Radiology

ORIGINAL RESEARCH • GASTROINTESTINAL IMAGING

Artificial Intelligence to Predict Lymph Node Metastasis at CT in Pancreatic Ductal Adenocarcinoma

Yun Bian, MD, PhD • Zhilin Zheng, BS • Xu Fang, MM • Hui Jiang, MD, PhD • Mengmeng Zhu, MM • Jieyu Yu, MM • Haiyan Zhao, MM • Ling Zhang, PhD • Jiawen Yao, PhD • Le Lu, PhD • Jianping Lu, MD, PhD • Chengwei Shao, MD, PhD**

From the Departments of Radiology (Y.B., X.F., M.Z., J. Yu, H.Z., J.L., C.S.) and Pathology (H.J.), Changhai Hospital, 168 Changhai Road, Shanghai 200433, China; Ping An Technology, Shanghai, China (Z.Z.); and PAII Inc, Bethesda, Md (L.Z., J. Yao, L.L.). Received February 11, 2022; revision requested April 15; revision received July 14; accepted July 22. **Address correspondence to** C.S. (email: *chengweishaoch@163.com*).

Supported in part by the National Science Foundation for Scientists of China (81871352, 82171915, and 82171930), Natural Science Foundation of Shanghai Science and Technology Innovation Action Plan (21ZR1478500, 21Y11910300), Clinical Research Plan of SHDC (SHDC2020CR4073), and 234 Platform Discipline Consolidation Foundation Project (2019YPT001, 2020YPT001).

* J.L. and C.S. are co-senior authors.

Conflicts of interest are listed at the end of this article.

See also the editorial by Chu and Fishman in this issue.

Radiology 2022; 000:1–10 • https://doi.org/10.1148/radiol.220329 • Content codes: GI CT AI

Background: Although deep learning has brought revolutionary changes in health care, reliance on manually selected cross-sectional images and segmentation remain methodological barriers.

Purpose: To develop and validate an automated preoperative artificial intelligence (AI) algorithm for tumor and lymph node (LN) segmentation with CT imaging for prediction of LN metastasis in patients with pancreatic ductal adenocarcinoma (PDAC).

Materials and Methods: In this retrospective study, patients with surgically resected, pathologically confirmed PDAC underwent multidetector CT from January 2015 to April 2020. Three models were developed, including an AI model, a clinical model, and a radiomics model. CT-determined LN metastasis was diagnosed by radiologists. Multivariable logistic regression analysis was conducted to develop the clinical and radiomics models. The performance of the models was determined on the basis of their discrimination and clinical utility. Kaplan-Meier curves, the log-rank test, or Cox regression were used for survival analysis.

Results: Overall, 734 patients (mean age, 62 years ± 9 [SD]; 453 men) were evaluated. All patients were split into training (n = 545) and validation (n = 189) sets. Patients who had LN metastasis (LN-positive group) accounted for 340 of 734 (46%) patients. In the training set, the AI model showed the highest performance (area under the receiver operating characteristic curve [AUC], 0.91) in the prediction of LN metastasis, whereas the radiologists and the clinical and radiomics models had AUCs of 0.58, 0.76, and 0.71, respectively. In the validation set, the AI model showed the highest performance (AUC, 0.92) in the prediction of LN metastasis, whereas the radiologists and the clinical and radiomics models had AUCs of 0.65, 0.77, and 0.68, respectively ($P < .001$). AI model–predicted positive LN metastasis was associated with worse survival (hazard ratio, 1.46; 95% CI: 1.13, 1.89; $P = .004$).

Conclusion: An artificial intelligence model outperformed radiologists and clinical and radiomics models for prediction of lymph node metastasis at CT in patients with pancreatic ductal adenocarcinoma.

Online supplemental material is available for this article.

Pancreatic ductal adenocarcinoma (PDAC) is predicted to become the second leading cause of cancer-related death by 2030, with positive peripancreatic lymph nodes (LNs) found in up to 80% of patients with PDAC (1).

Preoperative assessment of LN status is essential. For prognosis, LN metastasis is an independent predictor of PDAC survival (2), with a 5-year survival rate of 40% in patients with no pathologically confirmed lymph node metastasis (pN0) compared with 10% for those with pathologically confirmed increasing metastatic involvement (pN1 and pN2) (2,3). For early detection, especially for tumors less than or equal to 2 cm in size, it is critical to identify LN disease status and further stratify the patients with stage 1 PDAC (T1N0M0). Compared with standard surgical resection, extended resection, including dissection of LNs, might result in reduced quality of life with no survival benefit (4,5). To date, the National Comprehensive Cancer Network (NCCN) guidelines on the optimal extent of LN dissection have not been achieved. For neoadjuvant therapy, the optimal sequence of chemotherapy and surgery is still debatable in patients with resectable PDAC (6). NCCN guidelines suggest considering neoadjuvant therapy for large regional LNs (7).

However, preoperative prediction of LN metastasis is challenging. CT is recommended as the first-line imaging modality for evaluating LN metastasis in PDAC. However, a meta-analysis showed that using CT to assess extraregional LN metastasis in pancreatic and periampullary cancer yielded a pooled sensitivity of 25% and a positive predictive value of 28% (8). Recently, radiomics has shown some promise in predicting LN metastasis

Abbreviations

AI = artificial intelligence, AUC = area under the receiver operating characteristic curve, DSC = Dice similarity coefficient, LN = lymph node, LVSI = lymphovascular space invasion, PDAC = pancreatic ductal adenocarcinoma

Summary

An artificial intelligence model outperformed radiologists and clinical and radiomics models for prediction of lymph node metastasis at CT in patients with pancreatic ductal adenocarcinoma.

Key Results

- In a validation set of 189 patients with surgically resected, pathologically confirmed pancreatic ductal adenocarcinoma, an artificial intelligence (AI) model had better performance ($P < .001$) for identifying metastatic lymph nodes than a clinical model and a radiomics model (areas under the receiver operating characteristic curve [AUCs] of 0.92, 0.77, and 0.68, respectively).
- The AI model also outperformed radiologists (validation set AUC, 0.92 vs 0.65; $P < .001$) in the prediction of lymph node metastasis.

in patients with PDAC (9–13), but this technique is manual and limited by small sample sizes ($n < 250$), lack of reproducibility, and challenges to implementation in a real-world setting. Furthermore, all radiomics features have only been extracted from tumors. LNs, the true objects of the LN metastasis prediction task, have not yet been considered.

Recently, there has been growing interest in applying deep learning for LN metastasis prediction from cancer imaging data (14,15). Similar to the aforementioned radiomics studies, the reliance on manual tumor delineation and failure to account for LN features limit the clinical translation of these methods. One study developed an artificial intelligence (AI) model to automatically detect and predict LN metastasis (16), but the ground-truth label of LN metastasis was diagnosed from radiologic images, not pathologic results.

Thus, the aim of this study was to develop and validate a fully automated preoperative deep learning–based AI tool for tumor and LN segmentation with CT imaging and to predict LN metastasis in patients with PDAC.

Materials and Methods

Patients

This retrospective single-center study was reviewed and approved by the Biomedical Research Ethics Committee of our institution (no. CHEC2021164) and was performed in accordance with the ethical standards of the 1964 Declaration of Helsinki. The requirement for patient informed consent was waived by the institutional review board due to the retrospective nature of the study and because all procedures performed were part of routine care. This article fundamentally follows the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis, or TRIPOD, guidelines (17).

