



Time-to-event analysis of iron homeostasis to inform survival in cancer patients

Xiaoqing Fan¹, Kangna Cao¹, Xu Steven Xu², Xiaoyu Yan^{1*}

¹Guangdong-Hong Kong-Macao Joint Laboratory for New Drug Screening, School of Pharmacy, The Chinese University of Hong Kong, Hong Kong SAR, P. R. China. ²Clinical Pharmacology and Quantitative Science, Genmab Inc., Princeton, NJ, USA



INTRODUCTION

Iron is a crucial component required for a variety of vital physiological processes, such as DNA synthesis and energy metabolism that is indispensable to cell growth and division[1]. Emerging evidence has demonstrated that altered iron homeostasis is a hallmark observed in numerous types of tumors. In cancer cells, enhanced iron uptake and/or decreased iron efflux is essential to increase the intracellular iron amount to support their increased demands of survival, rapid cell division, and metastasis[2]. However, the amount of iron required by diverse tumors is different due to their remarkable heterogeneity. How the expression level of proteins that control iron uptake and efflux changes and coordinates to generate a unified effect on iron homeostasis is an important open question. These quantitative interconnections are important for guiding novel interventions for cancer therapy that involve iron homeostasis perturbations, such as iron-chelator-based therapy and ferroptosis-inducing drugs[3-5].

OBJECTIVES

- To develop parametric time-to-event (TTE) models using the expression profile of the major proteins that control iron uptake and efflux to quantify the contribution of the expression level of these proteins on the outcome of cancer patients.
- To create a cancer-specific iron risk score (CIRS) based on TTE modeling results to predict the long-term individual survival outcomes of patients.

METHODS

- **Data collection:** Based on the literature review, we found eight proteins is involved in iron transport, including transferrin receptor 1 (TFR1), divalent metal transporter 1 (DMT1), lipocalin-2 (LCN2), cluster determinant 44 (CD44), heme carrier protein 1 (HCP1), transmembrane ion channels Zrt-/Irt-like protein (ZIP) 8, and ZIP14 for iron uptake, and ferroportin (FPN) for iron efflux (Table 1) [6-9]. To investigate the major mediators for iron uptake or export in cancer patients, we collected human data from UALCAN (<https://ualcan.path.uab.edu/index.html>), a comprehensive database that integrates publicly available cancer OMICS data, including The Cancer Genome Atlas [TCGA] and Clinical Proteomic Tumor Analysis Consortium [CPTAC] database [10]. The RNA expressions of the iron transport-related proteins in cancer patients were compared with the corresponding normal subjects from TCGA. We only included cancer types that contained more than 10 paired normal samples to ensure the strength of statistical power. Therefore, 6903 TCGA patient samples and 709 normal subject samples across 16 cancer types were analyzed.

Table 1. Summary of iron transport-related proteins.

Function	Name	Gene symbol	Ensembl	Description
Uptake	CD44	CD44	ENSG00000026508	Hyaluronate/cluster determinant 44
	DMT1	SLC11A2	ENSG00000110911	Divalent metal transporter 1
	HCP1	SLC46A1	ENSG00000076351	Heme carrier protein
	LCN2	SLC22A17	ENSG00000092096	Solute carrier family 22 member 17, lipocalin-2
	TFR1	TFRC	ENSG00000072274	Transferrin receptor 1
	ZIP14	SLC39A14	ENSG00000104635	Solute carrier family 39 member 14, transmembrane ion channel
	ZIP8	SLC39A8	ENSG00000138821	Solute carrier family 39 member 8, transmembrane ion channel
Efflux	FPN	SLC40A1	ENSG00000138449	Solute carrier family 40 member 1, ferroportin

- **TTE Model Development:** To evaluate the impact of patient iron transport-related proteins (TFR1, CD44, and FPN) expression level on the observed overall survival (OS) and to stratify the iron homeostasis-based risk, parametric TTE models were developed and evaluated based on the RNA expression data.

