Probabilistic Integration of Renal Cancer Radiology and Pathology Using Graph Neural Networks

Shangqi Gao[1,2†], Shangde Gao[4†], Ines Machado[1,3], and Mireia Crispin-Ortuzar[1,2,3⋆]

[1] Early Cancer Institute, University of Cambridge, Cambridge, UK
[2] Department of Oncology, University of Cambridge, Cambridge, UK
[3] Cancer Research UK Cambridge Centre, University of Cambridge, Cambridge, UK
[4] College of Computer Science and Technology, Zhejiang University, China
sg2162@cam.ac.uk, gaosde@zju.edu.cn, im549@cam.ac.uk, mc973@cam.ac.uk
https://github.com/CrispinLab

Abstract. Kidney tumors can be highly heterogeneous from the microscopic to the macroscopic scale. To address this, we propose a sparsity-informed probabilistic integration of radiomics and pathomics for kidney cancer analysis. We construct radiology and pathology graphs to model spatial correlations, then use a probabilistic method and graph neural networks to identify biomarkers and aggregate spatial features. Our validation shows that this integrated approach significantly outperforms traditional methods in kidney survival analysis, achieving a notable improvement of 0.084 in the concordance index, enabling better prognostic assessments for kidney cancer patients. The source code has been released by https://github.com/shangqigao/RadioPath.

Keywords: Kidney Cancer · Multimodal Data Integration · Radiomics · Pathomics · Graph Neural Networks

1 Introduction

Kidney cancer exhibits significant heterogeneity across multiple scales, as revealed by genomic, histologic, and radiomic analyses [1–3]. Integrating these sources of heterogeneity within a single framework could reveal new insights into the nature and evolution of these tumors [4]. In particular, the fusion of the two most widely available imaging data sources, radiomics and pathomics—radiopathomics—remains mostly underexplored, limiting the potential for precision medicine approaches.

Radiological imaging can capture the macroscopic spatial heterogeneity of tumors, aiding prognosis by extracting radiomic features. Previous kidney cancer radiomics studies have used machine learning and deep learning to analyze

† These authors contributed equally to this work.
⋆ Corresponding author: mc973@cam.ac.uk

hand-crafted or deep features from radiological images for survival prediction. Khodabakhshi et al. [5] combined hand-crafted radiomic and clinical features to predict renal cell carcinoma survival, while Hon et al. [7] used deep neural networks for a similar approach in non-small-cell lung cancer. Recent foundation models in medical imaging show promise in understanding organs and tumors [25, 24], but their applications for prognosis have not yet been extensively studied.

Digital pathology captures tissue architecture and molecular characteristics, essential for cancer diagnosis and prognosis. Deep learning models analyze whole slide images (WSIs) patch by patch to understand histopathological patterns. Tabibu et al. [6] used CNNs to identify tumor regions and extract morphological features for ccRCC survival prediction, while Chen et al. [11] showed that transformer-based pathomic features outperform CNN-based ones in prognosis. Recently, foundation models in pathology have achieved state-of-the-art performance in tasks like subtyping [8], captioning [10], and survival prediction [9], demonstrating their potential for improved cancer analysis.

The complementary strengths of radiological imaging and digital pathology offer promise for personalized cancer treatment and prognosis. Radiopathomics, the integration of radiomics and pathomics, has been explored in breast [12], gastric [13], lung [14], and prostate cancer [15]. Ning et al. [16] combined CNN-extracted radiomic and pathomic features with genomic and clinical data to predict clear cell renal cell carcinoma survival, but their approach was limited by small datasets. The potential of advanced foundation models in radiology and pathology has not yet been studied due to challenges in model design and data heterogeneity.

A major challenge in radiopathomics is large-scale spatial aggregation, where traditional methods like global average pooling or gated attention risk losing sensitivity due to the curse of spatial dimension [16, 9]. To address this, we explore probabilistic integration of renal radiology and pathology using graph neural networks (GNNs). Our contributions are: (1) We introduce a sparsity-informed probabilistic model that statistically reconnect radiomics, pathomics, survival data, and spatial importance, as shown in Fig. 1 (a); (2) To enhance spatial sparsity, we incorporate a Student's t prior and employ graph neural networks to effectively learn spatial importance; and (3) we validate the proposed probabilistic integration method by predicting overall survival in kidney cancer using various radiological and pathological foundation models.

