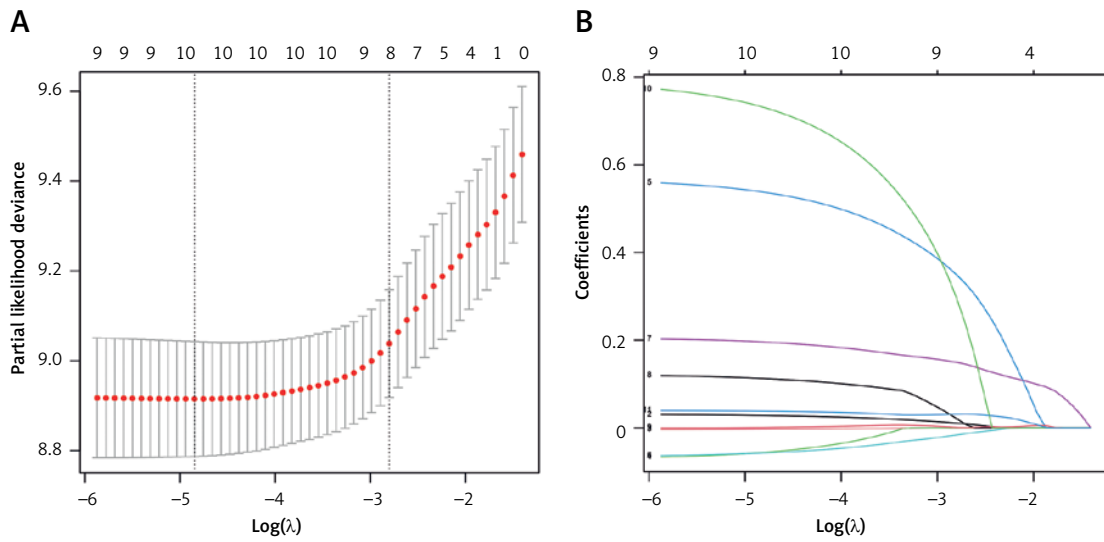
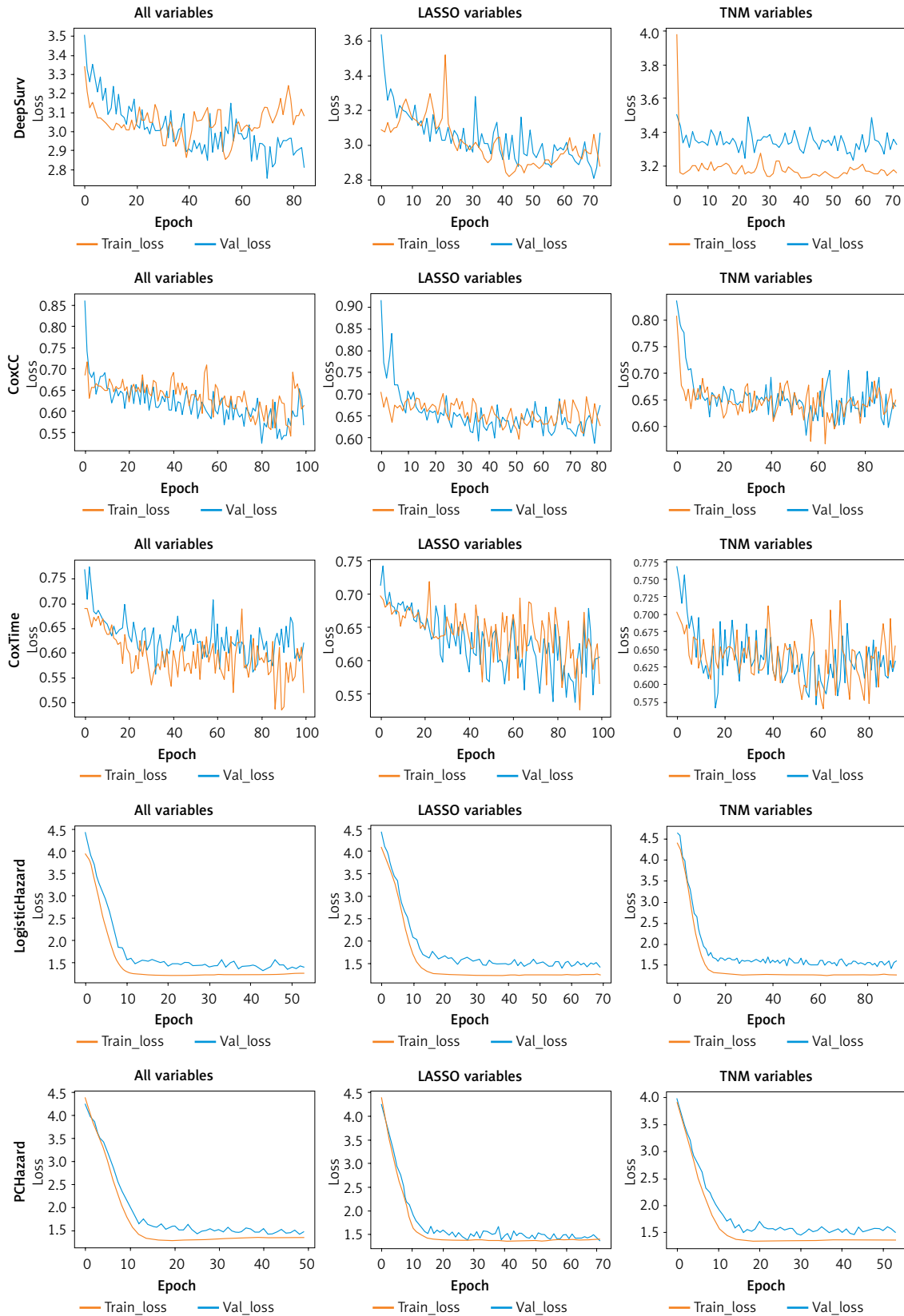