Data were obtained from consecutive patients treated for PDAC at a university teaching hospital from January 2015 to April 2020. A total of 1119 patients with pathologically confirmed PDAC were included. The inclusion and exclusion criteria are presented in Figure 1 and Appendix E1 (online).

CT Protocol

Multiphasic CT was performed according to a pancreas-specific protocol with a 320-section, multidetector-row CT scanner (Aquilion ONE; Canon Medical Systems). The CT scan parameters were as follows: 120 kV; effective mAs, 150; beam collimation, 160 × 0.5 mm; matrix, 350 × 350; gantry rotation time, 0.5 second; volume CT dose index, 15.4 mGy; dose-length product, 308 mGy · cm. An initial cross-sectional nonenhanced CT scan was obtained, followed by a dynamic contrast-enhanced CT scan. The scan delay time was determined based on the test bolus. The contrast agent (90–95 mL of 355 mgI/mL iopromide [Ultravist 370; Bayer Schering Pharma]) was injected at a rate of 5.5 mL/sec with a power injector (Medrad Mark V Plus; Bayer) via the forearm vein, followed by an injection of 98 mL normal saline to flush the tube. After contrast agent injection, contrast-enhanced CT was performed in the arterial (20–25 seconds), portal venous (60–70 seconds), and delayed (110–130 seconds) phases. The section thickness was 0.8 mm and the intervals of the scan were 1.0 mm. The scanning range extended from the level of the diaphragm to the pelvis.

Pathologic Image Analysis

Pathologic data were obtained from medical records. T and N categories were reevaluated by a specialized pathologist. All pathologic results were recorded for the following factors: *(a)* tumor size and invasion (T) and LN involvement (N) evaluated according to the American Joint Committee on Cancer TNM staging manual, eighth edition (18), *(b)* differentiation grade, *(c)* lymphovascular space invasion (LVSI), and *(d)* perineural invasion.

Radiologic Image Analysis

We used the original cross-sectional images of the arterial and portal phases for radiologic analysis. All images were analyzed by two abdominal radiologists (Y.B. and X.F., with 30 and 10 years of experience, respectively) who were blinded to the clinical and pathologic details. In case of disagreement, the senior radiologist (C.S.) decided the results.

All tumors were evaluated for the following characteristics: *(a)* CT-determined tumor size (ie, maximum cross-sectional diameter of the tumor) (19), *(b)* tumor location (pancreatic head, neck, body, and tail), and *(c)* CT-determined LN metastasis, which was considered if one of six criteria was met, including LN short-axis diameter greater than 10 mm, nonuniform density, nonuniform enhancement, internal necrosis, LN fusion, ill-defined borders, or involvement of surrounding organs or blood vessels (18).

LN Metastasis Prediction AI Model

The LN and tumor annotation and automatic segmentation protocols are detailed in Appendixes E2 and E3 (online) and Figures E1–E3 (online). The interobserver and intraobserver median Dice similarity coefficients (DSCs) of manually delineated LNs were reported. The median DSCs between automatic segmentation and manually delineated tumor and LNs were also reported. After segmenting LN instances from the whole CT

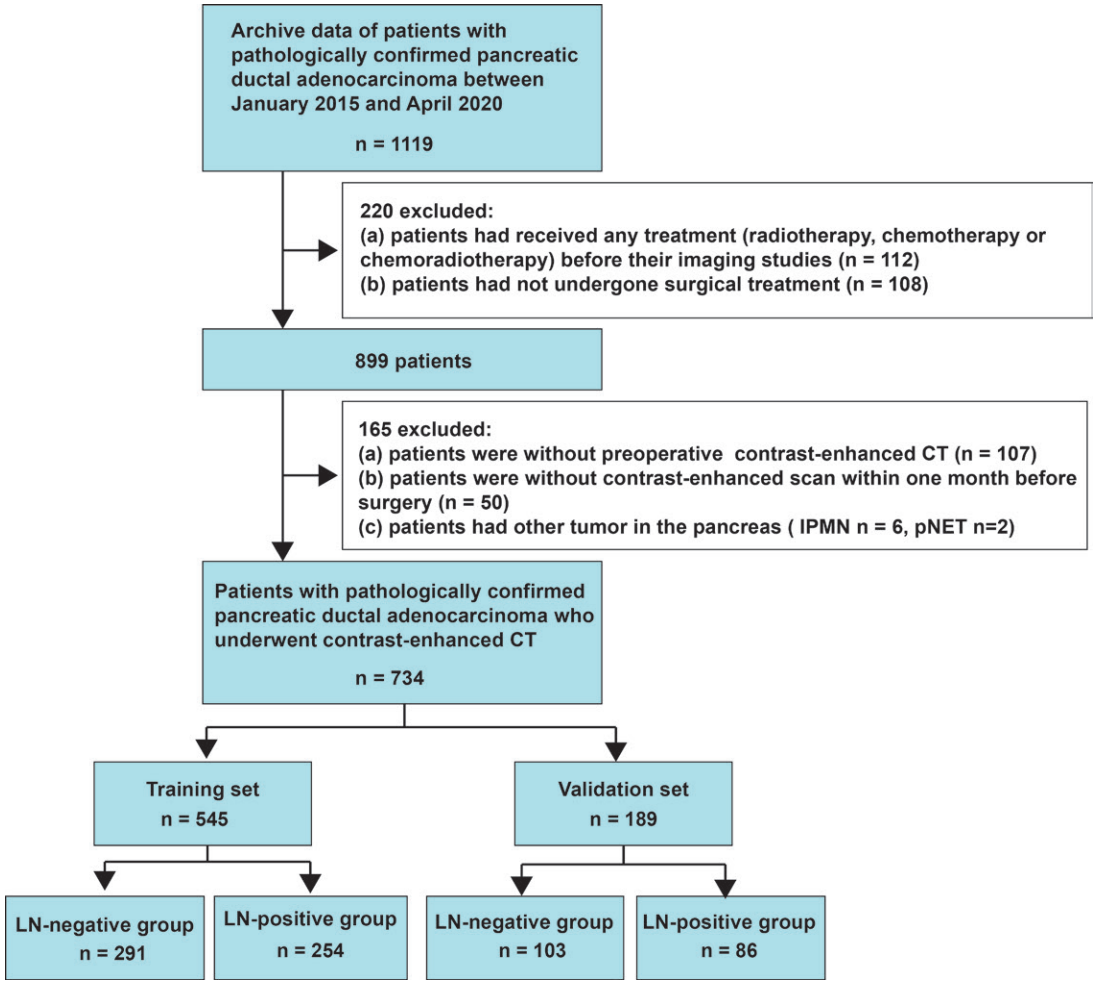

Figure 1: Patient flowchart for this study. IPMN = intraductal papillary mucinous neoplasm, LN = lymph node, pNET = pancreatic neuroendocrine tumor.