$$h(t) = \lambda_{OS} \cdot e^{\alpha t} \cdot (1 + \frac{E_{maxTFR1} \cdot TFR1^{\gamma_1}}{TFR1_{50}^{\gamma_1} + TFR1^{\gamma_1}}) \cdot (1 + \frac{E_{maxCD44} \cdot CD44^{\gamma_2}}{CD44_{50}^{\gamma_2} + CD44^{\gamma_2}}) \cdot (1 - \frac{E_{maxFPN} \cdot FPN^{\gamma_3}}{FPN_{50}^{\gamma_3} + FPN^{\gamma_3}}) \quad (Eq\ 1)$$

$$h(t) = \lambda_{OS} \cdot \alpha \cdot (\lambda_{OS} \cdot t)^{\alpha-1} \cdot (1 + \frac{E_{maxTFR1} \cdot TFR1_t^{\gamma_1}}{TFR1_{50}^{\gamma_1} + TFR1_t^{\gamma_1}}) \cdot (1 + \frac{E_{maxCD44} \cdot CD44_t^{\gamma_2}}{CD44_{50}^{\gamma_2} + CD44_t^{\gamma_2}}) \cdot (1 - \frac{E_{maxFPN} \cdot FPN_t^{\gamma_3}}{FPN_{50}^{\gamma_3} + FPN_t^{\gamma_3}}) \quad (Eq\ 2)$$

- Where h(t) is the hazard of death, λ_{OS} is the hazard scale factor, α is the hazard shape factor, $E_{maxTFR1}$, $E_{maxCD44}$, and E_{maxFPN} are maximum effect parameters of TFR1, CD44, and FPN on OS, respectively. $TFR1_{50}$, $CD44_{50}$, and FPN_{50} are half-expression levels of TFR1, CD44, and FPN that are needed to reach the maximum effect on OS, respectively. γ_1 , γ_2 , and γ_3 are the slope-factor for the effect of TFR1, CD44, and FPN, respectively.
- **Model evaluation:** Using the final covariate model and the distribution of the TCGA datasets, simulation-based (1000 replicates) visual predictive check (VPCs) were compared for different models. For each simulation, Kaplan–Meier proportions of the surviving patients were computed across time for evaluation of median OS and the 90% confidence intervals around simulations were included based on the 5th and 95th percentiles of the estimated covariate effect parameters.

- **Iron homeostasis-based risk stratification:** To exploit the potential of iron homeostasis profiles at baseline to predict long-term survival outcomes across various cancers, we developed a cancer-specific iron risk score (CIRS) based on the TTE modeling estimation results. We divided the patients into tertile groups based on stratified CIRS. Specifically, the risk of a cancer patient was deemed to be substantially associated with the expression level of TFR1, CD44, and FPN and was classified as (Table 2): (i) For TFR1 and CD44 expression levels higher than $TFR1_{50}$ and $CD44_{50}$ values and the FPN expression level lower than FPN_{50} values, the patient is regarded as high risk; (ii) For TFR1 and CD44 expression levels lower than $TFR1_{50}$ and $CD44_{50}$ values while the FPN expression level higher than FPN_{50} values, the patient is regarded as low risk; (iii) The patient outside these two groups is classified into the intermediate risk. Then, the relationship between CIRS and patient survival was plotted with the Kaplan–Meier survival curves using the Survival and Surminer R packages.

Table 2. Cancer-specific iron risk score (CIRS) development based on the TTE modeling estimation results.

Proteins	Risk		
	High	Low	Intermediate
TFR1	> $TFR1_{50}$	< $TFR1_{50}$	The patient outside the high and low-risk group
CD44	> $CD44_{50}$	< $CD44_{50}$	
FPN	< FPN_{50}	> FPN_{50}	

