

Research Paper

A Refined Staging Model for Resectable Pancreatic Ductal Adenocarcinoma Incorporating Examined Lymph Nodes, Location of Tumor and Positive Lymph Nodes Ratio

Yunda Song^{1,2*}, Zhenxin Chen^{1,2*}, Luohai Chen^{3*}, Chaobin He^{1,2}, Xin Huang^{1,2}, Fangting Duan¹, Jun Wang^{1,4}, Xiangming Lao^{1,2}, Shengping Li^{1,2}✉

1. State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou 510060, P. R. China.
2. Department of Hepatobiliary and Pancreatic Oncology, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China
3. Department of Gastroenterology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, P. R. China
4. Department of Ultrasonics, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China

*These authors contributed equally to this work

✉ Corresponding author: Shengping Li, Department of Hepatobiliary and Pancreatic Oncology, Sun Yat-sen University Cancer Center, No. 651 Dongfeng Road East, Guangzhou 510060, People's Republic of China. Tel: +862087343572; Fax: +862087343572; Email: lishp@sysucc.org.cn

© Ivyspring International Publisher. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY-NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2018.03.20; Accepted: 2018.07.25; Published: 2018.09.08

Abstract

Background: Nodal status and tumor site are prognostic factors for resectable pancreatic ductal adenocarcinoma (PDAC). Parameters for nodal status are diverse, and the number of examined lymph nodes (eNs) needed for good prognosis are uncertain. We try to modify staging system of resectable PDAC with parameters mentioned above by recursive partitioning analysis.

Methods: Patients from the Surveillance, Epidemiology, and End Results (SEER) database were divided into training cohort and internal validation cohort, randomly. PDAC patients from Sun Yat-sen University Cancer Center were regarded as external validation cohort. The training cohort was used to refine staging model by recursive partitioning analysis, while the internal validation cohort and the external validation cohort were applied to assess discriminatory capacity of staging model. For parameters included in the modified model, their effects were studied.

Results: The number of eNs, tumor site and tumor size were risk factors for positive nodal status. Lymph nodes ratio (LNR), tumor site, eNs and T stages of 8th the American Joint Committee on Cancer (AJCC) were selected to develop a refined model, dividing patients into 5 groups of different outcomes, preceding 8th AJCC classification. Besides, we found that (1) for small PDAC (diameter < 1cm), lymph node metastasis was rarely found; (2) enough eNs were needed to ensure better prognosis of node-negative patients; (3) tumors in the head of pancreas were prone to lymph nodes metastasis; (4) for node-positive patients, LNR was a better nodal parameter compared to positive lymph nodes (pNs).

Conclusion: Our improved staging system helps to illuminate the interactions among tumor site, size and eNs.

Key words: resectable pancreatic ductal adenocarcinoma, staging scheme, positive lymph nodes ratio, examined lymph nodes, head of pancreas, body and tail of pancreas

Introduction

Pancreatic ductal adenocarcinoma (PDAC), characterized with insidious onset and early metastasis, is one of the most life-threatening diseases with an extremely low 5-year survival rate, only 6% in

USA [1]. Even after resection with curative intent, most patients suffer from recurrence [2].

The prognosis of PDAC is evaluated by American Joint Committee on Cancer (AJCC) staging

system. Adding a N2 classification for nodal status (N0: no pN; N1: 1-3 pNs; N2: ≥ 4 pNs) and only considering tumor size regardless of extrapancreatic extension (T1: ≤ 2 cm, T2: >2 cm and ≤ 4 cm, T3: >4 cm), the discriminatory power of 8th AJCC schemes is still comparable with 7th AJCC schemes [3, 4].

Significant differences in lymphatic backflow and innervations between the head and body/tail of the pancreas have prognostic effect on PDAC [5], which is not reflected on the 8th AJCC staging scheme.

Nodal status is a potent prognostic factor [6-9]. Ignoring total number of examined lymph nodes (eNs), pNs perform well when the number of eNs is more than 20 [10, 11]. Besides, 15 eNs are needed to ensure node-negative status of PDAC [12]. However, the number of eNs is insufficient in many cases, which results in false negative nodal status [13], and the number of eNs needed for good prognosis is uncertain.

The lymph nodes ratio (LNR, pNs divided by eNs) is selected as a parameter for nodal status and occasionally proved to be superior to pNs, especially for node-positive PDAC [14-16].

Our aim is to identify the number of eNs needed for good prognosis, to modify the staging schemes for resectable PDAC by recursive partitioning analysis (RPA), and to explore the interactions among parameters included.

Methods

Patients

The Surveillance, Epidemiology, and End Results (SEER) database was used. We identified patients with PDAC (ICD-O-3 codes: 8140, 8150, 8210, 8211, 8251, 8260, 8261, 8263, 8480, 8481, 8490, 8500, and 8503) from 2004 to 2014. Exclusion criteria included: history of prior malignancy, T4 stage in 8th AJCC scheme, distant metastasis and missing information regarding overall survival (OS), tumor sizes, number of pNs and eNs. Finally, 8480 patients, which were further divided into training cohort (n=5936) and internal validation cohort (n=2544) randomly, were included.