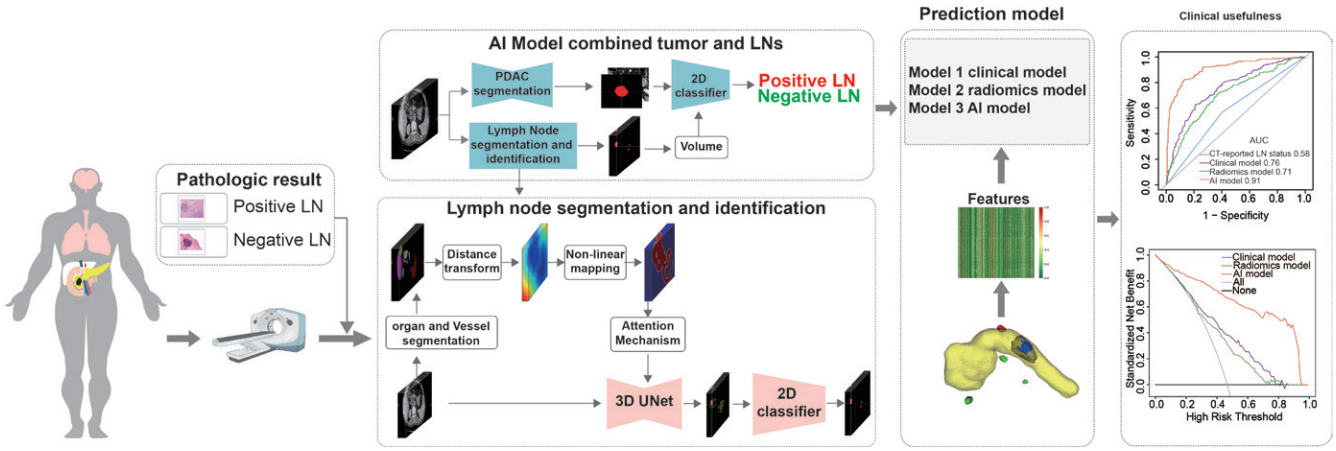

Figure 2: Artificial intelligence (AI) workflow and study flowchart. AUC = area under the receiver operating characteristic curve, LN = lymph node, PDAC = pancreatic ductal adenocarcinoma, 2D = two-dimensional, 3D = three-dimensional.

image, we identified them as having positive metastasis (LG-positive) or negative metastasis (LG-negative) (bottom branch in the upper panel of Fig 2). The network was trained using cropped subvolumes centered at the annotated LN instances and used to identify automatically segmented LNs as positive or negative at the inference time. To aggregate the results of LN instance-level segmentation and identification into patient-level LN metastasis status prediction, we conducted receiver operating characteristic analysis. The volume of the largest predicted positive LN in a patient was found to have the best results on the tuning set.

To further boost performance, we incorporated the imaging characteristics of the primary tumor into the model (top branch in the upper panel of Fig 2). A tumor and LN–combined, patient-level, LN metastasis status prediction model was built based on the automated segmentation results of the LNs and the annotated tumors. Note that the inference process was fully automated, in which the LN metastasis prediction was derived using automated segmentation of both the tumor and LNs. Details of the AI model are presented in Appendix E4 (online).

Comparison with Other Prediction Models

We developed two different integrated prediction models for comparison: a clinical model, combining pathologic and radiologic factors (T category, LVSI, perineural invasion, and CT-determined LN status), and a radiomics model, combining the arterial radiomics score and CT-determined LN status.

Radiomics Workflow

Our radiomics workflow included *(a)* automatic tumor segmentation, *(b)* feature extraction, and *(c)* feature reduction and selection. The radiomics approach is detailed in Appendix E5 (online).

Statistical Analysis

The prediction models, including the clinical and radiomics models, were built from the training set using the following steps. First, we examined the differences in all variables between the LN-positive and LN-negative groups. Second, univariable regression analysis and multivariable logistic regression analysis were conducted to develop a clinical and radiomics model for the prediction of LN metastasis, and a nomogram was constructed. We constructed a multivariable model using a stepwise regression method based on the Akaike information criterion to determine the best-fitting model (20). The differentiation performance of established models was quantified by plotting a receiver operating characteristic curve, and the area under the receiver operating characteristic curve (AUC) was calculated. Receiver operating characteristic curves of these models were compared using the DeLong test (21). Third, each model's clinical usefulness was tested using decision curve analysis by quantifying the net benefit at different threshold probabilities. Finally, overall survival times were calculated from the date of surgery to death from any cause or the last follow-up, where death was set as an event and those patients lost to follow-up were censored. Kaplan-Meier estimates were applied to graph survival curves, and the log-rank test was performed to analyze the differences between the curves. Univariable and multivariable Cox regression analyses were conducted to analyze the relationship between AI-predicted LN metastasis status and overall survival. A two-tailed P value less than .05 was considered indicative of a statistically significant difference. All analyses were performed using R version 3.3.3 (The R Foundation).

The major components of our code are available in open-source repositories or libraries, including PyTorch *(https://pytorch.org/)*, nnUNet *(https://github.com/MIC-DKFZ/nnUNet)*, and SimpleITK *(https://simpleitk.org/)* for the Euclidean distance transform. The prediction AI model, inference code, and illustrative examples of CT images, tumors, and LN masks are publicly available on GitHub *(https://github.com/DeepMedImg/DeepCT-LNM-Example)*.

Results

Patient Characteristics

Of a total of 1119 patients with pathologically confirmed PDAC, 734 consecutive patients (mean age, 62 years ± 9 [SD]; 453 men) with PDAC were included after applying the inclusion and exclusion criteria. The training set comprised 545 consecutive patients from January 2015 to April 2019, and the independent validation set comprised 189 consecutive patients from May 2019 to April 2020. Of the 734 study patients, there were 394 (54%) in the LN-negative group and 340 (46%) in the LN-positive group. A total of 466 positive LNs were included. There was a difference in T categories, LVSI, perineural invasion, and CT-determined LN status between the LN-positive and LN-negative groups in both the training and validation sets ($P < .05$). CT-determined tumor size differed between groups only in the training set ($P < .05$). Age differed between groups only in the validation set ($P < .05$). Patient characteristics are shown in Table 1. Interobserver agreement analysis is detailed in Appendix E6 (online).

Univariable Analysis

The results of univariable analysis are presented in Table E1 (online). T category, LVSI, perineural invasion, and CT-determined LN status were associated with a higher risk of LN metastasis in both the training and validation sets ($P < .05$). CT-determined size ($P = .02$) in the training set and age ($P = .03$) in the validation set were associated with a higher risk of LN metastasis.

Manual Delineation Reproducibility and Automated Segmentation Accuracy

The median interobserver and intraobserver DSCs of the LNs were 0.24 (IQR, 0.13–0.41) and 0.49 (IQR, 0.36–0.60), respectively. The median DSCs of tumor and LN segmentation in the validation set were 0.68 (IQR, 0.59–0.76) and 0.59 (IQR, 0.36–0.70), respectively. The automated model detected the tumor in 189 of 189 (100%) patients in the validation set. For LN metastasis detection, when considering a DSC greater than or equal to 0.5 as a true-positive indicator of LN detection, the tru-positive rate and positive predictive value were 80.1% and 40.5%, respectively, with four false-positive results per volume. DSC values for segmentation of organs and vessels are detailed in Appendix E3 (online).

LN Metastasis Prediction Accuracy

The AI model showed the highest discrimination between the LN-negative and LN-positive groups, with an AUC of 0.91 (95% CI: 0.89, 0.94) in the training set and 0.92 (95% CI: 0.88, 0.96) in the validation set (Table 2). The sensitivity, specificity, accuracy, positive predictive value, and negative predictive value for the training set were 82.7%, 86.9%, 85.0%, 84.7%, and 85.2%, respectively, whereas those for the validation set were 80.2%, 89.3%, 85.2%, 86.3%, and 84.4%, respectively. Two illustrative cases diagnosed using the AI model are shown in Figures 3 and 4.