Supplementary Figure S1. Distribution of missing values of all patients
LNE – number of lymph nodes examined, LNP – number of lymph nodes positive.

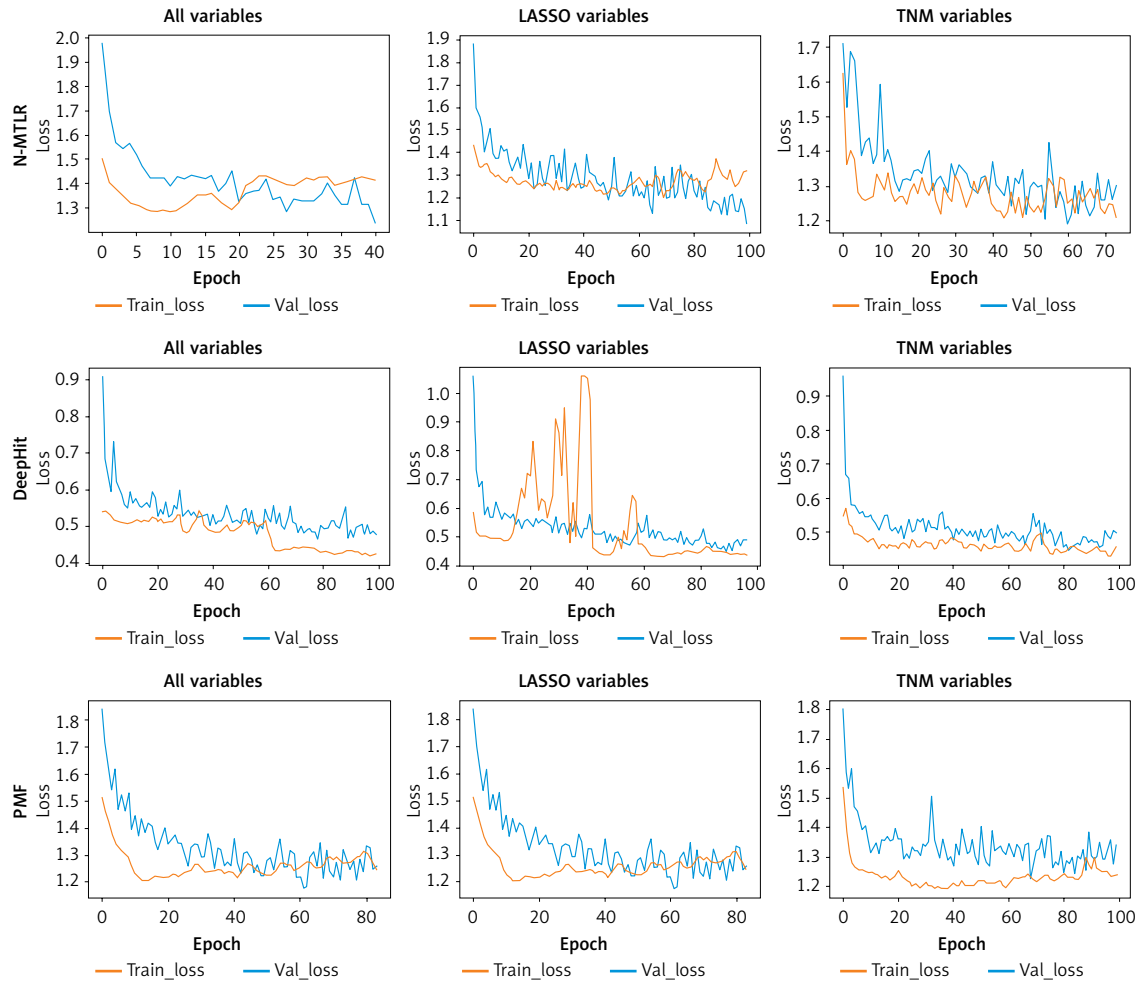


Supplementary Figure S2. Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression was used to filter clinical features. **A** – Search for the right number of predictor variables and λ value. **B** – LASSO coefficient profiles of clinical features



Supplementary Figure S3. Training curves of 8 neural network survival models with callback function to enable stopping trains automatically

CoxCC – Cox case-control corresponding methods. PCHazard – Piecewise Constant Hazard. N-MTLR – Neural Multi-Task Logistic Regression. PMF – Probability Mass Function. LASSO – Least Absolute Shrinkage and Selection Operator.



Supplementary Figure S3. Cont.

CoxCC – Cox case-control corresponding methods. PCHazard – Piecewise Constant Hazard. N-MTLR – Neural Multi-Task Logistic Regression. PMF – Probability Mass Function. LASSO – Least Absolute Shrinkage and Selection Operator.

Supplementary Table SI. Coefficients of variables in LASSO Cox regression

Variable	Coefficient
Sex	0.000
Age	0.028
Size	-0.003
Site	-0.057
Grade	0.539
Lymph nodes examined	-0.058
Lymph nodes positive	0.196
T	0.113
N	0.001
M	0.734
Stage	0.038

LASSO – Least Absolute Shrinkage and Selection Operator.

Supplementary Table SII. C-index 95% confidence interval from 1000 times bootstrap of different deep learning survival models