The external validation cohort consisted of patients from Sun Yat-sen University Cancer Center (SYSUCC, from 2001 to 2016). All the patients (n=92) were pathologically diagnosed with PDAC. Patients with missing data on tumor sizes, nodal status or survival were excluded.

Statistical Analysis

OS was regarded as primary outcome. Kaplan-Meier method and log-rank tests were applied to assess prognostic effect. Cox proportional hazards regression was used for univariate analysis

and multivariable analysis. To identify risk factors of positive nodal status, Logistic regression was applied [17].

To modify the staging scheme, RPA, which was able to classify patients into groups of maximum separation [18, 19], was applied. R package named "rpart" was used to develop RPA model, classifying patients into different subgroups with different prognosis using selected parameters automatically, and "rpart.plot" package was applied to visualize result of RPA. For subgroups with similar median survival time, we merged them into bigger groups manually to generate our refined staging system.

Concordance index (c-index), ranging from 0.5 to 1, was used for quantifying discriminatory capacity. The higher c-index, the better discriminatory capacity [20]. When assessing prognostic homogeneity, we compared OS within each stage of our model by 8th AJCC stages and likelihood ratio χ^2 test was also applied.

$P < 0.05$ was regarded as statistical significance. IBM SPSS Statistics for Windows v.19.0 and R v.3.4.1 were used for statistical analysis.

Results

Characteristics of Patients from SEER Database

Characteristics of 8480 patients from SEER database were shown in Table S1. Most of tumors were located in head of pancreas (84.5%), with the greatest dimension between 2cm and 4cm (59.6%), while median OS was 20 months.

For node-positive patients, more eNs were found ($p < 0.001$), locations of tumor tended to be head of pancreas ($p < 0.001$) and sizes of tumor were larger ($p < 0.001$) (Table S2).

Tumors in head of pancreas tended to have more eNs ($p < 0.001$), more pNs ($p < 0.001$) and smaller tumor sizes ($p < 0.001$). Besides, LNR tended to be higher for PDAC in head (LNR: head of pancreas: 0.175, 95% confidence interval (CI) 0.171-0.180; LNR: body and tail of pancreas: 0.124, 95% CI 0.113-0.134, $p < 0.001$), which implies that tumors in head are prone to lymph node metastasis (Table S3).

Characteristics of Patients from SYSUCC

Characteristics of 92 patients from SYSUCC were shown in Table 1. Median age at diagnosis was 59 years old. Nearly 60% patients were male. Most of tumors were located in head of pancreas (83.7%). The median OS was 23 months.

Risk Factors for Positive Nodal Status

For patients from SEER database, we identified tumor site (HR=2.132, 95% CI 1.876-2.423, $p < 0.001$),

tumor size (HR=1.016, 95% CI 1.013-1.020, $p<0.001$) and the number of eNs (HR=1.050, 95% CI 1.044-1.056, $p<0.001$) as risk factors for positive nodal status, while race ($p=0.057$), sex ($p=0.190$), age of diagnosis ($p=0.381$) and histologic type of PDAC ($p=0.174$) were not.

Table 1. Characteristics of patients from Sun Yat-sen University Cancer Center.

Variable	Median(IQR)/N%
Age(years)	59(50-66)
Sex	
Male	56(60.9%)
Female	36(39.1%)
Tumor site	
head	77(83.7%)
body or tail	15(16.3%)
Tumor differentiation	
well-moderate	6(6.5%)
Moderate	36(39.1%)
Moderate-poor	31(33.7%)
poor	15(16.3%)
not determined	4(4.3%)
Tumor size	
≤2cm	9(9.8%)
>2cm and ≤4cm	54(58.7%)
>4cm	29(31.5%)
number of positive lymph node	0(0-1)
0 positive lymph node	50(54.3%)
1-3 positive lymph nodes	34(40.0%)
≥4 positive lymph nodes	8(8.7%)
number of examined lymph nodes	8.5(4-14)

Abbreviations: IQR = interquartile range

Prognostic Factors

In training cohort, for node-positive patients, univariate analysis identified age of diagnosis ($p<0.001$), year of diagnosis ($p<0.001$), grade of PDAC ($p=0.002$), tumor size ($p<0.001$), the number of pNs ($p<0.001$) and eNs ($p<0.001$) and LNR ($p<0.001$) as risk factors, while sex ($p=0.119$), histologic type ($p=0.319$), race ($p=0.819$), tumor site ($p=0.355$), marital status ($p=0.197$) were not. Multivariable analysis showed that age of diagnosis (HR=1.012, 95% CI 1.008-1.015, $p<0.001$), year of diagnosis (HR=0.982, 95% CI 0.969-0.996, $p=0.011$), grade (HR=1.036, 95% CI 1.014-1.059, $p=0.001$), tumor size (HR=1.002, 95% CI 1.001-1.003, $p=0.001$), pNs (HR=1.025, 95% CI 1.006-1.044, $p=0.011$), eNs (HR=0.991, 95% CI 0.985-0.997, $p=0.005$) and LNR (HR=1.912, 95% CI 1.912-2.563, $p<0.001$) were independent risk factors for node-positive PDAC.

