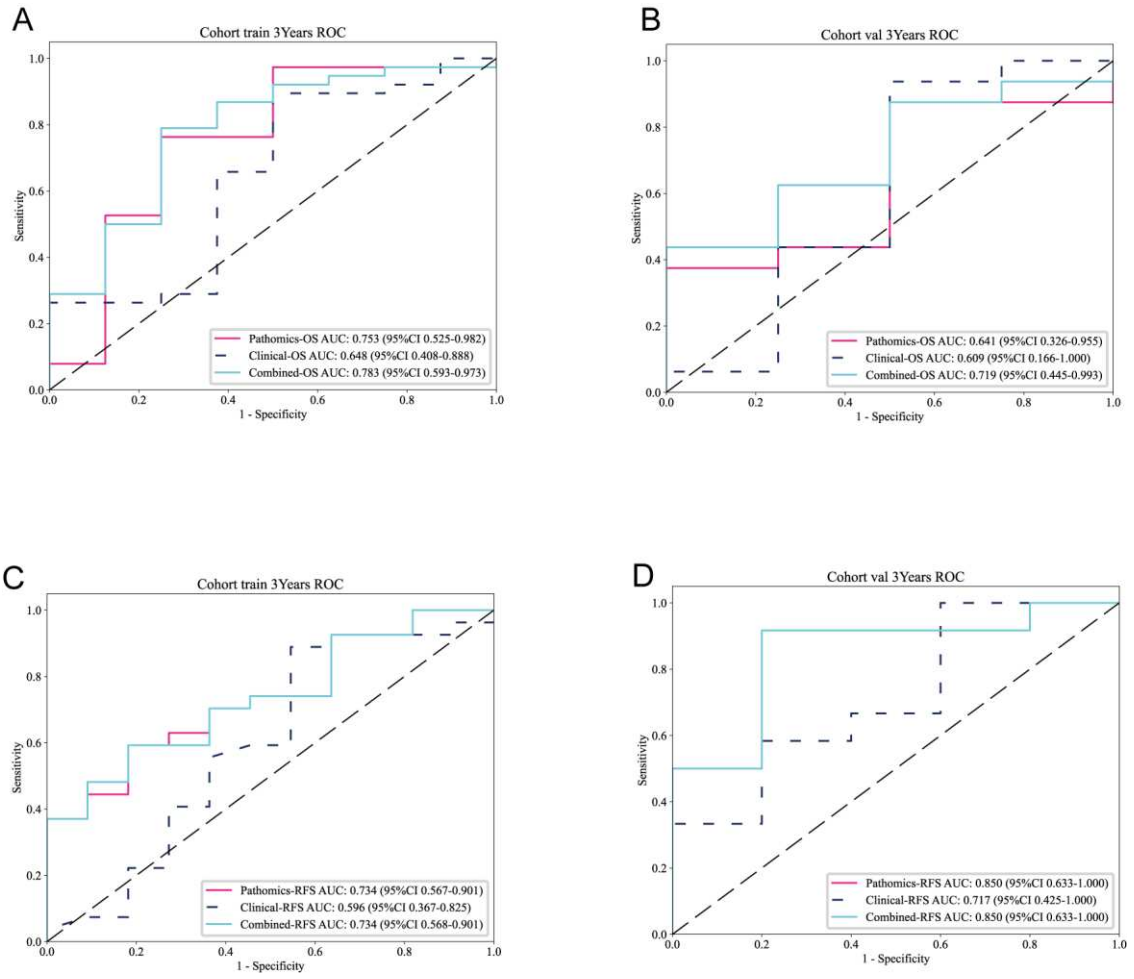
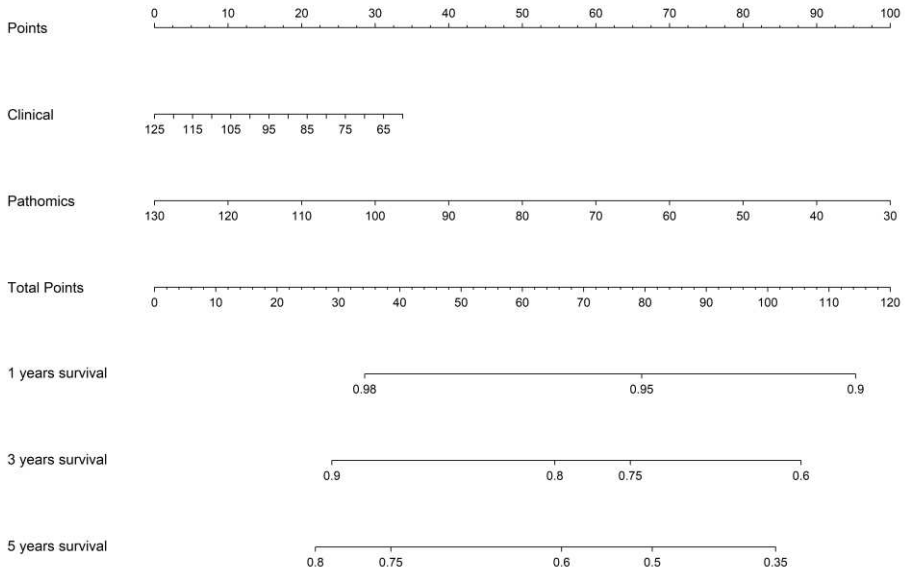