Models	TNM variables						LASSO variables						All variables					
	Internal validation		External validation		Internal validation		External validation		Internal validation		External validation		Internal validation		External validation			
	Mean	High	Low	High	Mean	High	Low	High	Mean	High	Low	High	Mean	High	Low	High		
DeepSurv	0.6950	0.6925	0.6975	0.6660	0.6630	0.6691	0.8256	0.8240	0.8272	0.7543	0.7517	0.7568	0.8315	0.8297	0.8332	0.7719	0.7693	0.7745
CoxCC	0.6677	0.6652	0.6702	0.6785	0.6758	0.6813	0.7090	0.7067	0.7114	0.6847	0.6818	0.6876	0.7554	0.7533	0.7575	0.6985	0.6957	0.7012
CoxTime	0.6716	0.6692	0.6743	0.6720	0.6692	0.6748	0.7333	0.7311	0.7355	0.6995	0.6968	0.7022	0.7430	0.7409	0.7451	0.7193	0.7165	0.7220
Logistic-Hazard	0.6766	0.6743	0.6789	0.6665	0.6639	0.6691	0.7217	0.7196	0.7237	0.6866	0.6840	0.6892	0.7320	0.7301	0.7340	0.7055	0.7030	0.7080
PCHazard	0.5494	0.5489	0.5500	0.5358	0.5354	0.5363	0.5495	0.5489	0.5501	0.5353	0.5349	0.5358	0.5499	0.5493	0.5504	0.5357	0.5353	0.5362
N-MTLR	0.5430	0.5403	0.5458	0.5551	0.5521	0.5580	0.5230	0.5204	0.5254	0.6070	0.6041	0.6098	0.5271	0.5243	0.5298	0.6430	0.6400	0.6459
DeepHit	0.4783	0.4756	0.4811	0.4133	0.4105	0.4162	0.6306	0.6282	0.6330	0.6788	0.6759	0.6818	0.6890	0.6868	0.6913	0.6493	0.6466	0.6521
PMF	0.5181	0.5154	0.5208	0.5233	0.5198	0.5269	0.5451	0.5424	0.5477	0.5832	0.5804	0.5860	0.6251	0.6227	0.6275	0.5857	0.5828	0.5885
Cox	0.6717	0.6692	0.6742	0.6710	0.6681	0.6738	0.7343	0.7321	0.7365	0.6632	0.6602	0.6662	0.7354	0.7332	0.7376	0.6638	0.6608	0.6669