2 Methodology

2.1 Statistical modelling and variational inference

Statistical modeling. Given extracted multi-omics $\mathcal{X} = \{X_r, X_p\}$ and target $y \in \mathbb{R}^k$, where $X_r \in \mathbb{R}^{m_r \times n_r}$ and $X_p \in \mathbb{R}^{m_p \times n_p}$ denotes radiomics and pathomics respectively, we can solve the statistical model in Fig. 1 by maximum a *posteriori* (MAP) estimation:

$$p(s, v|\mathcal{X}, y) \propto p(\mathcal{X}, y|s, v)p(s, v). \tag{1}$$

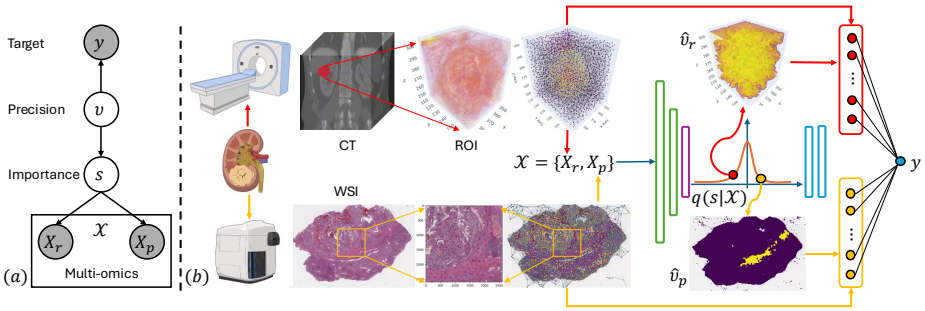


Fig. 1: Probabilistic integration of radiomics and pathomics. (a) Probabilistic graphical model for the integration of radiomics and pathomics. (b) Pipeline of inferring spatial importance and precision, and aggregating spatial features for survival prediction.

Here, s denotes the spatial importance conditioned on its precision v. However, the joint prior distribution, $p(s, v) = p(s|v)p(v)$ is unknown and we need to select proper priors at first. To enforce the sparsity of s, we assign $p(s, v)$ with a Normal-Gamma prior $\mathcal{NG}(\mu_0, \lambda_0, \alpha_0, \beta_0)$, where, $p(s|v) = \mathcal{N}(\mu_0, \text{diag}(\lambda_0 v)^{-1})$, $p(v) = \mathcal{G}(\alpha_0, \beta_0)$, μ_0 denotes the mean and $\lambda_0 v$ is a scaled precision controlling the sparsity. In this case, the marginal distribution $p(s) = \int p(s, v)dv$ is a Student's t distribution, which is known as a long-tailed distribution often used for modeling sparsity.

Variational inference. The MAP problem is still intractable since $p(\mathcal{X}, y|s, v)$ is unknown in practice. To tackle the difficulty, we introduce a variational distribution $q(s, v|\mathcal{X}, y) = q(s|\mathcal{X})q(v|s)$ to approximate the posterior distribution $p(s, v|\mathcal{X}, y)$ by minimizing the Kullback–Leibler (KL) divergence of the former and the latter, which is equivalent to maximize the evidence lower bound,

$$\max \mathbb{E}_{q(s,v|\mathcal{X},y)} \left[\ln p(y|\mathcal{X}, v)p(\mathcal{X}|s)\right] - KL(q(s|\mathcal{X})q(v|s)||p(s|v)p(v)). \quad (2)$$

The second term of Eq. (2) can be further factorized as,

$$\mathbb{E}_{q(v|s)} \left[KL(q(s|\mathcal{X})||p(s|v))\right] + \mathbb{E}_{q(s|\mathcal{X})} \left[KL(q(v|s)||p(v))\right]. \quad (3)$$

For feasibility, we assume $q(s|\mathcal{X}) = \mathcal{N}(\mu_s, \sigma_s^2)$ and $q(v|s) = \mathcal{G}(\alpha_v, \beta_v)$, Then, minimizing Eq. (3) over $q(v|s)$ leads to,

$$\alpha_v = 2\alpha_0 + 1 \text{ and } \beta_v = 2\beta_0 + \lambda_0(\mu_s - \mu_0)^2 + \lambda_0\sigma_s^2. \quad (4)$$