Table 1: Baseline Characteristics of Patients with Pancreatic Ductal Adenocarcinoma

Characteristic	Training Set			Validation Set		
	LN-Negative Group (n = 291)	LN-Positive Group (n = 254)	P Value	LN-Negative Group (n = 103)	LN-Positive Group (n = 86)	P Value
Age (y)*	62 ± 9	62 ± 9	.98	63 ± 10	59 ± 10	.03
Body mass index (kg/m^2)*	23 ± 3	24 ± 7	.07	23 ± 3	23 ± 2	.39
Sex			.12			.83
M	170 (58)	165 (65)		65 (63)	53 (62)	
F	121 (42)	89 (35)		38 (37)	33 (38)	
T category			<.001			.02
T1	55 (19)	29 (11)		16 (15)	6 (7)	
T2	100 (34)	128 (50)		46 (45)	55 (64)	
T3	135 (46)	91 (36)		40 (39)	22 (26)	
T4	1 (1)	6 (3)		1 (1)	3 (3)	
Differentiation grade			.74			.63
1	4 (1)	3 (1)		2 (2)	1 (1)	
2	212 (73)	178 (70)		77 (75)	60 (70)	
3–4	75 (26)	73 (29)		24 (23)	25 (29)	
Lymphovascular space invasion			<.001			<.001
Negative	242 (83)	133 (52)		87 (84)	47 (55)	
Positive	49 (17)	121 (48)		16 (16)	39 (45)	
Perineural invasion			<.001			.01
Negative	60 (21)	12 (5)		22 (21)	7 (8)	
Positive	231 (79)	242 (95)		81 (79)	79 (92)	
Surgery			.67			.93
Pancreatoduodenectomy	169 (58)	157 (62)		59 (57)	51 (59)	
Distal pancreatectomy	109 (37)	86 (34)		41 (40)	32 (37)	
Total pancreatectomy	13 (5)	11 (4)		3 (3)	3 (3)	
CT-determined LN status			<.001			<.001
Negative	175 (60)	111 (44)		65 (63)	29 (34)	
Positive	116 (40)	143 (56)		38 (37)	57 (66)	
CT-determined tumor size (cm)*	2.74 ± 1.33	3.02 ± 1.31	.002	2.76 ± 1.23	2.84 ± 1.06	.54
Pancreas location			.38			.78
Head	169 (58)	157 (62)		59 (57)	51 (59)	
Neck, body, and tail	122 (42)	97 (38)		44 (43)	35 (41)	

Note.—Except where indicated, data are numbers of patients, with percentages in parentheses. For the differentiation grade, 1 indicates well-differentiated, 2 indicates moderately differentiated, 3 indicates poorly differentiated, and 4 indicates undifferentiated tumor.
LN = lymph node.
* Data are means ± SDs.

Table 2: Performance of the Integrated Prediction Models

Performance	Clinical Model		Radiomics Model		AI Model	
	Training Set	Validation Set	Training Set	Validation Set	Training set	Validation Set
AUC*	0.76 (0.72, 0.80)	0.77 (0.70, 0.84)	0.71 (0.67, 0.75)	0.68 (0.60, 0.75)	0.91 (0.89, 0.94)	0.92 (0.88, 0.96)
Sensitivity	62.6 (159/254)	79.1 (68/86)	72.4 (184/254)	79.1 (68/86)	82.7 (210/254)	80.2 (69/86)
Specificity	79.0 (230/291)	68.0 (70/103)	62.2 (181/291)	49.5 (51/103)	86.9 (253/291)	89.3 (92/103)
Accuracy	71.4 (389/545)	73.0 (138/189)	67.0 (365/545)	63.0 (119/189)	85.0 (463/545)	85.2 (161/189)
PPV	72.3 (159/220)	67.3 (68/101)	62.6 (184/294)	56.7 (68/120)	84.7 (210/248)	86.3 (69/80)
NPV	70.8 (230/325)	79.6 (70/88)	72.1 (181/251)	73.9 (51/69)	85.2 (253/297)	84.4 (92/109)

Note.—Except where indicated, data are percentages, with numbers of patients in parentheses. Performance is presented as AUC, sensitivity, and specificity values according to the optimal selected cutoff. AI = artificial intelligence, AUC = area under the receiver operating characteristic curve, NPV = negative predictive value, PPV = positive predictive value.
* Data are AUCs, with 95% CIs in parentheses.

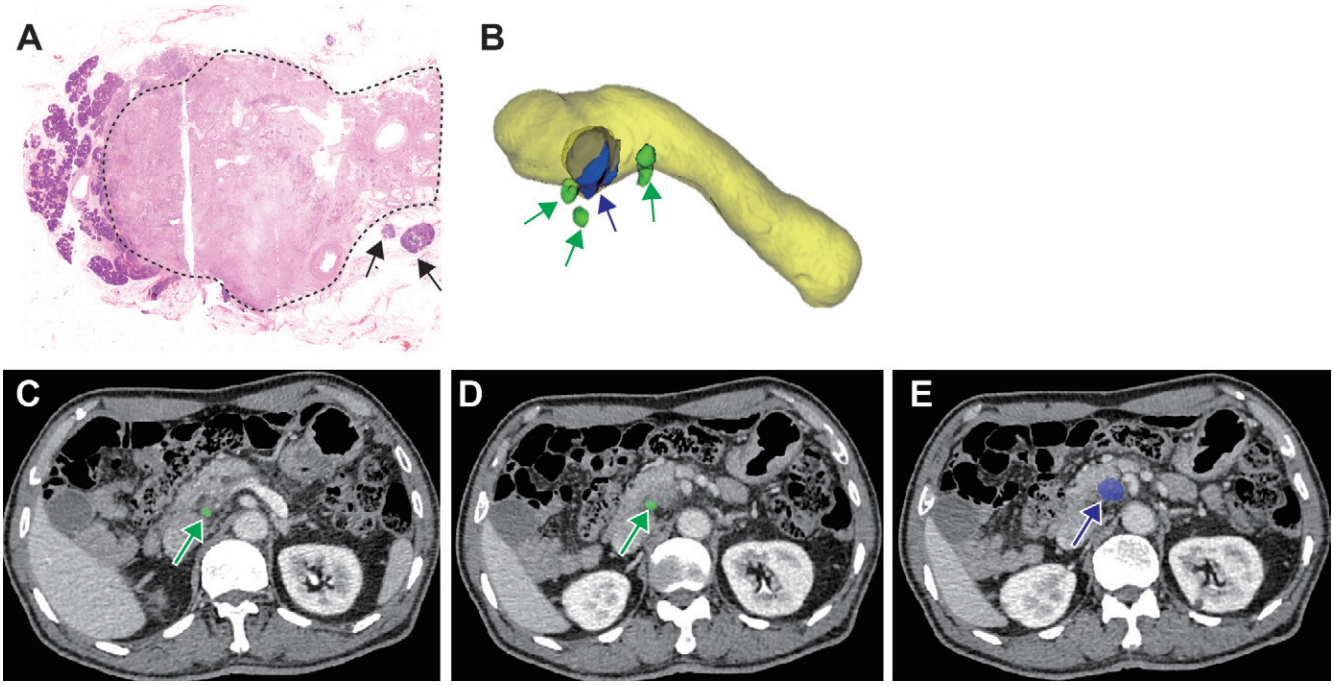

Figure 3: Images in a 71-year-old man with pancreatic ductal adenocarcinoma in the lymph node (LN)–negative group. **(A)** Hematoxylin-eosin stain shows that the peritumoral LNs (arrows) are negative (hematoxylin-eosin staining; magnification, ×1). The dashed line indicates the tumor periphery. **(B)** Graphic shows LNs (green arrows), tumor (blue arrow), and pancreas (yellow) segmented and diagnosed by artificial intelligence (AI). **(C)** Axial portal phase CT scan segmented by the AI model shows a negative LN (arrow) located at the pancreatic head. **(D)** Axial portal phase CT scan segmented by the AI model shows a negative LN (arrow) located at the pancreatic head. **(E)** Axial portal phase CT scan segmented by the AI model shows an infiltrative low-attenuation mass (arrow) at the pancreatic head.