For nodes-negative patients, similarly, age of diagnosis (HR=1.019, 95%CI 1.013-1.024, $p<0.001$), year of diagnosis (HR=0.950, 95%CI 0.929-0.971, $p<0.001$), tumor site (HR=1.119, 95%CI 1.027-1.378, $p=0.021$), eNs (HR=0.987, 95%CI 0.980-0.994, $p<0.001$) and tumor size (HR=1.009, 95%CI 1.007-1.011, $p<0.001$) were independent risk factors for node-negative PDAC.

Refined Model in Training Cohort

For node-positive patients, risk factors of positive nodal status (T stage of 8th AJCC, eNs and site of tumor) and parameters for nodal status (LNR and pNs) were put into RPA. For node-negative patients, risk factors of positive nodal status (T stage of 8th AJCC, eNs, site of tumor) were used. LNR, T stage of 8th AJCC and eNs were selected as parameters of our staging system by RPA for node-positive patients, while T stage, eNs and tumor locations for node-negative patients. After RPA was done for node-positive and node-negative patients respectively, we combine them together and merge subgroups with similar median survival time into the same group, as it was in Figure S1. For convenience, we summarized results of RPA in Table 2, which was easier to understand and use. Remarkably, for node-positive patients, LNR was the first parameter to be taken into consideration by RPA, which implied that LNR was of the great importance in prognosis of patients with positive nodal status, corresponding with multivariable analysis above.

Table 2. Refined staging scheme for resectable PDAC by recursive partitioning analysis. MS: median survival (R0: LNR=0, R1: $0<LNR<0.24$, R2: $0.24\leq LNR<0.45$ and R3: $LNR\geq 0.45$; T1: ≤ 2 cm, T2: >2 cm and ≤ 4 cm, T3: >4 cm; head: head of pancreas, body&tail: body and tail of pancreas; RPA: recursive partitioning analysis).

R	T	eNs	site	stage
R0	T1	Any	Body&Tail	RPA-I
		eNs>2.5	Head	RPA-I
	T2	eNs<2.5	Head	RPA-III
		eNs>5.5	Any	RPA-II
		eNs<5.5	Body&Tail	RPA-II
			Head	RPA-III
T3	eNs>5.5	Any	RPA-III	
		Body&Tail	RPA-III	
	eNs<5.5	Body&Tail	RPA-III	
		Head	RPA-IV	
R1	T1	Any	Any	RPA-II
	T2, T3	eNs>9.5	Any	RPA-III
		eNs<9.5	Any	RPA-IV
R2	T1	Any	Any	RPA-III
	T2, T3	Any	Any	RPA-IV
R3	Any	Any	Any	RPA-V

LNR, which was selected by RPA, was divided as follow: R0 (LNR=0), R1 ($0<LNR<0.24$), R2 ($0.24\leq LNR<0.45$) and R3 ($LNR\geq 0.45$) (Figure 1a). Diagnostic effects of T stage were shown in Figure 1b. Our staging scheme classified patients into 5 groups: RPA-I (R0T1 in body&tail of pancreas and R0T1 in head of pancreas with eNs >2.5), RPA-II (R0T2 with eNs >5.5, R0T2 in the body&tail of pancreas with eNs<5.5 and R1T1), RPA-III (R0T1 in head of pancreas with eNs <2.5, R0T2 in the head of pancreas with eNs <5.5, R0T3 with eNs >5.5, R0T3 in the body&tail of pancreas with eNs<5.5, R1T2-T3 with eNs >9.5 and R2T1), RPA4-IV (R0T3 in the head of pancreas with

eNs <5.5, R1T2-T3 with eNs <9.5 and R2T2-T3) and RPA-V (R3T1-T3) (Table 2, Figure S1). Notably, for node-negative patients, 6 eNs or more implied good prognosis.

488, 1303, 2376, 1174 and 595 patients were included in RPA-I, RPA-II, RPA-III, RPA-IV and RPA-V groups with a median survival of 47, 26, 20, 15 and 12 months ($p < 0.001$), respectively (Figure 1c). Adjusted with sex, age of diagnosis, histologic type,

race, year of diagnosis, marital status and grade, our model showed that the advanced stage correlated with elevated risk of poor prognosis (RPA-II vs RPA-I: HR 1.644, 95% CI 1.412-1.914, $p < 0.001$; RPA-III vs RPA-I: HR 2.261, 95% CI 1.959-2.610, $p < 0.001$; RPA-IV vs RPA-I: HR 3.263, 95% CI 2.809-3.790, $p < 0.001$; RPA-V vs RPA-I: HR 4.026, 95% CI: 3.427-4.731, $p < 0.001$).