Here, the hyperparameters are set to $\alpha_0 = 2$, $\beta_0 = 1 \times 10^{-2}$, $\mu_0 = 0$, and $\lambda_0 = 1$. After that, $q(s|\mathcal{X})$ is optimized by

$$\min_{q(s|\mathcal{X})} \mathbb{E}_{q(v|s)} \left[KL(q(s|\mathcal{X})||p(s|v))\right] - \mathbb{E}_{q(v|s)} \left[\ln p(y|\mathcal{X}, v)\right] - \mathbb{E}_{q(s|\mathcal{X})} \left[\ln p(\mathcal{X}|s)\right]. \quad (5)$$

Here, the first term drives the sparsity of s, the second term enforces the consistency between predictions and targets, and the last term induces a variational auto-encoder for self-supervised learning.

2.2 Deep inference by graph neural networks

Modelling of spatial correlation. Spatial correlation is crucial for spatial aggregation and regression. To model the spatial correlation of radiological images and whole slide images, we construct radiological and pathological graphs by (1) extracting patch-level feature using foundation models, (2) clustering based on spatial distance and feature distance, and (3) building spatial connection by constructing graphs. The detailed process of constructing graphs is similar to previous works [17].

Pooling module. Multi-omic graphs contain over ten thousand nodes, making direct spatial regression computationally inefficient and prone to the curse of dimensionality. To address this, we use graph neural networks for aggregation. Specifically, we employ graph convolutional layers (GCNconv) to propagate spatial information and apply top-k pooling to select important nodes, where k, the pooling ratio, is set to be 0.2. As shown in Fig. 1, our pooling module consists of two pooling layers (green bars) integrating GCNconv and top-k pooling for efficient node selection.

Multi-omics encoder and decoder. The distribution $q(s|\mathcal{X})$ is estimated by an encoder which maps multi-omics $\mathcal{X} = \{X_r, X_p\}$ to sparse importance score $s \in \mathbb{R}^{k^2(m_r+m_p)}$, where k denotes the pool ratio of the pooling module. To mitigate the different feature dimension between radiomics and pathomics, two encoders, i.e., $E_{\theta_r^e}$ and $E_{\theta_p^e}$, are built to project them into the same dimensional features, namely, $s^T = [s_r^T, s_p^T]$, where $s_r = E_{\theta_r^e}(X_r)$ and $s_p = E_{\theta_p^e}(X_p)$. Thus, $q(s|\mathcal{X})$ is parameterized by $q_{\theta_E}(s|\mathcal{X})$, where $\theta_E = \{\theta_r^e, \theta_p^e\}$. Similarly, the distribution $p(\mathcal{X}|s)$ is estimated by two decoders which map the sparse importance score s to the multi-omics $\mathcal{X} = \{X_r, X_p\}$. Symmetrically, two decoder branches (blue bars), i.e., $D_{\theta_r^d}$ and $D_{\theta_p^d}$, are built to recover radiomics and pathomics, namely, $X_r = D_{\theta_r^d}(s_r) + N_r$ and $X_p = D_{\theta_p^d}(s_p) + N_p$, where N_r and N_p denote Gaussian noise. Hence, $p(\mathcal{X}|s)$ is parameterized as,

$$p_{\theta_D}(\mathcal{X}|s) = \mathcal{N}(X_r|D_{\theta_r^d}(s_r), I_r/\tau_r)\mathcal{N}(X_p|D_{\theta_p^d}(s_p), I_p/\tau_p), \tag{6}$$

where, $\theta_D = \{\theta_r^d, \theta_p^d\}$, τ_r and τ_p are hyperparameters set to 1×10^{-3}.