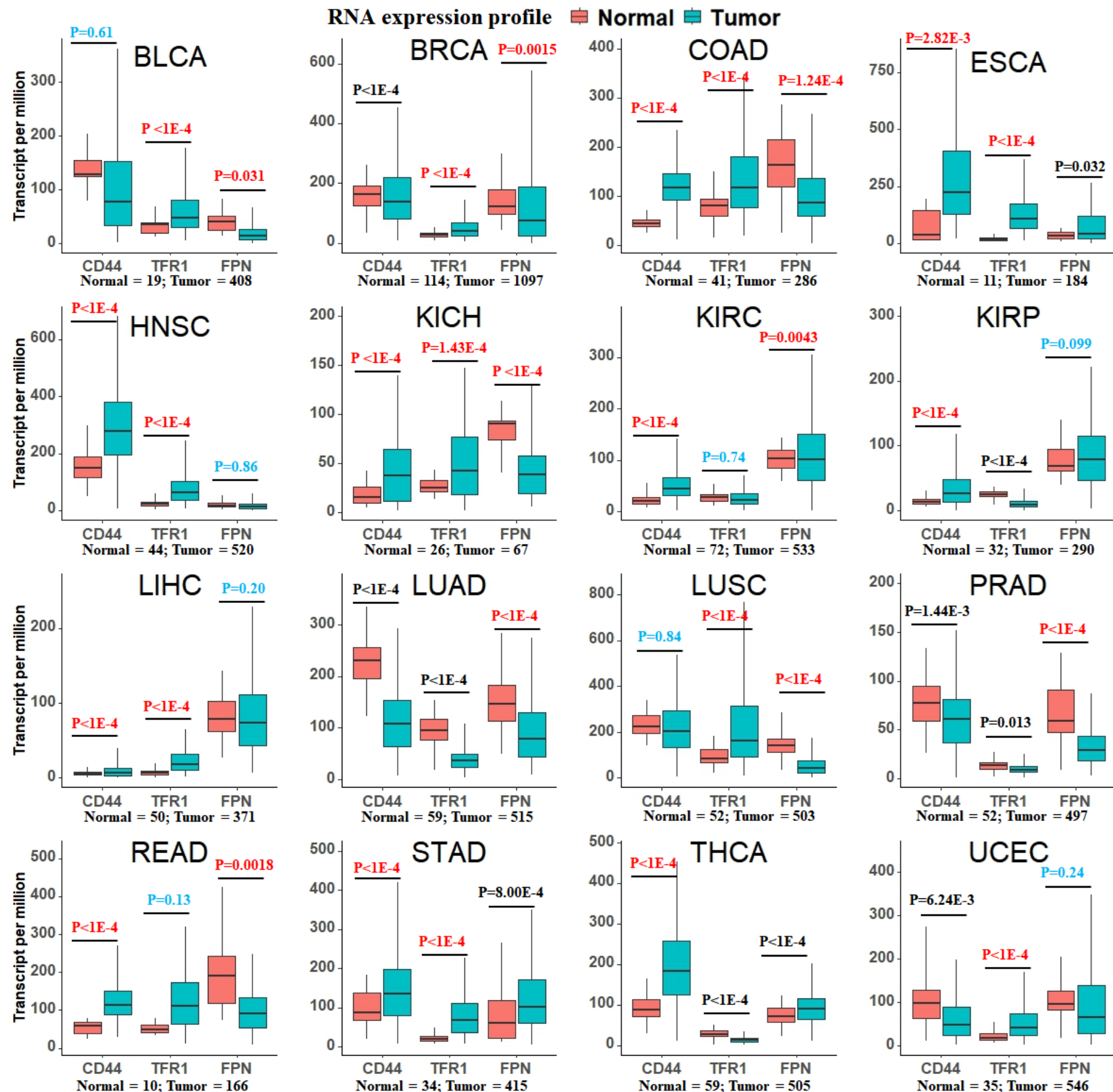
- **Software:** NONMEM7.5 FOCEI

REFERENCES

- [1] Trends Pharmacol Sci. 2021; 42: 640-56.
- [2] Annu Rev Nutr. 2018; 38: 97-125.
- [3] Biometals. 2020; 33: 201-15.
- [4] Cancer Res. 2020; 80: 5435-48.
- [5] Nat Cell Biol. 2024; 26(9):1386-1393.
- [6] Nat Chem. 2020; 12: 929-38.
- [7] Trends Pharmacol Sci. 2024; 45: 145-56.
- [8] Semin Cancer Biol. 2018; 53: 125-38.
- [9] Nat Rev Mol Cell Biol. 2024; 25: 133-55.
- [10] Neoplasia. 2022; 25: 18-27.