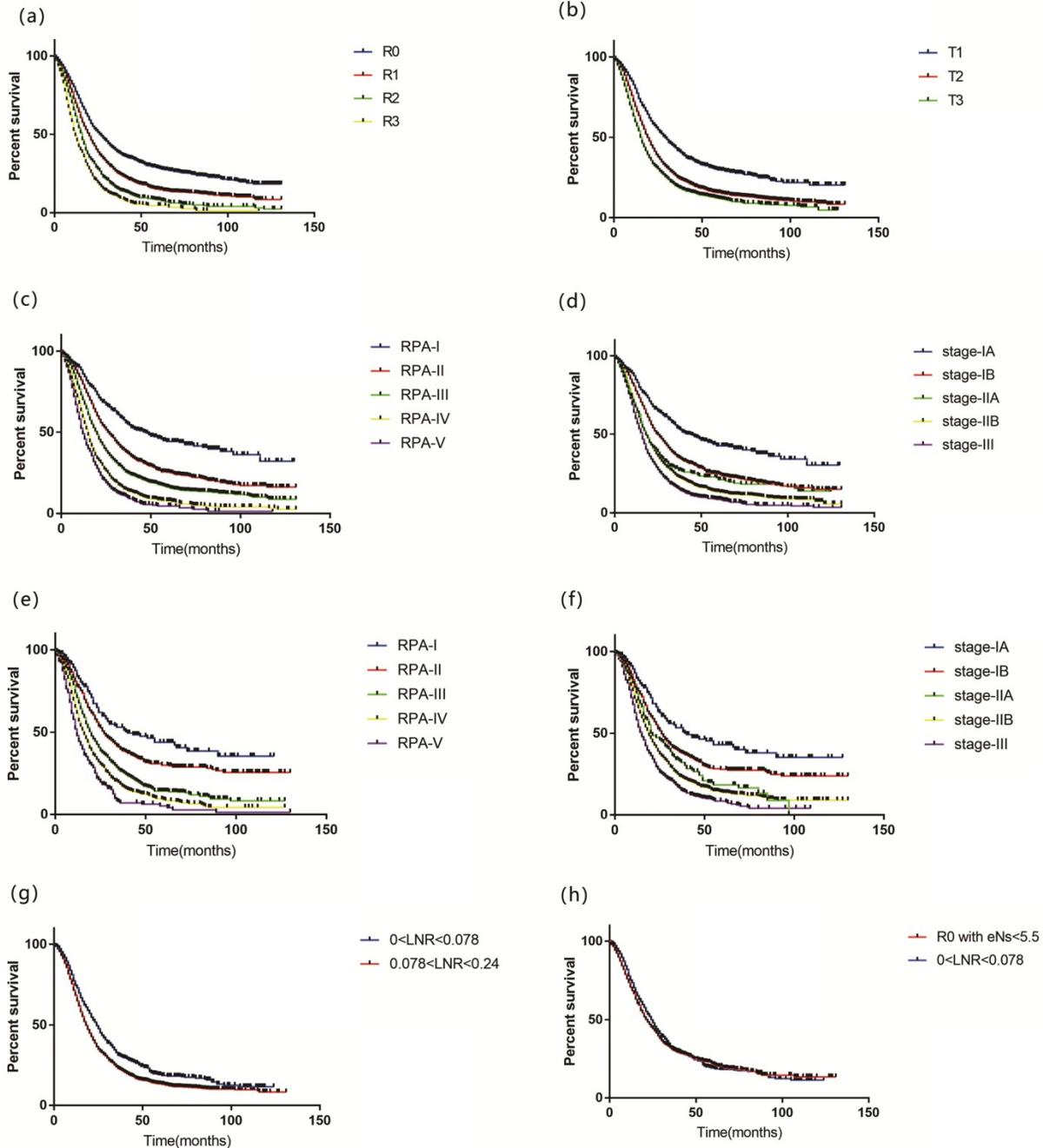


Figure 1. Overall survival for patients by (a) R stage, (b) T stage, (c) RPA model, (d) 8th AJCC model in training set; (e) RPA model, (f) 8th AJCC model in internal validation set; (g) $LNR > 0.078$ and $LNR < 0.078$ in R1 stage of RPA model; (h) $0 < LNR < 0.078$ and R0 with eNs < 5.5 of RPA model. (R0: LNR=0, R1: $0 < LNR < 0.24$, R2: $0.24 \leq LNR < 0.45$ and R3: $LNR \geq 0.45$; T1: ≤ 2 cm, T2: > 2 cm and ≤ 4 cm, T3: > 4 cm; RPA: recursive partitioning analysis; LNR: positive lymph nodes ratio; eNs: examined lymph nodes; AJCC: American Joint Committee on Cancer)

Evaluation of Performance

In training set, compared to the 8th AJCC staging scheme (Figure 1d) (c-index: 0.577, 95% CI 0.567-0.587), our staging system displayed a better performance of discriminatory ability (c-index: 0.606, 95% CI 0.596-0.616, $p < 0.001$). For prognostic homogeneity, our staging scheme (likelihood ratio χ^2 test= 547.2) was also better than 8th AJCC scheme (likelihood ratio χ^2 test= 345.4) (Table 3). Our model can classify patients of 8th AJCC stages into subgroups with different OS. Stratified by 8th AJCC stages, OS was homogeneous for patients in our scheme (Table 4).

Table 3. Comparison of the performance of our refined model and 8th AJCC staging system

	c-index		likelihood ratio *2 test	
	8 th AJCC	RPA model	8 th AJCC	RPA model
training cohort of SEER	0.577	0.606	345.4	547.2
internal validation cohort of SEER	0.583	0.597	157.2	208.1
external validation cohort of SYSUCC	0.558	0.565	2.64	4.24

Characteristics were comparable between training and internal validation cohort (Table S4). For internal validation cohort, c-index of our staging system (c-index: 0.597, 95% CI: 0.581-0.613) was higher than 8th AJCC (c-index: 0.583, 95% CI 0.567-0.599, $p = 0.004$) (Figure 1e, 1f). Stage-IB, stage-IIB and stage-III in 8th AJCC staging scheme (likelihood ratio χ^2 test= 157.2) can be classified by our staging system (likelihood ratio χ^2 test=208.1) into subgroups with distinct OS, while OS was homogeneous in each stage of our system stratified by 8th AJCC scheme (Table 3, Table S5).