Sparsity-Informed Spatial Aggregation (SPARRA). The distribution $p(y|\mathcal{X}, \upsilon)$ induces a task-specific predictor consists of spatial regression and multivariate regression. Specifically, the *spatial regression* aims to integrate spatial feature vectors using the precision $\upsilon \in \mathbb{R}^{k^2(m_r+m_p)}$, which is achieved by $\mathbf{x} = [\hat{\upsilon}_r^T f_{pool}(X_r), \hat{\upsilon}_p^T f_{pool}(X_p)]$, where $\hat{\upsilon}_\cdot$ denotes a normalized inverse precision, namely $\hat{\upsilon}_\cdot = \upsilon_\cdot^{-1}/\|\upsilon_\cdot^{-1}\|_1$, and $f_{pool}(X_\cdot) \in \mathbb{R}^{k^2(m_r+m_p)\times n_h}$ denotes the output of the pooling module with the hidden dimension as n_h. Note that while SPARRA is built on multi-omics data, it remains effective for a single modality by limiting multi-omics to either radiomics or pathomics. The *multivariate regression* aims to predict targets from the aggregated feature vector $\mathbf{x}$, which allows us to predict y by a linear layer $f_{\theta_{mlp}}(\mathbf{x})$, then $p(y|\mathcal{X}, \upsilon)$ is expressed as,

$$p_{\theta_y}(y|\mathcal{X}, \upsilon) = \frac{1}{Z}\exp\{-D(y, f_{\theta_{mlp}}(\mathbf{x}))\}, \tag{7}$$

where, Z is a normalizing constant, $\theta_y = \{\theta_E, \theta_{mlp}\}$ denotes the set of all parameters related to the prediction, and $D(\cdot, \cdot)$ denotes the negative partial log likelihood for survival analysis.

Training loss function. Based on the parameterized $q_{\theta_E}(s|\mathcal{X})$, $p_{\theta_D}(\mathcal{X}|s)$, and $p_{\theta_y}(y|\mathcal{X}, \upsilon)$ by neural networks, we obtain a loss of training neural networks from Eq. (5) as follows,

$$\min_{\theta_D, \theta_y} \mathbb{E}_{q(\upsilon|s)} \left[KL(q_{\theta_E}(s|\mathcal{X})||p(s|\upsilon)) - \ln p_{\theta_y}(y|\mathcal{X}, \upsilon) \right] - \mathbb{E}_{q_{\theta_E}(s|\mathcal{X})} \left[\ln p_{\theta_D}(\mathcal{X}|s) \right] \quad (8)$$

3 Experiments

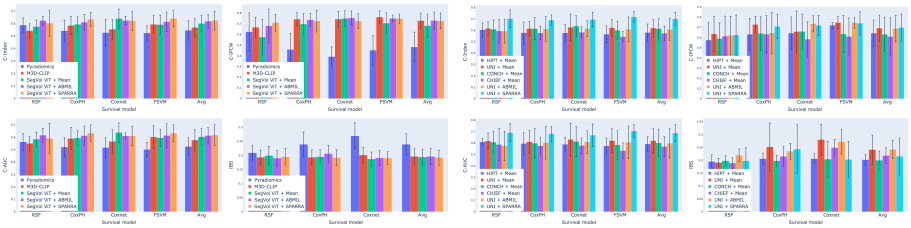
3.1 Datasets and implementation details

Datasets. The TCGA-RCC datasets (TCGA-KIRC [18], TCGA-KIRP [19], TCGA-KICH [20]) contain 947 subjects with digital slides, while TCIA includes 272 matched subjects with CT scans. For radiomics extraction, we trained a SegResNet [22] using KiTS23 [21] for kidney tumor segmentation and tested it on the 272 CT scans. After manual validation and clinical data matching, 205 subjects with WSIs, CT scans, and clinical records were selected, comprising 180 from TCGA-KIRC, 15 from TCGA-KIRP, and 10 from TCGA-KICH. We then performed 5-fold cross-validation, splitting the 205 samples into an 8:2 ratio, with 164 for training and 41 for testing in each fold.

Feature extraction and graph construction. For radiomics, we used three feature extractors: Pyradiomics [23] for hand-crafted features, M3D-CLIP [24] for projecting volumetric CT scans into feature vectors, and SegVol ViT [25], which maps 3D CT scans to 4D feature tensors while preserving spatial information. For pathomics, we included HIPT [11], a transformer trained on large-scale WSIs, along with UNI [8], CONCH [10], and CHIEF [9], foundation models achieving state-of-the-art performance in computational pathology. Using the extracted radiomic and pathomic features, we constructed radiological and pathological graphs for CT scans and WSIs, respectively.