C-index, concordance index. Mean – mean C-index. Low – low C-index 95% confidence. High – high C-index 95% confidence. TNM variables: T + N + M + Stage. LASSO variables: Age + Size + Site + Grade + Lymph nodes examined + Lymph nodes positive + T + N + M + Stage. All variables: Sex + Age + Size + Site + Grade + Lymph nodes examined + Lymph nodes positive + T + N + M + Stage. LASSO – Least Absolute Shrinkage and Selection Operator. CoxCC – Cox Case-Control Corresponding Methods. PCHazard – Piecewise Constant Hazard. N-MTLR, Neural Multi-Task Logistic Regression. PMF – Probability Mass Function.

Additional Table S1. Clinical data of patients in two cohorts

Sex	Age	Size	Site	Grade	Lymph nodes examined	Lymph nodes positive	T	N	M	Stage	Duration	Event
Male	54	50	Ascending colon	II	11	0	T3	N0	M0	II	51	0
Female	74	30	Rectum	II	7	0	T2	N0	M0	I	60	1
Male	63	70	Rectum	III	6	0	T3	N0	M0	IIA	90	0
Female	68	50	Rectum	II	4	0	T2	N0	M0	I	90	0
Male	57	20	Rectum	II	2	0	T1	N0	M0	I	98	0
Female	55	70	Ascending colon	III	25	0	T4a	N0	M0	II	49	0
Female	79	55	Rectum	III	7	1	T3	N1a	M1a	IVA	31	1
Female	67	50	Rectum	III	7	1	T4a	N1a	M0	IIIB	89	0
Female	72	40	Rectum	II	4	0	T3	N0	M0	IIA	68	1
Female	78	12	Rectum	III	5	2	T3	N1b	M0	IIIB	36	1
Female	66	50	Rectum	II	1	0	T3	N0	M0	IIA	85	0
Male	80	95	Rectum	III	5	0	T3	N0	M0	IIA	27	1
Female	60	60	Rectum	II	2	2	T3	N1b	M0	IIIB	7	1
Female	51	60	Rectosigmoid junction	II	6	3	T4a	N1b	M0	IIIB	92	0
Male	78	35	Sigmoid colon	II	15	3	T4a	N1	M0	III	60	0
Female	78	80	Rectum	II	2	0	T3	N0	M0	IIA	48	1
Female	56	65	Rectum	II	4	1	T3	N1a	M0	IIIB	93	0
Female	73	70	Ascending colon	III	23	2	T4a	N1	M0	III	18	1
Male	79	70	Hepatic flexure	II	30	2	T3	N1	M0	III	77	0
Female	59	50	Rectum	III	2	0	T3	N0	M0	IIA	99	1
Male	74	80	Rectum	II	13	0	T3	N0	M0	IIA	68	1
Female	75	40	Rectum	III	6	1	T3	N1a	M0	IIIB	85	0
Male	67	80	Rectum	II	15	2	T3	N1b	M0	IIIB	92	0
Male	36	4	Rectum	II	1	0	T2	N0	M0	I	95	0
Male	58	35	Ascending colon	II	11	0	T2	N0	M0	I	55	0
Male	46	45	Rectum	II	2	0	T2	N0	M0	I	86	0
Male	46	70	Descending colon	II	75	1	T4b	N1	M0	III	22	1
Female	78	50	Large intestine, NOS	II	12	1	T4a	N1	M0	III	54	0
Female	67	40	Rectum	II	5	0	T2	N0	M0	I	89	0
Female	76	40	Large intestine, NOS	III	24	24	T3	N2	M0	III	50	0
Male	48	30	Rectum	II	6	0	T3	N0	M0	IIA	104	0
Female	59	50	Ileocecal junction	II	19	0	T4b	N0	M0	IIC	12	1
Male	73	80	Rectum	III	11	0	T3	N1c	M0	IIIB	14	1
Male	76	50	Sigmoid colon	III	18	17	T4b	N2b	M0	IIIC	9	1

Additional Table SI. Cont.