For patients from SYSUCC, our staging model also displayed a better performance of discriminatory ability (c-index of our refined model was 0.565, while c-index of 8th AJCC staging scheme was 0.558) and prognostic homogeneity (likelihood ratio χ^2 test of our refined model was 4.24, while likelihood ratio χ^2 test of 8th AJCC staging system was 2.64).

Tumor Size and Nodal Status

Higher proportion of positive nodal status was found in high T stage ($p < 0.001$) (Table S6). Patients of T1 tended to have less pNs than T2-3 patients

($p < 0.001$, Figure 2a). A cumulative frequency diagram of PDAC with diameter less than 4cm showed that when diameters were less than 10mm, few events of metastasis were found (Figure 2b).

Effects of Tumor Site on Prognosis

For patients of R0T1 in our scheme, OS was significantly different between different sites of tumor ($p = 0.004$, Figure 2c). Besides, for patients of T1 or R0 with eNs <5.5 , OS was significantly different between different sites ($p < 0.001$, $p = 0.040$, respectively. Figure 2d, 2h). However, for patients of T3, R0 and R3, OS of different sites were homogeneous ($p = 0.283$, $p = 0.460$ and $p = 0.651$, respectively (Figure 2e, 2f and 2g).

Comparisons: R0 with eNs <5.5 Patients and 0 $<$ LNR <0.078 Patients

Patients of R1 in our model were further divided using LNR by RPA. It was revealed that prognosis of patients with 0 $<$ LNR <0.078 was better than others ($p < 0.001$, Figure 1g). There was no significant difference in OS between patients of R0 with eNs <5.5 and patients with 0 $<$ LNR <0.078 ($p = 0.341$, Figure 1h), which supported that if eNs were less than 6, negative nodal status may be false-negative (Table S7).

Discussion

Our staging model shows better performance than 8th AJCC scheme for resectable PDAC in internal validation cohort from SEER database and external validation cohort from SYSUCC. Site of tumor and eNs were introduced into our scheme, and the number of pNs is replaced by LNR.

Nodal status is a potent prognostic factor for resectable PDAC [6]. pNs were selected as parameter for nodal status in 8th AJCC staging scheme. However, there are different clinical meanings for finding a positive node with different eNs. LNR, considering both pNs and eNs, was said to have preeminent prognostic effects [14-16]. Multivariable analysis confirmed that for node-positive patients, prognostic effect of LNR (HR=1.912) was more powerful than pNs (HR=1.025) in this study. After putting both pNs and LNR together into RPA, LNR is chosen in our model, implying that LNR may be a better parameter for node-positive patients.

Table 4. Prognostic homogeneity comparison between 8th AJCC staging system and our staging system in training set

	8 th AJCC-IA		8 th AJCC-IB		8 th AJCC-IIA		8 th AJCC-IIB		8 th AJCC-III		p-value
	n	MS	n	MS	n	MS	n	MS	n	MS	
RPA-I	488	47									
RPA-II			946	27			315	26	42	30	$p = 0.406$
RPA-III	35	16	184	20	356	20	1447	20	354	18	$p = 0.460$
RPA-IV					49	13	554	14	571	16	$p = 0.600$
RPA-V							136	10	459	13	$p = 0.119$
p-value	$P < 0.001$		$p < 0.001$		$p = 0.001$		$p < 0.001$		$p < 0.001$		

Abbreviations: AJCC= American Joint Committee on Cancer; MS=median survival; RPA=recursive partitioning analysis.