Univariate feature selection. Given the high dimensionality of extracted features, we performed univariate Cox regression to select relevant features, retaining those with a p-value below 0.2 to avoid excluding rare but important ones. This loose threshold ensures key features are preserved for regression. However, our SPARRA model does not require feature selection, as it automatically identifies important features.

Multivariate models and metrics. To evaluate the integration of radiomics and pathomics, we used four survival models: random survival forest (RSF), Cox proportional hazards model with ridge regression (CoxPH), Cox proportional hazards model with elastic net (Coxnet), and fast survival support vector machine (FSVM). For RSF, we used 5-fold cross-validation on training samples to select the optimal max depth, while for CoxPH, Coxnet, and FSVM, we selected regularization parameters through 5-fold cross-validation. Survival



(a) Radiomics (b) Pathomics

Fig. 2: Kidney cancer survival analysis based on (a) radiomics and (b) pathomics.

prediction performance was assessed using four metrics: Concordance index (C-Index), Concordance index with inverse probability of censoring weights (C-IPCW), cumulative/dynamic AUC (C-AUC), and Integrated Brier Score (IBS).

Implementation. We trained the models using the ADAM optimizer with an initial learning rate of 3×10^{-4}, decayed following a cosine annealing schedule. The total number of training steps, the rate of weight decay, and the probability of dropout, and the batch size were set to 20, 1×10^{-5}, 0.5, and 32, respectively. All models were trained and tested on an NVIDIA L40 GPU with 48GB of memory.

3.2 Survival analysis by radiomics

In this section, we evaluated kidney cancer survival prediction using radiomics by comparing feature extractors (Pyradiomics, M3D-CLIP, and SegVol ViT) and aggregation methods (mean pooling, attention-based multiple instance learning (ABMIL), and the proposed SPARRA). During training, ABMIL and SPARRA were optimized by minimizing the negative partial log-likelihood [9]. For testing, we aggregated features using Mean, ABMIL, and SPARRA, then performed survival analysis and reported results from 19 studies, including C-Index (↑), C-IPCW (↑), and C-AUC (↑) for four survival models, IBS (↓) for three models, and average scores across models.

Fig. 2a compares survival prediction performance using four survival models and metrics. Among Pyradiomics, M3D-CLIP, and SegVol ViT + Mean, SegVol ViT outperforms the others in 11 out of 19 studies, highlighting its ability to better represent tumor texture. For SegVol ViT, SPARRA outperforms Mean and ABMIL in 8 studies, while ABMIL leads in 7 studies, indicating that SPARRA and ABMIL deliver comparable performance. This may be due to radiomic features extracted from the different tumor subregions all containing information relevant for survival prediction.

3.3 Survival analysis by pathomics

In this section, we evaluated kidney cancer survival prediction using pathomics by comparing feature extractors (HIPT, UNI, CONCH, and CHIEF) and ag-

gregation methods (Mean, ABMIL, and SPARRA). The training and testing settings were consistent with those in Section 3.2.

Fig. 2b compares survival prediction performance using different pathomics and survival models. For Mean aggregation, UNI outperforms HIPT, CONCH, and CHIEF in 13 out of 19 studies, indicating that UNI better captures tumor microenvironments. When comparing aggregation methods based on UNI, SPARRA outperforms Mean and ABMIL in 12 out of 19 studies, showing significant improvements in C-Index and C-AUC. This highlights SPARRA's ability to identify sparse biomarkers crucial for survival prediction. Notably, ABMIL performs worse than Mean across the four metrics due to overfitting on pathomics.

3.4 Survival analysis by radiopathomics

In this section, we evaluated the integration of radiomics and pathomics for kidney cancer survival prediction. Using SegVol ViT for radiomics and UNI for pathomics, we compared three aggregation methods—Mean, ABMIL, and SPARRA—following the same settings as in Sections 3.2 and 3.3.

Table 1 presents the results for kidney cancer survival analysis using radiopathomics. SPARRA outperforms the other aggregation methods in 14 out of 19 studies, showing significant improvements in C-Index, C-AUC, and IBS. In contrast, ABMIL performs worse than mean aggregation due to overfitting on pathomics, as shown in Section 3.3. SPARRA, by incorporating a sparse prior, effectively regularizes feature aggregation and prevents overfitting.