RESULTS

- **The pan-cancer expression landscape of the iron transport-related proteins.**



Notes: BLCA: Bladder urothelial carcinoma; BRCA: Breast invasive carcinoma; COAD: Colon adenocarcinoma; ESCA: Esophageal carcinoma; HNSC : Head and Neck squamous cell carcinoma; KICH: Kidney chromophobe; KIRC: Kidney renal clear cell carcinoma; KIRP : Kidney renal papillary cell carcinoma; LIHC: Liver hepatocellular carcinoma; LUAD: Lung adenocarcinoma; LUSC: Lung squamous cell carcinoma; PRAD : Prostate adenocarcinoma; READ: Rectum adenocarcinoma; STAD: Stomach adenocarcinoma; THCA: Thyroid carcinoma; UCEC: Uterine corpus endometrial carcinoma.

Table 3. Final parameters from time-to-event analysis of overall survival in patients across 15 cancers from TCGA with baseline measurements of TFR1, CD44, and FPN.

Cancer types	Model Structure	Parameters										
		λ_{η}	α	$E_{maxTFR1}$	$TFR1_{50}$	γ_1	$E_{maxCD44}$	$CD44_{50}$	γ_2	E_{maxFPN}	FPN_{50}	γ_3
BLCA	Gompertz	0.0073	-0.13	27.71	6.01	121.5	0.5	4.65	112.1	0.15	16.5	121.1
BRCA	Gompertz	0.0016	0.19	6.89	2.84	102.2	2.10	0.42	11.15	0.55	7.56	106.5
COAD	Gompertz	0.0076	0.032	10.31	9.85	137.5	0.25	0.12	3.20	0.27	11.05	124.5
HNSC	Weibull	0.020	-0.21	2.61	6.45	117.1	1.67	17.58	66.57	0.24	6.27	155.2
KICH	Weibull	0.0056	0.050	12.72	20.28	140.8	1.67	5.62	136.7	0.71	6.49	149.9
KIRC	Gompertz	0.0071	-0.049	14.31	1.75	180.6	1.96	18.19	15.36	0.92	25.52	1.64
KIRP	Weibull	0.043	0.21	2.14	10.29	18.09	0.86	1.35	134.1	0.58	9.93	149.8
LIHC	Weibull	0.19	-0.086	1.12	12.78	6.66	0.71	9.79	155.3	0.47	21.36	113.5
LUAD	Weibull	0.045	0.056	3.51	0.57	19.69	0.32	10.86	138.7	0.48	8.44	123.8
LUSC	Weibull	0.027	-0.08	5.42	5.22	12.91	0.80	148	125.7	0.22	57.33	127.6
PRAD	Weibull	0.00049	1.88	2.15E5	23.9	217.7	1.38	3.92	158.8	0.64	6.64	175.3
READ	Weibull	0.0004	0.57	54.42	8.39	138.9	4.5	23.22	161.2	0.62	95.64	121.5
STAD	Gompertz	0.045	0.066	4.44	0.91	17.61	1.87	10	100.6	0.58	8.21	130.2
THCA	Weibull	0.0014	0.39	19.44	4.07	217.4	0.84	6.07	15.88	0.70	16.29	70.91
UCEC	Weibull	0.0093	-0.0022	2.26	9.03	7.62	1.36	0.03	6.47	0.39	39.7	3.52

