

Supplementary Material

When RNA is Missing: A Multimodal Fusion Benchmark for BCG Progression Prediction in HR-NMIBC

Saeedeh Hashemi, Tahlita Zuiverloon, Nadieh Khalili

S1. Fold-Level Results

Table S1. Fold-level C-index values for all evaluated model configurations. Fold-level C-index values for all evaluated model configurations. Uni = unimodal, Bi = bimodal, Tri = trimodal; Raw tab = raw tabular clinical encoding; BERT = BioClinical ModernBERT; Concat = concatenation fusion; P = pathology, R = RNA, C = clinical.

Con-fig	Mo-dality	Clinical encoder	Fusion	Mean $\pm$ std	Fold 0	Fold 1	Fold 2	Fold 3	Fold 4
Uni	P	-	-	0.585 $\pm$ 0.046	0.582	0.568	0.666	0.561	0.550
Uni	R	-	-	0.610 $\pm$ 0.086	0.586	0.518	0.567	0.744	0.633
Uni	C	Raw tab	-	0.648 $\pm$ 0.059	0.557	0.656	0.719	0.636	0.669
Uni	C	CONCH	-	0.662 $\pm$ 0.072	0.655	0.670	0.734	0.545	0.707
Uni	C	Modern-BERT	-	0.705 $\pm$ 0.058	0.611	0.723	0.741	0.690	0.758
Bi	R+P	-	Concat	0.631 $\pm$ 0.092	0.617	0.527	0.616	0.781	0.615
Bi	R+P	-	Gated	0.628 $\pm$ 0.095	0.607	0.528	0.609	0.787	0.610
Bi	R+P	-	Low-rank	0.580 $\pm$ 0.087	0.527	0.485	0.586	0.715	0.589
Bi	C+R	Raw tab	Concat	0.673 $\pm$ 0.083	0.628	0.575	0.768	0.645	0.749
Bi	C+R	Raw tab	Gated	0.676 $\pm$ 0.080	0.639	0.577	0.764	0.649	0.752
Bi	C+R	Raw tab	Low-rank	0.564 $\pm$ 0.047	0.508	0.628	0.585	0.529	0.573
Bi	C+P	Raw tab	Concat	0.655 $\pm$ 0.049	0.622	0.693	0.719	0.604	0.637
Bi	C+P	Raw tab	Gated	0.718 $\pm$ 0.032	0.763	0.689	0.726	0.674	0.740
Bi	C+P	Raw tab	Low-rank	0.698 $\pm$ 0.034	0.713	0.674	0.750	0.665	0.689
Bi	C+P	CONCH	Concat	0.625 $\pm$ 0.024	0.606	0.615	0.661	0.636	0.606
Bi	C+P	CONCH	Low-rank	0.593 $\pm$ 0.046	0.614	0.564	0.660	0.585	0.541
Bi	C+R	CONCH	Concat	0.648 $\pm$ 0.100	0.679	0.485	0.676	0.757	0.644
Bi	C+R	CONCH	Gated	0.623 $\pm$ 0.076	0.640	0.494	0.693	0.635	0.654
Bi	C+R	CONCH	Low-rank	0.581 $\pm$ 0.090	0.539	0.504	0.609	0.727	0.525
Bi	C+P	BERT	Concat	0.608 $\pm$ 0.064	0.690	0.582	0.658	0.577	0.534
Bi	C+P	BERT	Gated	0.599 $\pm$ 0.053	0.643	0.588	0.660	0.572	0.532
Bi	C+P	BERT	Low-rank	0.550 $\pm$ 0.045	0.528	0.524	0.630	0.545	0.525
Bi	C+R	BERT	Concat	0.638 $\pm$ 0.084	0.680	0.491	0.647	0.693	0.680
Bi	C+R	BERT	Gated	0.640 $\pm$ 0.092	0.626	0.491	0.719	0.711	0.653
Bi	C+R	BERT	Low-rank	0.639 $\pm$ 0.095	0.585	0.562	0.566	0.765	0.719
Tri	C+R+P	Raw tab	Concat	0.710 $\pm$ 0.078	0.712	0.575	0.758	0.753	0.752
Tri	C+R+P	Raw tab	Gated	0.726 $\pm$ 0.094	0.738	0.577	0.834	0.760	0.721

Tri	C+R+P	Raw tab	Low-rank	0.686±0.093	0.670	0.535	0.777	0.738	0.709
Tri	C+R+P	CONCH	Concat	0.642±0.047	0.654	0.571	0.638	0.703	0.642
Tri	C+R+P	CONCH	Gated	0.672±0.097	0.659	0.557	0.706	0.816	0.620
Tri	C+R+P	CONCH	Low-rank	0.594±0.077	0.629	0.459	0.609	0.645	0.629
Tri	C+R+P	BERT	Concat	0.633±0.089	0.657	0.496	0.693	0.718	0.600
Tri	C+R+P	BERT	Gated	0.625±0.069	0.670	0.504	0.664	0.643	0.642
Tri	C+R+P	BERT	Low-rank	0.597±0.098	0.611	0.438	0.706	0.617	0.613

S2. Missing-RNA Robustness

Table S2. Fold-level C-index under complete and missing RNA at inference (Complete/Missing).

Config	Drop-out	Fold 0	Fold 1	Fold 2	Fold 3	Fold 4
Concat	0.30	0.654/0.710	0.618/0.661	0.757/0.747	0.766/0.732	0.746/0.689
Concat	0.50	0.739/0.730	0.697/0.700	0.780/0.780	0.721/0.621	0.705/0.696
Gated	0.30	0.697/0.720	0.658/0.678	0.839/0.787	0.757/0.681	0.596/0.609
Gated	0.50	0.724/0.724	0.676/0.673	0.750/0.750	0.652/0.651	0.560/0.600
Low-rank bilinear	0.30	0.726/0.763	0.522/0.700	0.812/0.755	0.741/0.665	0.707/0.714
Low-rank bilinear	0.50	0.807/0.761	0.656/0.683	0.757/0.771	0.659/0.618	0.689/0.689
Surv-pathway	0.30	0.625/0.658	0.630/0.614	0.685/0.648	0.721/0.713	0.673/0.660
Surv-pathway	0.50	0.657/0.637	0.603/0.697	0.661/0.599	0.759/0.687	0.660/0.641
Surv-biocategory	0.30	0.694/0.716	0.582/0.592	0.666/0.664	0.791/0.750	0.677/0.634
Surv-biocategory	0.50	0.692/0.657	0.626/0.687	0.693/0.696	0.776/0.675	0.669/0.600
Surv-scF	0.30	0.706/0.721	0.667/0.672	0.644/0.666	0.683/0.640	0.701/0.771
Surv-scF	0.50	0.706/0.676	0.591/0.677	0.651/0.676	0.760/0.741	0.650/0.753

S3. Number at Risk

Table S3 reports the numbers at risk at selected follow-up times for the Kaplan-Meier analyses presented in Figure 2.

Table S3. Numbers at risk corresponding to the Kaplan-Meier curves shown in Figure 2.

Model	Risk Group	0	50	100	150	200
Trimodal gated fusion	Low risk	136	91	31	7	0
	High risk	136	62	18	1	1
Clinical unimodal (BERT5)	Low risk	136	88	31	5	1
	High risk	136	65	18	3	0
Pathology unimodal	Low risk	136	86	28	5	1
	High risk	136	67	21	3	0
RNA unimodal	Low risk	136	80	23	2	0
	High risk	136	73	26	6	1