Fig. 3 shows the mean scores and standard deviation bands from 5-fold cross-validation. Radiomics-based models exhibit narrower bands, indicating greater stability in survival prediction. Compared to Mean and ABMIL, SPARRA-based models reduce standard deviations in pathomics. Particularly, SPARRA can significantly improve the C-Index and C-AUC of CoxPH by integrating radiomics with pathomics. This demonstrates that SPARRA identifies more robust fea-

Radiopathomics		Metric	RSF	CoxPH	Coxnet	FSVM	Avg.
Extractor	Aggr.						
SegVol ViT UNI	Mean	C-Index	$0.614_{0.085}$	$0.635_{0.077}$	$0.614_{0.122}$	$0.630_{0.086}$	$0.623_{0.093}$
		C-IPCW	$\mathbf{0.640_{0.178}}$	$0.746_{0.065}$	$\mathbf{0.759_{0.037}}$	$0.746_{0.067}$	$\mathbf{0.723_{0.087}}$
		C-AUC	$0.619_{0.104}$	$0.628_{0.113}$	$0.621_{0.157}$	$0.627_{0.122}$	$0.624_{0.124}$
		IBS	$\mathbf{0.191_{0.019}}$	$0.293_{0.083}$	$0.283_{0.061}$	N/A	$0.256_{0.054}$
SegVol ViT UNI	ABMIL	C-Index	$0.597_{0.091}$	$0.624_{0.093}$	$0.614_{0.083}$	$0.623_{0.091}$	$0.615_{0.090}$
		C-IPCW	$0.612_{0.188}$	$\mathbf{0.764_{0.052}}$	$0.640_{0.201}$	$\mathbf{0.750_{0.055}}$	$0.692_{0.124}$
		C-AUC	$0.567_{0.132}$	$0.620_{0.130}$	$0.604_{0.106}$	$0.621_{0.136}$	$0.603_{0.126}$
		IBS	$0.225_{0.028}$	$0.247_{\mathbf{0.032}}$	$0.271_{\mathbf{0.043}}$	N/A	$0.248_{\mathbf{0.034}}$
SegVol ViT UNI	SPARRA	C-Index	$\mathbf{0.684_{0.063}}$	$\mathbf{0.716_{0.064}}$	$\mathbf{0.706_{0.071}}$	$\mathbf{0.721_{0.063}}$	$\mathbf{0.707_{0.065}}$
		C-IPCW	$0.612_{0.179}$	$0.742_{0.094}$	$0.735_{0.083}$	$0.747_{0.103}$	$0.709_{0.115}$
		C-AUC	$\mathbf{0.674_{0.054}}$	$\mathbf{0.711_{0.077}}$	$\mathbf{0.681_{0.105}}$	$\mathbf{0.714_{0.064}}$	$\mathbf{0.695_{0.075}}$
		IBS	$0.196_{0.037}$	$\mathbf{0.219}_{0.071}$	$\mathbf{0.198}_{0.067}$	N/A	$\mathbf{0.204}_{0.058}$

Table 1: Kidney cancer survival analysis based on radiopathomics. Bold denotes the best performance.

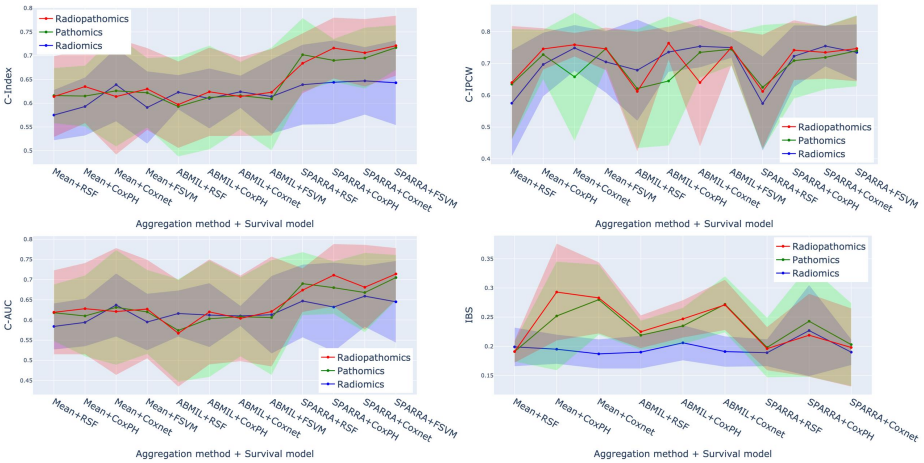


Fig. 3: Kidney cancer survival analysis based on radiopathomics.

tures for survival analysis. None of the radiopathomic models significantly outperforms the corresponding pathomics model. This is due to the comparably lower predictive power of radiomics features in this dataset.