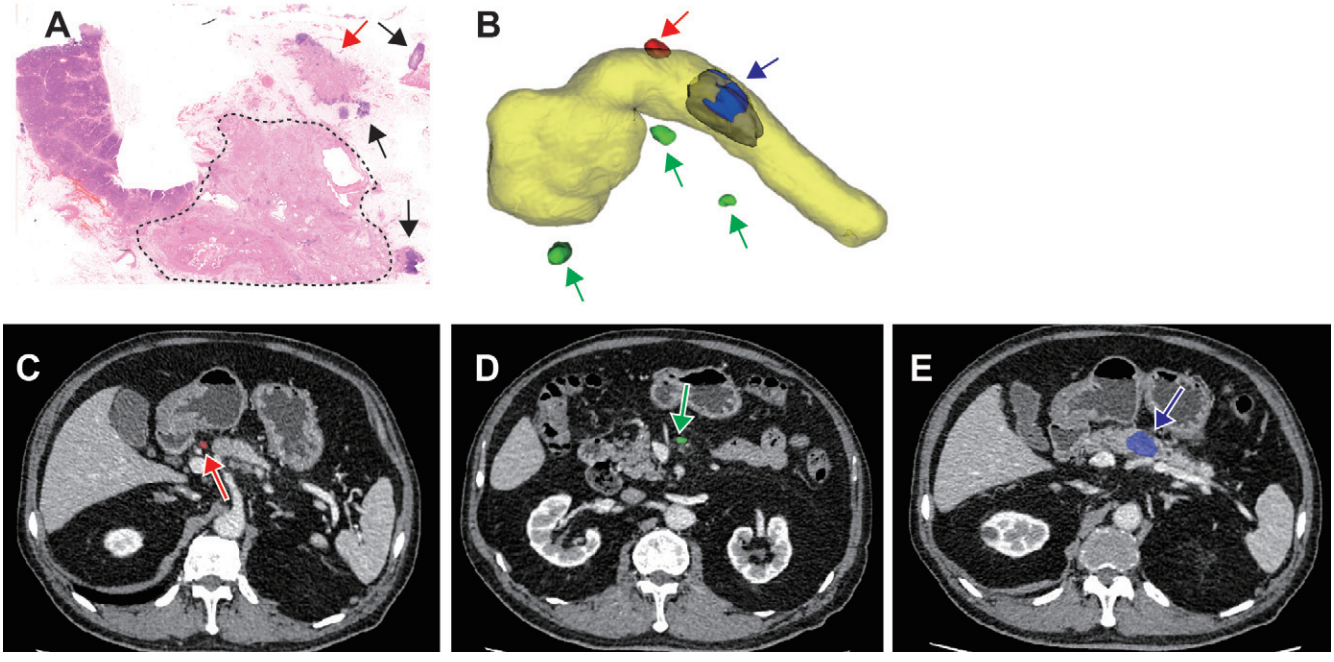

Figure 4: Images in a 78-year-old man with pancreatic ductal adenocarcinoma in the lymph node (LN)–positive group. **(A)** Hematoxylin-eosin stain shows a positive LN (red arrow) and three negative LNs (black arrows) located in the peritumoral area (hematoxylin-eosin staining; magnification, ×1). The dashed line indicates the tumor periphery. **(B)** Graphic shows three negative LNs (green arrows) and a positive LN (red arrow) located in the peritumoral area and tumor (blue arrow) located in the pancreatic body segmented and diagnosed by artificial intelligence (AI). **(C)** Axial portal phase CT scan segmented by the AI model shows a positive LN (arrow) located at the pancreatic neck. **(D)** Axial portal phase CT scan segmented by AI shows a negative LN (arrow) located at the pancreatic body. **(E)** Axial portal phase CT scan shows an infiltrative low-attenuation mass (arrow) at the pancreatic body.

For CT-determined LN status, the AUC, sensitivity, specificity, accuracy, positive predictive value, and negative predictive value for the training set were 0.60 (95% CI: 0.53, 0.63), 56.3%, 60.1%, 58.4%, 55.2%, and 61.2%, respectively, while those for the validation set were 0.65 (95% CI: 0.57, 0.73), 66.3%, 63.1%, 64.6%, 60.0%, and 69.2%, respectively.

The integrated clinical model identified T categories, LVSI, perineural invasion, CT-determined LN status, and CT-determined size as independent predictors (Table E2 [online]).

The radiomics features results are detailed in Appendix E5 (online) and Figure E4 (online). The integrated radiomics model identified the arterial radiomics score and CT-determined LN status as independent predictors (Table E3 [online]). Nomograms for the clinical and radiomics models are shown in Figure E5 (online).

There was a significant difference in AUCs between the AI model and the other integrated models in the validation set according to the DeLong test ($P < .001$). Receiver operating characteristic curves for CT-determined LN status according to radiologists and the clinical, radiomics, and AI models are shown in Figure 5A and 5B, and the performance of the three prediction models is outlined in Table 2.

Clinical Utility of the Prediction Models

Compared with scenarios in which no prediction model was used (ie, treat-all or treat-none scheme), the AI model provided a better net benefit for predicting LN metastasis than the clinical and radiomics models for threshold probabilities greater than 0.4% (Fig 5C).

Association between AI-predicted LN Metastasis Status and Overall Survival

In the training set, 253 of 291 patients (87%) were accurately identified in the LN-negative group and 210 of 254 patients (83%) were accurately identified in the LN-positive group using