- **TTE models predict long-term survival in patients across diverse cancers.**

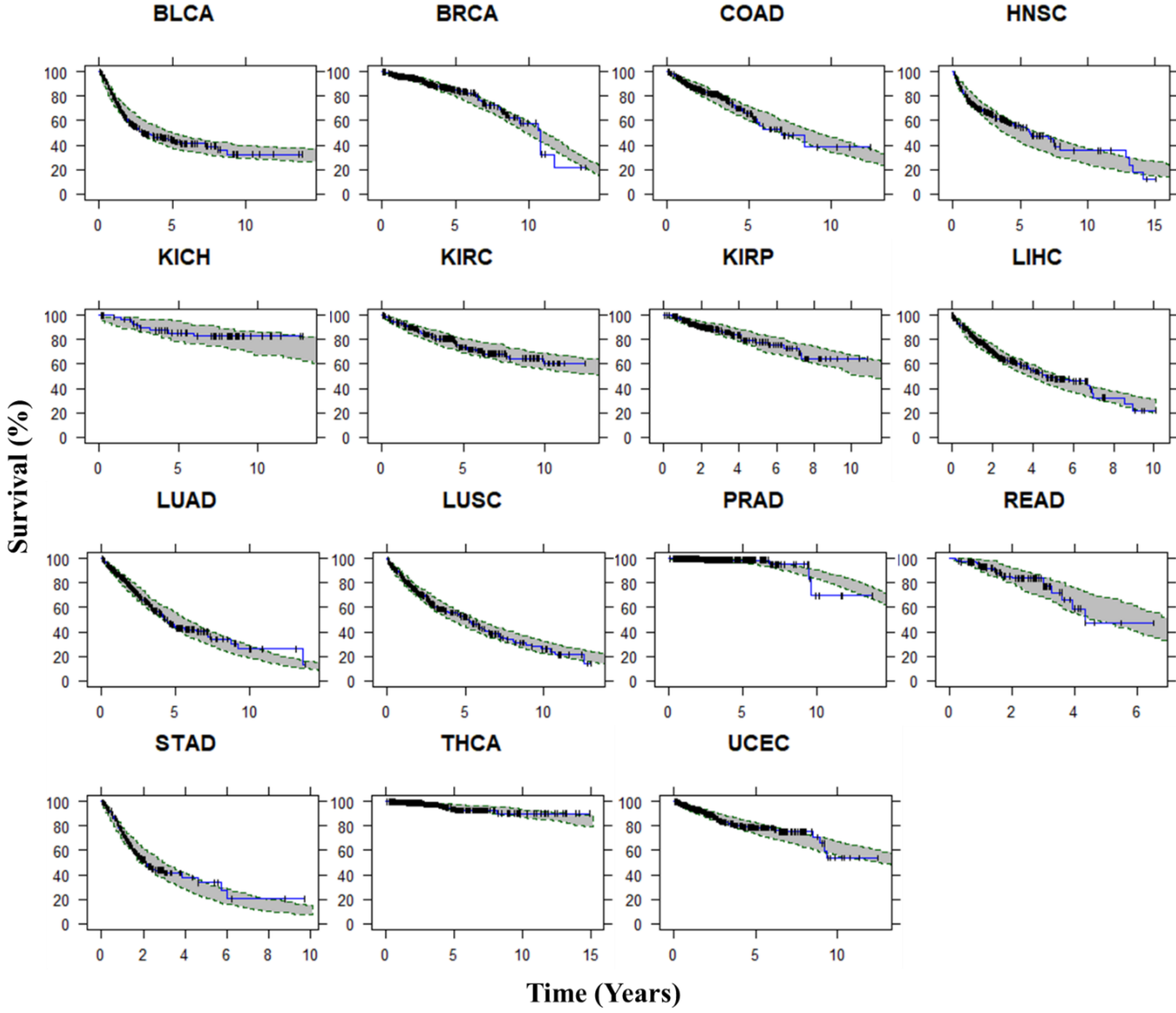


Figure 2. Visual predictive check for the time-to-event models of survival in different types of cancers. The solid line in each panel represents the observed Kaplan–Meier curve. The shaded area shows 95% confidence interval of the model-predicted Kaplan–Meier curve based on 1000-time simulations.