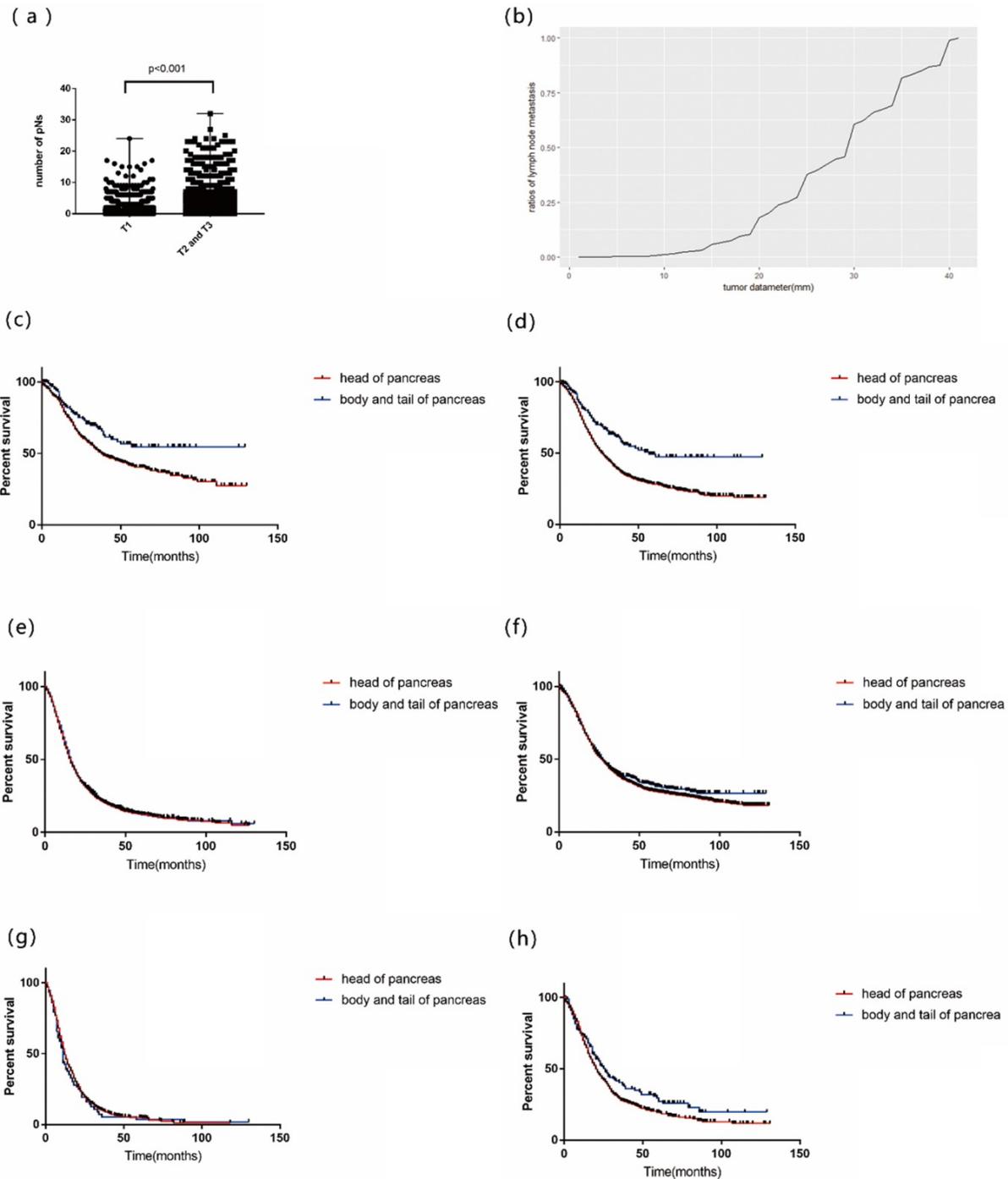


Figure 2. (a) the number of positive lymph nodes (pNs) in patients of T1 stage and T2-3 stage; (b) A cumulative frequency diagram of tumor diameter and frequency of lymph nodes metastasis; Overall survival by tumor site for patients of (c)R0T1; (d) T1; (e) T3; (f) R0; (g) R3; (h) R0 with eNs < 5.5 . (R0: LNR=0, R1: $0 < LNR < 0.24$, R2: $0.24 \leq LNR < 0.45$ and R3: $LNR \geq 0.45$; T1: ≤ 2 cm, T2: > 2 cm and ≤ 4 cm, T3: > 4 cm; LNR: positive lymph nodes ratio; eNs: examined lymph nodes).

Previous study showed that 15 eNs were needed for accuracy of staging, decreasing the false negative rate of nodal status [12]. A study for gastric cancer incorporated node-negative patients with insufficient eNs into node-positive group, which implies that insufficient eNs may result in poor prognosis and false-negative nodal status [13]. We focus on the number of eNs needed for good prognosis, finding

that 6 eNs are needed for good prognosis of node-negative patients, suggesting that at least 6 eNs are expected after resection.

Characteristics of PDACs in different sites of pancreas are different [21]. Our study implied that for patients with small tumor and node-negative patients with insufficient eNs, tumors in the head of pancreas led to poor prognosis, possibly because of their

tendency to lymph nodes metastasis. However, the underlying molecule mechanisms are needed to be explored.

When diameters of PDAC are less than 1cm, few positive nodes are found. Local tumor growth is a vital step for metastasis [22]. As tumor grows, proliferating tumor cells, insufficient blood supply and poor lymph drainage result in activation of VEGF-C/Flt-4 axis and hypoxia signaling pathway, promoting lymphangiogenesis and making cancer cell more aggressive, which accelerates metastasis of cancer [23-25]. That corresponds with our finding.

RPA, dichotomizing populations into subsets step by step repeatedly [26] by suitable parameters and cut-off value to maximize the change [27, 28], is widely used in tumor staging [19, 29-31], while nomograms combine prognostic factors to predict clinical outcome [32, 33]. However, cumbersome calculation limits nomograms' uses in clinical practice [34]. Our RPA model has the following advantages. Firstly, without cumbersome calculation of scores which limits nomograms' uses, it is easy to use our model to evaluate prognosis of PDAC patients in clinical practices, which is helpful for decision-making of clinicians. Secondly, RPA, a kind of decision tree model [35], can handle lots of variables at a time and show importance ranking of variables [36]. The more important a variable is, the closer it is to the "root" of the decision tree. For example, in our model, LNR is the closest to the "root" of tree in node-positive patients (Figure S1), which means that compared to other parameters included, LNR is the most important prognostic variable, corresponding with our multivariable analysis. Thirdly, our RPA model can help to understand interactions among included parameters. For example, for R0T1 patients in our model, location of tumor is an important prognostic factor. However, for patients with high LNR, tumor site is not taken into consideration, which implies that tumor site may be more important for PDAC of early stage. Last but not the least, as both continuous variables and nominal variables can be handled by RPA. The best cut-off value for division of patients into different prognosis can be found for continuous variables. In our study, 6 eNs was found by RPA as cut-off value for eNs to improve prognosis of node-negative patients.