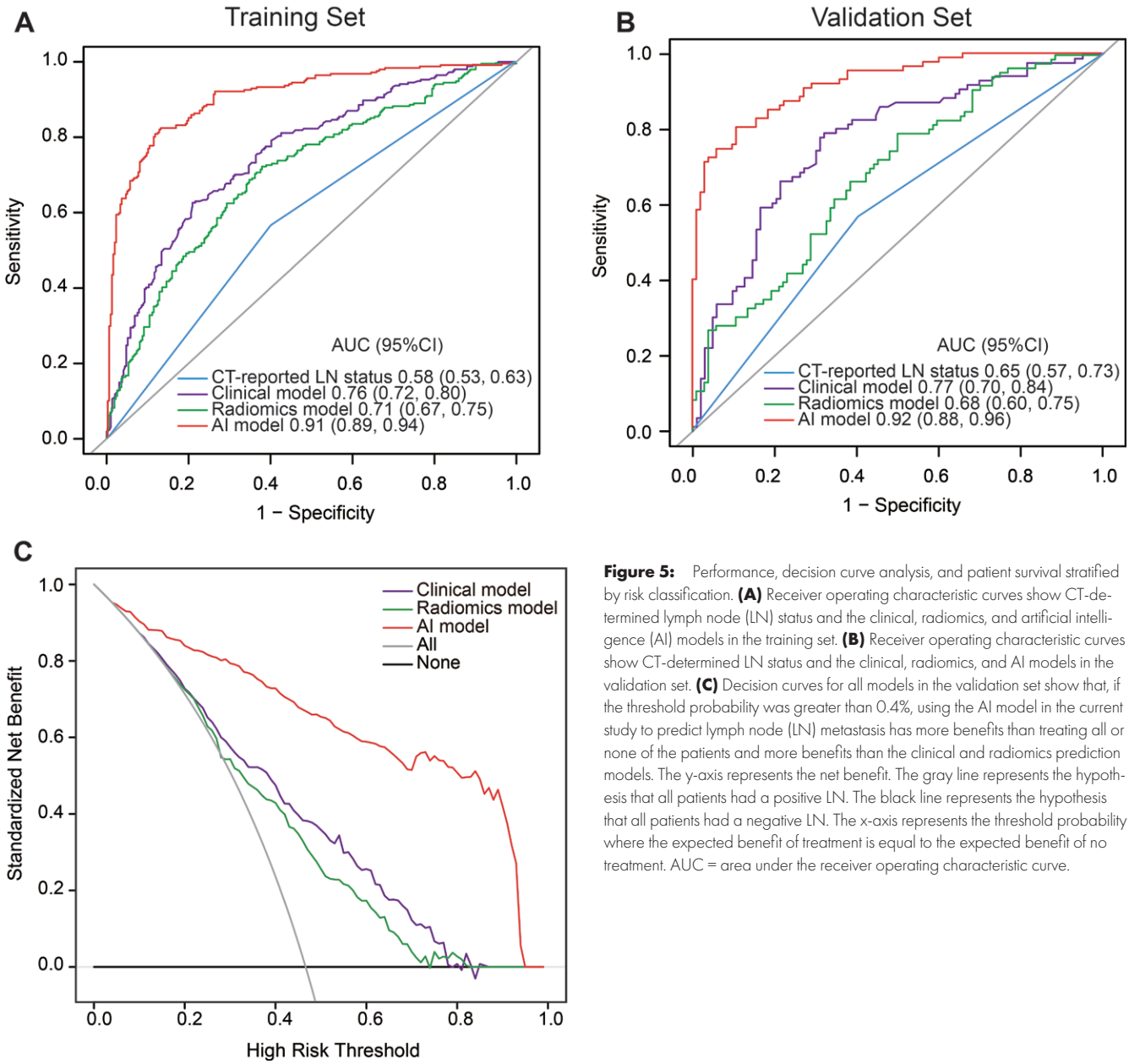

Figure 5: Performance, decision curve analysis, and patient survival stratified by risk classification. **(A)** Receiver operating characteristic curves show CT-determined lymph node (LN) status and the clinical, radiomics, and artificial intelligence (AI) models in the training set. **(B)** Receiver operating characteristic curves show CT-determined LN status and the clinical, radiomics, and AI models in the validation set. **(C)** Decision curves for all models in the validation set show that, if the threshold probability was greater than 0.4%, using the AI model in the current study to predict lymph node (LN) metastasis has more benefits than treating all or none of the patients and more benefits than the clinical and radiomics prediction models. The y-axis represents the net benefit. The gray line represents the hypothesis that all patients had a positive LN. The black line represents the hypothesis that all patients had a negative LN. The x-axis represents the threshold probability where the expected benefit of treatment is equal to the expected benefit of no treatment. AUC = area under the receiver operating characteristic curve.

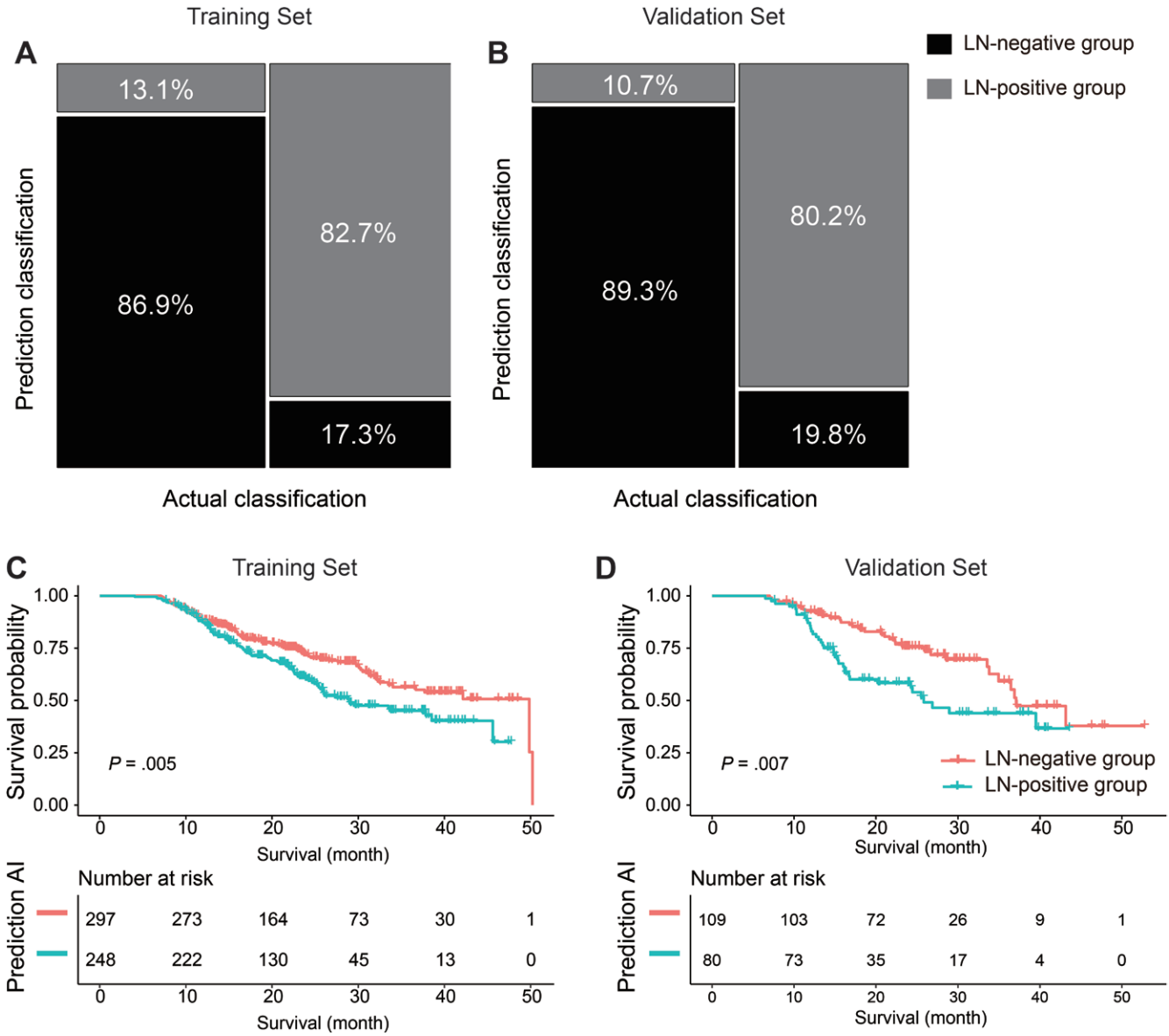

Figure 6: Classification and survival prediction of the artificial intelligence (AI) model. **(A)** Mosaic plot of the training set. **(B)** Mosaic plot of the validation set. **(C)** Kaplan-Meier survival curves for the prediction AI model show significantly longer survival for patients in the lymph node (LN)–negative group than those in the LN-positive group in the training set. **(D)** Kaplan-Meier survival curves for the prediction AI model show significantly longer survival for patients in the LN-negative group than those in the LN-positive group in the validation set.

the AI model (Fig 6A). In contrast, in the validation set, 92 of 103 patients (89%) were accurately predicted in the LN-negative group and 69 of 86 patients (80%) were accurately identified in the LN-positive group using the AI model. (Fig 6B).