Sex	Age	Size	Site	Grade	Lymph nodes examined	Lymph nodes positive	T	N	M	Stage	Duration	Event
Male	72	53	Large intestine, NOS	II	21	0	T3	N0	M0	II	28	1
Male	61	55	Sigmoid colon	II	28	0	T4a	N0	M0	II	52	0
Male	72	55	Rectum	II	4	0	T3	N0	M0	IIA	1	1
Male	66	50	Rectum	II	13	3	T3	N1b	M0	IIIB	48	1
Female	71	70	Rectum	II	13	0	T3	N0	M0	IIA	50	1
Female	57	70	Ascending colon	II	16	2	T4b	N1	M1	IV	2	1
Female	78	30	Sigmoid colon	II	16	0	T3	N0	M0	II	59	0
Male	62	23	Rectosigmoid junction	III	28	0	T4a	N0	M0	II	55	0
Male	50	55	Rectum	II	7	0	T3	N0	M0	IIA	99	0
Female	73	40	Rectum	III	3	1	T3	N1a	M0	IIIB	104	0
Male	50	50	Rectum	II	7	0	T3	N0	M0	IIA	85	0
Male	90	55	Rectum	II	1	0	T3	N0	M0	IIA	98	0
Female	52	45	Rectum	II	13	0	T3	N0	M0	IIA	89	0
Male	78	60	Ileocecal junction	II	17	0	T3	N0	M0	II	47	1
Male	66	110	Descending colon	II	10	0	T4b	N0	M0	IIC	110	0
Male	53	40	Rectum	II	9	0	T3	N0	M0	IIA	94	0
Male	49	120	Rectum	II	13	0	T3	N0	M0	IIA	85	0
Female	80	45	Ascending colon	II	11	0	T3	N0	M0	II	65	0
Male	71	40	Rectum	II	4	0	T3	N0	M0	IIA	89	0
Male	51	55	Sigmoid colon	II	18	0	T4a	N0	M0	II	59	0
Male	55	50	Transverse colon	III	18	0	T4a	N0	M0	IIB	86	0
Female	61	35	Rectum	III	2	1	T3	N1a	M0	IIC	19	1
Female	71	65	Ascending colon	II	31	0	T3	N0	M0	II	52	0
Female	74	70	Sigmoid colon	II	22	8	T3	N2	M0	III	60	0
Male	83	45	Rectum	II	2	1	T3	N1a	M0	IIIB	33	1
Male	49	50	Rectum	II	3	2	T3	N1b	M0	IIIB	17	1
Male	79	60	large intestine, NOS	II	10	2	T4a	N1b	M0	IIIB	11	1
Female	63	50	Rectum	II	8	4	T3	N2a	M0	IIIB	99	0
Female	37	60	Rectum	II	6	0	T3	N0	M0	IIA	84	0
Male	80	38	Rectum	II	4	0	T3	N0	M0	IIA	98	0
Male	76	20	Rectum	III	2	0	T3	N0	M0	IIA	26	1
Male	72	80	Rectum	III	6	0	T3	N0	M0	IIA	105	0
Male	57	45	Ascending colon	II	17	0	T3	N0	M0	II	60	0

Additional Table SI. Cont.

Sex	Age	Size	Site	Grade	Lymph nodes examined	Lymph nodes positive	T	N	M	Stage	Duration	Event
Female	85	55	Sigmoid colon	II	15	2	T4a	N1b	M0	IIIB	72	0
Female	79	60	Rectum	III	6	5	T3	N2a	M0	IIIB	6	1
Male	67	75	Rectum	II	6	4	T3	N2a	M0	IIIB	90	0
Male	63	50	Rectum	III	5	0	T3	N0	M0	IIA	86	0
Female	53	40	Large intestine, NOS	II	0	0	T4a	N0	M0	IIB	92	0
Female	61	35	Rectum	III	2	1	T3	N1a	M0	IIIB	19	1
Male	59	60	Rectum	II	4	1	T3	N1a	M0	IIIB	90	0
Female	53	50	Rectum	II	5	0	T3	N0	M0	IIA	99	0
Male	74	60	Rectum	II	10	0	T3	N0	M0	IIA	43	1
Male	79	70	Ascending colon	II	18	0	T3	N0	M0	II	53	0
Male	79	50	Sigmoid colon	II	7	0	T4a	N0	M0	IIB	34	1
Female	59	50	Rectum	II	9	3	T3	N1b	M0	IIIB	108	0
Female	51	100	Rectum	II	4	0	T3	N0	M0	IIA	103	0
Male	54	60	Rectum	II	11	1	T3	N1a	M0	IIIB	23	1
Male	54	100	Rectum	III	6	0	T3	N0	M0	IIA	86	0
Male	56	90	Rectum	III	6	0	T3	N0	M0	IIA	90	0