- **CIRS stratifies the risk of long-term survival in cancers.**

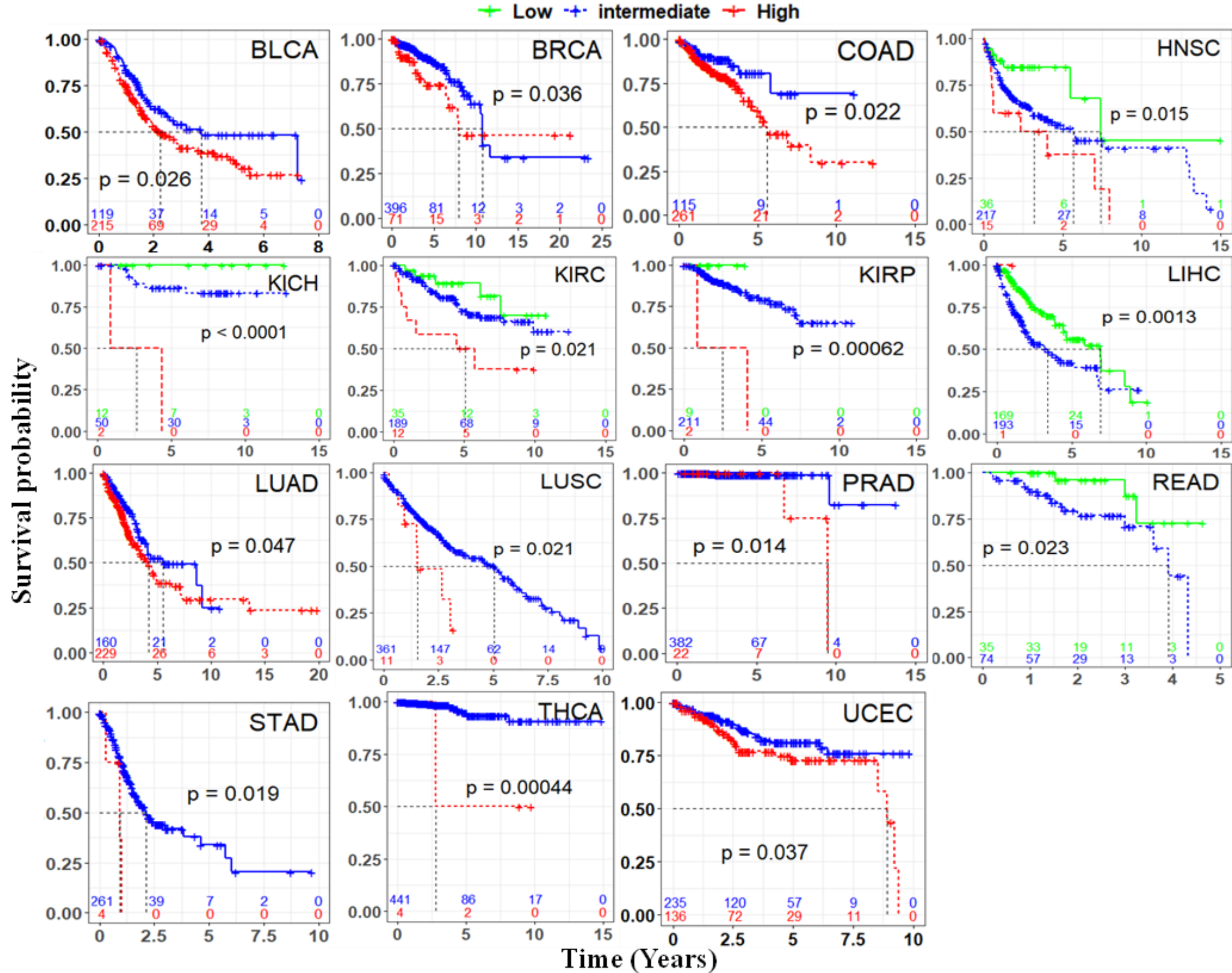


Figure 3. Cancer-specific iron risk score (CIRS) could be applied to stratify the survival probability in cancers. Kaplan-Meier survival plots of 15 cancer categories and all-cause mortality according to tertiles of CIRS during follow-up (green, bottom tertile; blue, median tertile; red, top tertile). Cancer types are labeled using the TCGA study abbreviations.

CONCLUSIONS

- This study provided useful insights to identify cancer patient subgroups at higher risk based on predefined CIRS using RNA-seq data, and guide optimal clinical interventions regarding the significant predictors, which would constitute a novel strategy for cancer diagnosis, staging, and lead to potential therapeutic strategies to target these tumor-specific events.

Corresponding Author: Xiaoyu Yan, PhD, xiaoyuyan@cuhk.edu.hk