The median follow-up duration was 23 months (IQR, 15–30 months) in the LN-negative group and 19 months (IQR, 13–26 months) in the LN-positive group, and 105 patients in the LN-negative group and 125 patients in the LN-positive group died. The log-rank test revealed a longer survival duration in the predicted LN-negative group than in the predicted LN-positive group (training set, $P = .005$; validation set, $P = .007$) (Fig 6C, 6D).

In the univariable analysis, positive AI-predicted LN metastasis was associated with worse survival (hazard ratio, 1.59; 95% CI: 1.25, 2.02; $P < .001$). In addition, multivariable Cox regression analysis revealed that AI-predicted LN metastasis was an independent preoperative predictor for worse overall survival (hazard ratio, 1.46; 95% CI: 1.13, 1.89; $P = .004$) (Table 3).

Discussion

Studies on the automatic segmentation and identification of both pancreatic lymph nodes (LNs) and tumor that also combine artificial intelligence (AI) for the automated imaging characteristics of both remain lacking in the literature. In this study, we developed and validated an AI model to predict LN metastasis in patients with pancreatic ductal adenocarcinoma. The AI model demonstrated favorable discrimination in both the primary set (area under the receiver operating characteristic curve [AUC], 0.91) and the validation set (AUC, 0.92) and outperformed radiologists, a clinical model, and a radiomics model.

Table 3: Clinical and Pathologic Characteristics and AI Model–predicted LN Metastasis Associated with Worse Overall Survival

	Univariable Analysis		Multivariable Analysis	
Variable	Hazard Ratio	*P* Value	Hazard Ratio	*P* Value
Sex				
M	1		1	
F	0.83 (0.65, 1.07)	.14	0.84 (0.65, 1.09)	.19
Age (per year)	1.01 (1.00, 1.02)	.22	1.01 (1.00, 1.02)	.08
Body mass index (per kg/m^2)	1.02 (1.01, 1.04)	.01	1.02 (1.01, 1.04)	.01
Pancreas location				
Head	1		1	
Neck, body, and tail	0.91 (0.71, 1.16)	.45	0.92 (0.71, 1.20)	.54
CT-determined tumor size (per cm)	0.99 (0.90, 1.08)	.75	1.06 (0.96,1.18)	.25
T category				
T1	1		1	
T2	1.61 (1.03, 2.51)	.04	1.48 (0.94, 2.32)	.09
T3–4	1.63 (1.04, 2.55)	.03	1.66 (1.04, 2.66)	.04
Differentiation grade				
1	1		1	
2	2.17 (0.54, 8.74)	.28	1.40 (0.34, 5.78)	.64
3–4	3.10 (0.76, 12.59)	.11	2.00 (0.48, 8.30)	.34
Lymphovascular space invasion				
Negative	1		1	
Positive	1.32 (1.03, 1.70)	.03	1.11 (0.85, 1.45)	.43
Perineural invasion				
Negative	1		1	
Positive	2.00 (1.27, 3.15)	.003	1.66 (1.04, 2.65)	.03
AI model–predicted LN metastasis				
Low risk	1		1	
High risk	1.59 (1.25, 2.02)	<.001	1.46 (1.13, 1.89)	.004

Note.—Data in parentheses are 95% CIs. There were 734 patients and 230 deaths. For the differentiation grade, 1 indicates well-differentiated, 2 indicates moderately differentiated, 3 indicates poorly differentiated, and 4 indicates undifferentiated tumor. AI = artificial intelligence, LN = lymph node.

In addition, the patients with LN metastasis identified by AI had approximately 50% worse survival (hazard ratio, 1.46; *P* = .004) after radical tumor resection, thereby providing important information for medical decision support.

Many studies have investigated risk stratification methods for LN metastasis, including T categories, LVSI, and perineural invasion (22–24). Yet preoperative identification of LN metastasis still mainly depends on the radiologist's report. However, the lower AUC values of 0.58–0.65 reported in the current study show that LN metastasis prediction by radiologists remains extremely difficult.

At present, few studies have assessed LN metastasis prediction using radiomics in patients with PDAC. Liu et al (10) developed a radiomics model with AUC values of 0.84 and 0.68 in the training and validation sets, respectively. The study by Li et al (11) reported a radiomics model with AUC values of 0.94 and 0.91 in the primary and validation sets, respectively. An et al (12) built a deep learning radiomics model with dual-energy CT and reported the best performance in predicting LN metastasis (AUC, 0.92). A similar radiomics model investigated by Gao et al (25) and Liang et al (26) also showed good performance. However, the primary limitation of these studies was the small number of patients with an over-optimized AUC. Another limitation of these studies was that the radiomics features in the prediction model were quite different from one another, which indicated challenges for reproducibility. The limitation persists in our studies. The performance of the radiomics model in the primary (AUC, 0.71) and validation (AUC, 0.68) sets was less desirable in the current study compared with our previous study (13). Thus, extracting more direct LN characteristics to develop the prediction model was more accurate in predicting LN metastasis in the current study.

Computer-aided LN segmentation has been reported in some solid tumors (27,28). In our study, the LN segmentation model improved a popular image segmentation model (ie, nnUNet) by introducing an attention mechanism guided by contextual organs and vessels, which explicitly used the spatial prior of pancreatic LN locations. Consequently, the network was encouraged to focus on regions surrounding certain organs and vessels and discard negative samples far from LN areas. In addition, automatic LN segmentation has the potential to decrease interobserver and intraobserver inconsistencies in delineating LNs and reduce the time and labor involved; in the current study, the median interobserver and intraobserver DSCs of LNs were only 0.24 and 0.49, respectively. These DSCs were lower than that (DSC, 0.59) between AI and manual segmentation, indicating that our automated segmentation model has similar reproducibility to a radiologist. This is probably the best an automated model can do in practice.

Our study had some limitations. First, although we included a relatively large number of patients with PDAC for LN metastasis prediction, this was a retrospective and single-center study with potential for bias. Second, we did not include carbohydrate antigen 19–9 because there are several issues with its clinical use, such as false-negative results in individuals with Lewis^{a-b-} phenotype or false-positive, occasional, and transient elevation in patients with benign diseases, making it a poor cancer-specific marker (29). Third, the median DSCs of PDAC tumor and LN segmentation were 0.68 and 0.59, indicating a modest agreement between the AI model and manual segmentation. Further work will test generalizability of the automated segmentation model to provide the agreement between AI and manual segmentation. Fourth, the

clinical model of this study, which included pathologic features, aided little in the preoperative prediction of LN metastasis. However, a routine pathology workflow for diagnosing LN metastasis is time-consuming and might result in misdiagnosis by a pathologist alone due to habituation. The clinical model was developed to solve these the problems. Fifth, there may have been difficulty in correlating LNs obtained at surgery with the specific nodes seen at preoperative scanning, representing a potential limitation. Finally, we did not evaluate the benefit of neoadjuvant therapy for patients with early PDAC and resectable PDAC.

In conclusion, our artificial intelligence (AI) model, a fully automated noninvasive tool combining both lymph node (LN) and tumor image characteristics, showed favorable accuracy for prediction of LN metastasis at CT in patients with pancreatic ductal adenocarcinoma (PDAC). The AI model outperformed radiologists, a clinical model, and a radiomics model. It could further provide clinically beneficial information for individualized diagnosis and treatment for patients with PDAC. To advance our study, we recommend prospective multicenter validation with a larger sample size to acquire high-level evidence for clinical applications. In addition, we aim to further investigate the accuracy of tumor and LN segmentation in the AI model. Finally, the benefit of resection for metastasis-positive LNs in PDAC needs evaluation.