Fig. 4 shows survival functions for low- and high-risk groups identified by CoxPH using different radiopathomic aggregation methods. Risk groups were defined by comparing individual risk scores to the mean. All log-rank test p-values are below 0.005, indicating significant differences between the groups. However, Mean and ABMIL-based models fail to distinguish between low- and high-risk groups early on, while the SPARRA-based CoxPH performs better, highlighting the prognostic potential of SPARRA.

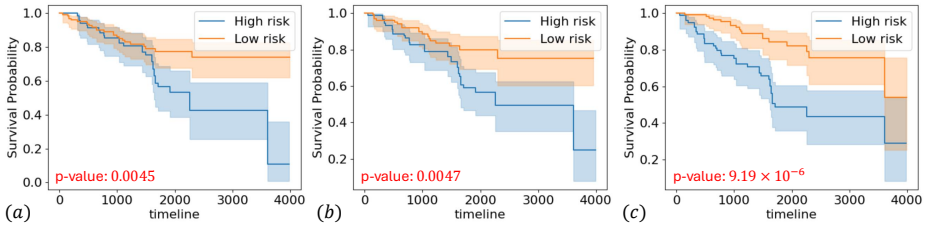


Fig. 4: Survival functions of low- and high-risk groups based on radiopathomics. The groups were identified by (a) Mean + CoxPH, (b) ABMIL + CoxPH, and (c) SPARRA + CoxPH.

4 Conclusion

In this work, we proposed a sparsity-informed probabilistic integration of radiomics and pathomics. We constructed a probabilistic graphical model to jointly model radiomics, pathomics, spatial importance, and survival data, incorporated a sparsity prior to enforce feature sparsity and used variational inference for model estimation. Our validation on kidney cancer survival analysis demonstrated the superior performance of this probabilistic integration in radiopathomics, enabling better prognostic assessments for kidney cancer patients.

Acknowledgments. This work was supported by the Cancer Research UK Cambridge Centre [CTRQQR-2021\100012; and C9685/A25117], The Mark Foundation for Cancer Research [RG95043], NIHR Cambridge Biomedical Research Centre (NIHR203312), and the EPSRC Tier-2 capital grant [EP/P020259/1]. M.C.O. was supported by the Joseph Mitchell Cancer Research Fund, the Academy of Medical Sciences [G117526] and NIHR [NIHR206092]. The results shown here are in whole or part based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Turajlic, S., Xu, H., Litchfield, K., et al.: Deterministic evolutionary trajectories influence primary tumor growth: TRACERx renal. Cell **173**(3), 595–610 (2018).
2. Wang, W., Cao, K., Jin, S., Zhu, X., et al.: Differentiation of renal cell carcinoma subtypes through MRI-based radiomics analysis. European Radiology **30**(10), 5738–5747 (2020).
3. Li, Y., Lih, T. S. M., Dhanasekaran, S. M., et al.: Histopathologic and proteogenomic heterogeneity reveals features of clear cell renal cell carcinoma aggressiveness. Cancer cell **41**(1), 139–163 (2023).
4. Hu, J., Wang, S. G., Hou, Y., et al.: Multi-omic profiling of clear cell renal cell carcinoma identifies metabolic reprogramming associated with disease progression. Nature Genetics **56**(3), 442–457 (2024).
5. Khodabakhshi, Z., Amini, M., Mostafaei, S., et al.: Overall Survival Prediction in Renal Cell Carcinoma Patients Using Computed Tomography Radiomic and Clinical Information. J Digit Imaging **34**(5), 1086–1098 (2012).
6. Tabibu, S., Vinod, P.K., Jawahar, C.V.: Pan-Renal Cell Carcinoma classification and survival prediction from histopathology images using deep learning. Sci. Rep. **9**, 10509 (2019).
7. Hou, K. Y., Chen, J. R., Wang, Y. C., et al.: Radiomics-Based Deep Learning Prediction of Overall Survival in Non-Small-Cell Lung Cancer Using Contrast-Enhanced Computed Tomography. Cancers (Basel) **14**(15), 3798 (2022).
8. Chen, R. J, Ding, T., Lu, M. Y., et al.: Towards a general-purpose foundation model for computational pathology. Nature Medicine **30**(3), 850–862 (2024).
9. Wang, X., Zhao, J., Marostica, E., et al.: A pathology foundation model for cancer diagnosis and prognosis prediction. Nature **634**(8035), 970–978 (2024).