Fig. 7. Three-year survival discrimination. ROC curves for exploratory 3-year OS and RFS prediction in the training and internal validation cohorts. The figure displays discrimination estimates and does not establish clinical utility.



389 Fig. 8. Overall-survival nomogram. Exploratory nomogram for overall survival based on the

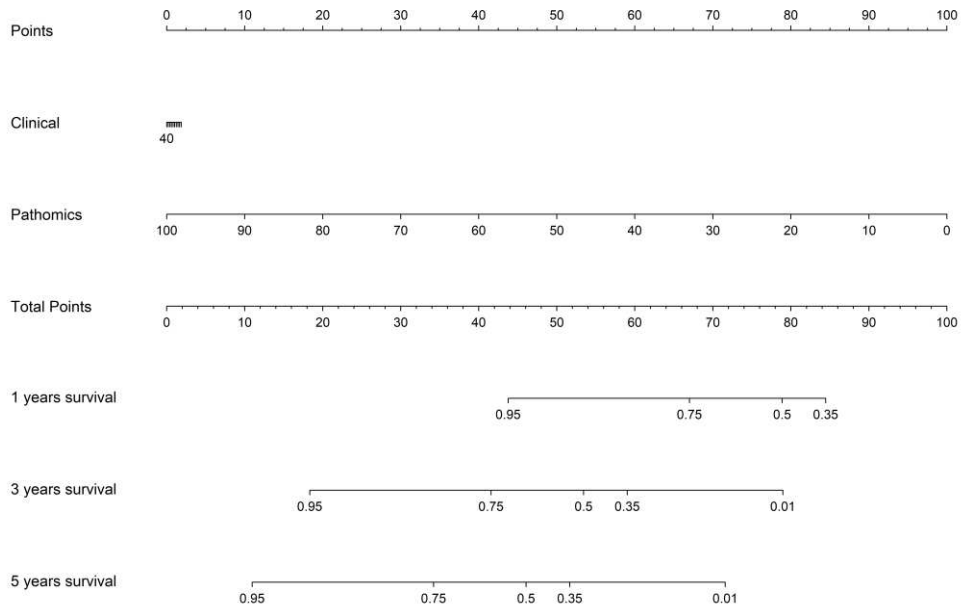
390 internally fitted model.



391

392 Fig. 9. Recurrence-free-survival nomogram. Exploratory nomogram for recurrence-free survival

393 based on the internally fitted model.



394

Table 1. Baseline characteristics of the full cohort and the training and internal validation cohorts.

Items	All patients (%)	Training cohort (%)	Validation cohort (%)	*P* value
Age (years)	52.44±13.55	50.74±13.71	56.48±12.55	0.104
Gender				0.892
Female	33(46.48)	24(48.00)	9(42.86)	
Male	38(53.52)	26(52.00)	12(57.14)	
Histological patterns				0.852
Tubular & cribriform	37(52.11)	25(50.00)	12(57.14)	
Solid	27(38.03)	20(40.00)	7(33.33)	
Mixed	7(9.86)	5(10.00)	2(9.52)	
Location				0.889
Cheek area	4(5.63)	2(4.00)	2(9.52)	
Palatal gland	27(38.03)	18(36.00)	9(42.86)	
Parotid gland	5(7.04)	4(8.00)	1(4.76)	
Sublingual gland	20(28.17)	14(28.00)	6(28.57)	
Submandibular gland	13(18.31)	10(20.00)	3(14.29)	
Labia	1(1.41)	1(2.00)	null	
Peri-orbital area	1(1.41)	1(2.00)	null	

Table 2. Patch-level performance under weak supervision in the training and internal validation cohorts.