Acknowledgments: We would like to acknowledge Jing Li, MD, PhD, Tiegong Wang, MD, PhD, and Fang Liu, MM, for technical support.

Author contributions: Guarantors of integrity of entire study, **Y.B., X.F., H.J., M.Z., J. Yu, H.Z., L.Z., L.L., J.L., C.S.**; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **Y.B., X.F., H.J., M.Z., L.L., J.L., C.S.**; clinical studies, **Y.B., H.J., L.L., C.S.**; experimental studies, **Y.B., Z.Z., H.J., J. Yu, H.Z., L.Z., J. Yao, L.L., J.L., C.S.**; statistical analysis, **Y.B., X.F., H.J., M.Z., J. Yu, H.Z., J. Yao, L.L., C.S.**; and manuscript editing, **Y.B., Z.Z., H.J., J. Yao, L.L., C.S.**

Data sharing: Data generated or analyzed during the study are available from the corresponding author by request.

Disclosures of conflicts of interest: **Y.B.** No relevant relationships. **Z.Z.** No relevant relationships. **X.F.** No relevant relationships. **H.J.** No relevant relationships. **M.Z.** No relevant relationships. **J. Yu** No relevant relationships. **H.Z.** No relevant relationships. **L.Z.** No relevant relationships. **J. Yao** No relevant relationships. **L.L.** No relevant relationships. **J.L.** No relevant relationships. **C.S.** No relevant relationships.

References

1. Katz MH, Hwang R, Fleming JB, Evans DB. Tumor-node-metastasis staging of pancreatic adenocarcinoma. CA Cancer J Clin 2008;58(2):111–125.
2. Lahat G, Lubezky N, Gerstenhaber F, et al. Number of evaluated lymph nodes and positive lymph nodes, lymph node ratio, and log odds evaluation in early-stage pancreatic ductal adenocarcinoma: numerology or valid indicators of patient outcome? World J Surg Oncol 2016;14(1):254.
3. Yamada M, Sugiura T, Okamura Y, et al. Clinical Implication of Node-negative Resectable Pancreatic Cancer. Ann Surg Oncol 2021;28(4):2257–2264.
4. Pedrazzoli S, DiCarlo V, Dionigi R, et al. Standard versus extended lymphadenectomy associated with pancreatoduodenectomy in the surgical treatment of adenocarcinoma of the head of the pancreas: a multicenter, prospective, randomized study. Lymphadenectomy Study Group. Ann Surg 1998;228(4):508–517.
5. Yeo CJ, Cameron JL, Lillemoe KD, et al. Pancreaticoduodenectomy with or without distal gastrectomy and extended retroperitoneal lymphadenectomy for periampullary adenocarcinoma, part 2: randomized controlled trial evaluating survival, morbidity, and mortality. Ann Surg 2002;236(3):355–366; discussion 366–368.
6. de Geus SW, Evans DB, Bliss LA, et al. Neoadjuvant therapy versus upfront surgical strategies in resectable pancreatic cancer: A Markov decision analysis. Eur J Surg Oncol 2016;42(10):1552–1560.
7. Tempero MA. NCCN Guidelines Updates: Pancreatic Cancer. J Natl Compr Canc Netw 2019;17(5.5):603–605.
8. Tseng DS, van Santvoort HC, Fegrachi S, et al. Diagnostic accuracy of CT in assessing extra-regional lymphadenopathy in pancreatic and periampullary cancer: a systematic review and meta-analysis. Surg Oncol 2014;23(4):229–235.
9. Bian Y, Guo S, Jiang H, et al. Relationship Between Radiomics and Risk of Lymph Node Metastasis in Pancreatic Ductal Adenocarcinoma. Pancreas 2019;48(9):1195–1203.
10. Liu P, Gu Q, Hu X, et al. Applying a radiomics-based strategy to preoperatively predict lymph node metastasis in the resectable pancreatic ductal adenocarcinoma. J XRay Sci Technol 2020;28(6):1113–1121.
11. Li K, Yao Q, Xiao J, et al. Contrast-enhanced CT radiomics for predicting lymph node metastasis in pancreatic ductal adenocarcinoma: a pilot study. Cancer Imaging 2020;20(1):12.
12. An C, Li D, Li S, et al. Deep learning radiomics of dual-energy computed tomography for predicting lymph node metastases of pancreatic ductal adenocarcinoma. Eur J Nucl Med Mol Imaging 2022;49(4):1187–1199.
13. Bian Y, Guo S, Jiang H, et al. Radiomics nomogram for the preoperative prediction of lymph node metastasis in pancreatic ductal adenocarcinoma. Cancer Imaging 2022;22(1):4.
14. Dong D, Fang MJ, Tang L, et al. Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study. Ann Oncol 2020;31(7):912–920.
15. Jin C, Jiang Y, Yu H, et al. Deep learning analysis of the primary tumour and the prediction of lymph node metastases in gastric cancer. Br J Surg 2021;108(5):542–549.
16. Lu Y, Yu Q, Gao Y, et al. Identification of Metastatic Lymph Nodes in MR Imaging with Faster Region-Based Convolutional Neural Networks. Cancer Res 2018;78(17):5135–5143.
17. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162(1):W1–W73.
18. Amin MB, Edge SB, Greene FL, et al, eds. AJCC Cancer Staging manual. 8th ed. New York, NY: Springer, 2017; 337–406. https://link.springer.com/book/9783319406176.
19. Watanabe H, Okada M, Kaji Y, et al. New response evaluation criteria in solid tumours-revised RECIST guideline (version 1.1) [in Japanese]. Gan To Kagaku Ryoho 2009;36(13):2495–2501.
20. Portet S. A primer on model selection using the Akaike Information Criterion. Infect Dis Model 2020;5:111–128.
21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845.
22. Yang YH, Liu JB, Gui Y, Lei LL, Zhang SJ. Relationship between autophagy and perineural invasion, clinicopathological features, and prognosis in pancreatic cancer. World J Gastroenterol 2017;23(40):7232–7241.
23. Chatterjee D, Rashid A, Wang H, et al. Tumor invasion of muscular vessels predicts poor prognosis in patients with pancreatic ductal adenocarcinoma who have received neoadjuvant therapy and pancreaticoduodenectomy. Am J Surg Pathol 2012;36(4):552–559.
24. Okano K, Asano E, Kushida Y, Kamada H, Mori H, Suzuki Y. Factors influencing lymph node metastasis in patients with ampullary adenocarcinoma. Dig Surg 2014;31(6):459–467.
25. Gao J, Han F, Jin Y, Wang X, Zhang J. A Radiomics Nomogram for the Preoperative Prediction of Lymph Node Metastasis in Pancreatic Ductal Adenocarcinoma. Front Oncol 2020;10:1654.
26. Liang X, Cai W, Liu X, Jin M, Ruan L, Yan S. A radiomics model that predicts lymph node status in pancreatic cancer to guide clinical decision making: A retrospective study. J Cancer 2021;12(20):6050–6057.
27. Alheejawi S, Xu H, Berendt R, Jha N, Mandal M. Novel lymph node segmentation and proliferation index measurement for skin melanoma biopsy images. Comput Med Imaging Graph 2019;73:19–29.
28. Li Z, Xia Y. Deep Reinforcement Learning for Weakly-Supervised Lymph Node Segmentation in CT Images. IEEE J Biomed Health Inform 2021;25(3):774–783.
29. Scarà S, Bottoni P, Scatena R. CA 19-9: Biochemical and Clinical Aspects. Adv Exp Med Biol 2015;867:247–260.