10. Lu, M. Y., Chen, B., Williamson, D. F. K., et al. A visual-language foundation model for computational pathology. Nature Medicine **30**(3), 863–874 (2024).
11. Chen, R. J., Chen, C., Li, Y., et al. Scaling vision transformers to gigapixel images via hierarchical self-supervised learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 16144–16155 (2022).
12. Xu, N., Guo, X., Ouyang, Z., et al.: Multiparametric MRI-based radiomics combined with pathomics features for prediction of the efficacy of neoadjuvant chemotherapy in breast cancer. Heliyon **10**(2), e24371 (2024).
13. Tan, Y., Feng, L. J., Huang, Y. H., et al.: A comprehensive radiopathological nomogram for the prediction of pathological staging in gastric cancer using CT-derived and WSI-based features. Transl. Oncol. **40**, 101864 (2024).
14. Tortora, M., Cordelli, E., Sicilia, R., et al.: RadioPathomics: Multimodal Learning in Non-Small Cell Lung Cancer for Adaptive Radiotherapy. IEEE Access 11, 47563–47578 (2023).
15. Zhang, Y. F., Zhou, C., Guo, S., et al.: Deep learning algorithm-based multimodal MRI radiomics and pathomics data improve prediction of bone metastases in primary prostate cancer. J Cancer Res Clin Oncol **150**(2), 78 (2024).
16. Ning, Z., Pan, W., Chen, Y., et al.: Integrative analysis of cross-modal features for the prognosis prediction of clear cell renal cell carcinoma. Bioinformatics **36**(9), 2888–2895 (2020).
17. Gao, S., Browning, L., Alham, N. K., et al. Characterising borderline areas in bladder tumour grading with Bayesian graph neural networks. In: 2024 IEEE International Symposium on Biomedical Imaging (ISBI), Athens, Greece, pp. 1–5 (2024).
18. Akin, O., Elnajjar, P., Heller, M. et al.: The Cancer Genome Atlas Kidney Renal Clear Cell Carcinoma Collection (TCGA-KIRC) (Version 3) [Data set] (2016). The Cancer Imaging Archive.
19. Linehan, M., Gautam, R., Kirk, S. et al.: The Cancer Genome Atlas Cervical Kidney Renal Papillary Cell Carcinoma Collection (TCGA-KIRP) (Version 4) [Data set] (2016). The Cancer Imaging Archive.
20. Linehan, M. W., Gautam, R., Sadow, C. A., Levine, S.: The Cancer Genome Atlas Kidney Chromophobe Collection (TCGA-KICH) (Version 3) [Data set] (2016). The Cancer Imaging Archive.
21. Heller, N., Isensee, F., Maier-Hein, K. H., et al.: The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge. Medical Image Analysis **67**, 101821 (2021).
22. Myronenko, A: 3D MRI Brain Tumor Segmentation Using Autoencoder Regularization. In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2018.
23. Griethuysen van, J. J. M., Fedorov, A., Parmar, C. et al.: Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research **77**(21), e104–e107 (2017).
24. Bai, F., Du, Y., Huang, T., et al.: M3D: Advancing 3D Medical Image Analysis with Multi-Modal Large Language Models. arXiv eprint: 2404.00578 (2024).
25. Du, Y., Bai, F., Huang, T., Zhao, B.: SegVol: Universal and Interactive Volumetric Medical Image Segmentation. In: The Thirty-eighth Annual Conference on Neural Information Processing Systems (2024).