When assessing prognostic homogeneity in internal validation set, stage-IA and IIA of 8th AJCC can't be stratified by our staging model (Table S5), as a result of insufficient number of patients. Prognostic effects of eNs are based on the hypothesis that insufficient eNs may result in high rate of false negative of nodal status, meaning that eNs show

predictive power when samples size is large enough.

There several limitations. First, coding errors are reported to remain in SEER database [37]. Besides, additional therapies, such as adjuvant chemotherapy, are not considered because of information unavailable. Thirdly, the number of patients in SYSUCC cohort may be not enough compared to patients from SEER database.

In short, we develop a refined staging scheme incorporating tumor size, tumor site, eNs and LNR, illustrating their interactions.

Supplementary Material

Supplementary figure and tables.

<http://www.jcancer.org/v09p3507s1.pdf>

Acknowledgements

This study is supported by National Natural Science Funds (nos. 81672390)

Competing Interests

The authors have declared that no competing interest exists.

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015; 65: 87-108.
2. Kamisawa T, Wood LD, Itoi T, Takaori K. Pancreatic cancer. *Lancet.* 2016; 388: 73-85.
3. Allen PJ, Kuk D, Castillo CF, Basturk O, Wolfgang CL, Cameron JL, et al. Multi-institutional Validation Study of the American Joint Commission on Cancer (8th Edition) Changes for T and N Staging in Patients With Pancreatic Adenocarcinoma. *Ann Surg.* 2017; 265: 185-91.
4. Kamarajah SK, Burns WR, Frankel TL, Cho CS, Nathan H. Validation of the American Joint Commission on Cancer (AJCC) 8th Edition Staging System for Patients with Pancreatic Adenocarcinoma: A Surveillance, Epidemiology and End Results (SEER) Analysis. *Ann Surg Oncol.* 2017; 24: 2023-30.
5. Ling Q, Xu X, Zheng SS, Kalthoff H. The diversity between pancreatic head and body/tail cancers: clinical parameters and in vitro models. *Hepatobiliary & pancreatic diseases international: HBPD INT.* 2013; 12: 480-7.
6. Lim JE, Chien MW, Earle CC. Prognostic factors following curative resection for pancreatic adenocarcinoma: a population-based, linked database analysis of 396 patients. *Ann Surg.* 2003; 237: 74-85.
7. Richter A, Niedergethmann M, Sturm JW, Lorenz D, Post S, Trede M. Long-term results of partial pancreaticoduodenectomy for ductal adenocarcinoma of the pancreatic head: 25-year experience. *World J Surg.* 2003; 27: 324-9.
8. Tarantino I, Warschkow R, Hackert T, Schmied BM, Buchler MW, Strobel O, et al. Staging of pancreatic cancer based on the number of positive lymph nodes. *Brit J Surg.* 2017; 104: 608-18.
9. Morales-Oyarvide V, Rubinson DA, Dunne RF, Kozak MM, Bui JL, Yuan C, et al. Lymph node metastases in resected pancreatic ductal adenocarcinoma: predictors of disease recurrence and survival. *Br J Cancer.* 2017; 117: 1874-82.
10. Strobel O, Hinz U, Gluth A, Hank T, Hackert T, Bergmann F, et al. Pancreatic adenocarcinoma: number of positive nodes allows to distinguish several N categories. *Ann Surg.* 2015; 261: 961-9.
11. Warschkow R, Widmann B, Beutner U, Marti L, Steffen T, Schiesser M, et al. The More the Better-Lower Rate of Stage Migration and Better Survival in Patients With Retrieval of 20 or More Regional Lymph Nodes in Pancreatic Cancer: A Population-Based Propensity Score Matched and Trend SEER Analysis. *Pancreas.* 2017; 46: 648-57.
12. Tomlinson JS, Jain S, Bentrem DJ, Sekeris EG, Maggard MA, Hines OJ, et al. Accuracy of staging node-negative pancreas cancer: a potential quality measure. *Arch Surg.* 2007; 142: 767-23; discussion 73-4.
13. Li BF, Li YF, Wang W, Qiu HB, Seeruttun SR, Fang C, et al. Incorporation of N0 Stage with Insufficient Numbers of Lymph Nodes into N1 Stage in the Seventh Edition of the TNM Classification Improves Prediction of Prognosis in Gastric Cancer: Results of a Single-Institution Study of 1258 Chinese Patients. *Annals of Surgical Oncology.* 2016; 23: 142-8.
14. Valsangkar NP, Bush DM, Michaelson JS, Ferrone CR, Wargo JA, Lillemoe KD, et al. N0/N1, PNL, or LNR? The effect of lymph node number on accurate