Model Name	Acc	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
Resnet50	0.880	0.954	0.9531-0.9539	0.865	0.887	0.789	0.931	train
Resnet50	0.508	0.507	0.5051-0.5090	0.463	0.558	0.544	0.477	validation
CrossFormer	0.681	0.740	0.7387-0.7407	0.629	0.707	0.512	0.796	train
CrossFormer	0.543	0.524	0.5223-0.5261	0.863	0.179	0.545	0.536	validation

ResNet50, Residual Network 50; AUC, area under the receiver-operating-characteristic curve; Acc, accuracy; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval. Patch-level estimates reflect weakly supervised patient-level metastasis labels assigned to patches.

Table 3. Exploratory patient-level metastasis prediction performance in the training and internal validation cohorts.

Model name	Acc	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
SVM	0.780	0.827	0.702 - 0.952	0.750	0.8	0.714	0.828	train
SVM	0.762	0.818	0.627 - 1.000	0.727	0.8	0.800	0.727	validation
RandomForest	0.620	0.945	0.889 - 1.000	0.050	1.0	1.000	0.612	train
RandomForest	0.476	0.709	0.498 - 0.921	0.000	1.0	NA	0.476	validation
ExtraTrees	0.860	0.926	0.849 - 1.000	0.650	1.0	1.000	0.811	train
ExtraTrees	0.571	0.623	0.376 - 0.870	0.545	0.6	0.600	0.545	validation

AUC, area under the receiver-operating-characteristic curve; Acc, accuracy; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval. Validation estimates should be interpreted cautiously because of the small cohort.

Table 4. Concordance indices of exploratory survival models in the training and internal validation cohorts.

Pathomics-OS	Clinical-OS	Combined-OS	Pathomics-RFS	Clinical-RFS	Combined-RFS	Cohort
0.682	0.570	0.701	0.729	0.612	0.729	train
0.642	0.632	0.717	0.725	0.560	0.725	validation

OS, overall survival; RFS, recurrence-free survival. Values describe internal performance estimates only.

Table 5. Exploratory 3-year overall-survival risk classification in the training and internal validation cohorts.

Signature	Acc	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
Pathomics-OS	0.739	0.753	0.5248 - 0.9818	0.737	0.75	0.933	0.375	train
Clinical-OS	0.804	0.648	0.4077 - 0.8884	0.868	0.50	0.892	0.444	train
Combined-OS	0.761	0.783	0.5931 - 0.9727	0.763	0.75	0.935	0.400	train
Pathomics-OS	0.450	0.641	0.3261 - 0.9551	0.312	1.00	1.000	0.267	validation
Clinical-OS	0.800	0.609	0.1661 - 1.0000	0.875	0.50	0.875	0.500	validation
Combined-OS	0.500	0.719	0.4448 - 0.9927	0.375	1.00	1.000	0.286	validation

AUC, area under the receiver-operating-characteristic curve; Acc, accuracy; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval. Validation estimates should be interpreted cautiously because of the small cohort.

Table 6. Exploratory 3-year recurrence-free-survival risk classification in the training and internal validation cohorts.

Signature	Acc	AUC	95% CI	Sensitivity	Specificity	PPV	NPV	Cohort
Pathomics-RFS	0.632	0.734	0.5669 - 0.9011	0.556	0.818	0.882	0.429	train
Clinical-RFS	0.737	0.596	0.3665 - 0.8254	0.852	0.455	0.793	0.556	train
Combined-RFS	0.632	0.734	0.5675 - 0.9005	0.556	0.818	0.882	0.429	train
Pathomics-RFS	0.824	0.850	0.6332 - 1.0000	0.833	0.800	0.909	0.667	validation
Clinical-RFS	0.765	0.717	0.4252 - 1.0000	0.917	0.400	0.786	0.667	validation
Combined-RFS	0.824	0.850	0.6332 - 1.0000	0.833	0.800	0.909	0.667	validation

AUC, area under the receiver-operating-characteristic curve; Acc, accuracy; PPV, positive predictive value; NPV, negative predictive value; CI, confidence interval. Validation estimates should be interpreted cautiously because of the small cohort.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile2.docx