- survival prediction in pancreatic ductal adenocarcinoma. *J Gastrointest Surg.* 2013; 17: 257-66.
15. Slidell MB, Chang DC, Cameron JL, Wolfgang C, Herman JM, Schulick RD, et al. Impact of total lymph node count and lymph node ratio on staging and survival after pancreatectomy for pancreatic adenocarcinoma: a large, population-based analysis. *Ann Surg Oncol.* 2008; 15: 165-74.
 16. Huebner M, Kendrick M, Reid-Lombardo KM, Que F, Therneau T, Qin R, et al. Number of lymph nodes evaluated: prognostic value in pancreatic adenocarcinoma. *J Gastrointest Surg.* 2012; 16: 920-6.
 17. Koskas M, Luton D, Graesslin O, Barranger E, Clavel-Chapelon F, Haddad B, et al. Direct Comparison of Logistic Regression and Recursive Partitioning to Predict Lymph Node Metastasis in Endometrial Cancer. *Int J Gynecol Cancer.* 2015; 25: 1037-43.
 18. Su X, Azuero A, Cho J, Kvale E, Meneses KM, McNees MP. An introduction to tree-structured modeling with application to quality of life data. *Nurs Res.* 2011; 60: 247-55.
 19. Chen YT, Huang ZP, Zhou ZW, He MM. Equipping the American Joint Committee on Cancer staging for resectable pancreatic ductal adenocarcinoma with tumor grade: a recursive partitioning analysis. *Med Oncol.* 2016; 33: 122.
 20. Mayr A, Hofner B, Schmid M. Boosting the discriminatory power of sparse survival models via optimization of the concordance index and stability selection. *BMC Bioinformatics.* 2016; 17: 288.
 21. Lau MK, Davila JA, Shaib YH. Incidence and survival of pancreatic head and body and tail cancers: a population-based study in the United States. *Pancreas.* 2010; 39: 458-62.
 22. Coumans FA, Siesling S, Terstappen LW. Detection of cancer before distant metastasis. *BMC Cancer.* 2013; 13: 283.
 23. Ji RC. Hypoxia and lymphangiogenesis in tumor microenvironment and metastasis. *Cancer letters.* 2014; 346: 6-16.
 24. Su JL, Yang PC, Shih JY, Yang CY, Wei LH, Hsieh CY, et al. The VEGF-C/Flt-4 axis promotes invasion and metastasis of cancer cells. *Cancer Cell.* 2006; 9: 209-23.
 25. Chang Q, Jurisica I, Do T, Hedley DW. Hypoxia predicts aggressive growth and spontaneous metastasis formation from orthotopically grown primary xenografts of human pancreatic cancer. *Cancer Res.* 2011; 71: 3110-20.
 26. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods.* 2009; 14: 323-48.
 27. Cook EF, Goldman L. Empiric comparison of multivariate analytic techniques: advantages and disadvantages of recursive partitioning analysis. *J Chronic Dis.* 1984; 37: 721-31.
 28. Kattan MW, Hess KR, Beck JR. Experiments to determine whether recursive partitioning (CART) or an artificial neural network overcomes theoretical limitations of Cox proportional hazards regression. *Comput Biomed Res.* 1998; 31: 363-73.
 29. Jiang Y, Su Y, Chen Y, Li Z. Refining the American Joint Committee on Cancer Staging Scheme For Resectable Pancreatic Ductal Adenocarcinoma Using Recursive Partitioning Analysis. *Journal of Cancer.* 2017; 8: 2765-73.
 30. Huang SH, Xu W, Waldron J, Siu L, Shen X, Tong L, et al. Refining American Joint Committee on Cancer/Union for International Cancer Control TNM stage and prognostic groups for human papillomavirus-related oropharyngeal carcinomas. *J Clin Oncol.* 2015; 33: 836-45.
 31. Adam MA, Thomas S, Roman SA, Hyslop T, Sosa JA. Rethinking the Current American Joint Committee on Cancer TNM Staging System for Medullary Thyroid Cancer. *Jama Surgery.* 2017; 152: 869-76.
 32. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncology.* 2015; 16: E173-E80.
 33. Brennan MF, Kattan MW, Klimstra D, Conlon K. Prognostic nomogram for patients undergoing resection for adenocarcinoma of the pancreas. *Annals of Surgery.* 2004; 240: 293-8.
 34. Kawai K, Sunami E, Yamaguchi H, Ishihara S, Kazama S, Nozawa H, et al. Nomograms for colorectal cancer: A systematic review. *World J Gastroenterol.* 2015; 21: 11877-86.
 35. Zhang Z. Decision tree modeling using R. *Ann Transl Med.* 2016; 4: 275.
 36. Xu W, Che J, Kong Q. Recursive Partitioning Method on Competing Risk Outcomes. *Cancer Inform.* 2016; 15: 9-16.
 37. Shah SK, Fleet TM, Williams V, Smith AY, Skipper B, Wiggins C. SEER Coding Standards Result in Underestimation of Positive Surgical Margin Incidence at Radical Prostatectomy: Results of a Systematic Audit. *J Urology.* 2011; 186: 855-9.