

LETTER TO EDITOR

Artificial intelligence–based 5-year survival prediction and prognosis of DNp73 expression in rectal cancer patients

Dear Editor,

Preoperative radiotherapy (pRT) is known to improve local control for rectal cancer patients besides surgery.^{1–3} However, there are many patients who do not respond to pRT but experience side effects. It is therefore urgently required to find promising pRT-related biomarkers for approaching precision medicine.

In this study, we investigated the application of artificial intelligence (AI) for discovering the predictive and prognostic power of the DNp73 expression in a cohort of 143 rectal cancer patients from the Swedish rectal cancer trial of pRT.² The DNp73 expression was identified by immunohistochemistry (IHC), and the procedure for the IHC image extraction was described in Ref. 4. While the manual pathology-based analysis of DNp73 expression did not provide any survival information ($P > .05$), the average AI-based validation results show very high accuracy rates ($\geq 93\%$) for the 5-year prediction and prognosis of the rectal cancer patients either with or without pRT.

The DNp73 expression was investigated in 96 biopsies, surgically resected normal and tumor samples from 77 patients without pRT and 59 patients with pRT (Figure 1A,B). The DNp73 staining was performed in the whole group of surgically resected distant normal ($n = 119$), adjacent normal ($n = 79$), and tumor samples ($n = 136$). Strong cytoplasmic DNp73 staining was present in the normal and tumor cells (Figure 1A,B). In the analysis of the clinicopathologic and biologic significance of DNp73 expression, we divided the patients into DNp73 weak and strong groups. The expression of DNp73 was significantly increased in the tumors either without or with pRT, when compared with the normal mucosa (Figure 1C, $P < .001$). The significant differences of the DNp73 expression were observed in the matched cases of the distant normal mucosa, adjacent normal mucosa, and tumor derived from the same patient (Figure 1D, $P = .002$). We found that the DNp73 expression in the biopsies was not related to any clinicopathologic variables including gender, age, differentiation, surgical type, local recurrence, distant recurrence,

and survival status (Table S1 in the Supporting Information, $P > .05$), while the DNp73 expression was related to local recurrence (Table S2 in the Supporting Information, $P = .042$) in the surgically resected tumor samples with pRT and surgical type (Table S2, $P = .021$) in the surgically resected tumor samples without pRT.

Because AI is considered as the foremost advanced approach in cancer research,^{5–10} we then used AI methods for exploring the DNp73 expression with respect to 5-year survival prediction and prognosis. The methods consist of 10 pretrained convolutional neural networks (CNNs) whose properties are listed in Table S3 in the Supporting Information. The data processing and network configuration are described as follows. Each whole IHC image was resized to match the input image size specified by each of the 10 networks (see the last column of Table S3). In performing the transfer learning, parameters of the networks were set as stochastic gradient descent with momentum = 0.9, minimum batch size = 10, maximum number of epochs = 6, initial learning rate = 0.0003, data were shuffled before every training epoch, learning rate drop factor = 0.1, learning rate drop period = 10, factor for the L_2 regularizer = 0.0001, and the method used for gradient thresholding = L_2 norm. The training and testing of the datasets for biopsies and surgically resected tumors without or with pRT were carried out by randomly selecting 90% of each dataset for training the CNN models and the remaining 10% for validation. Both training and validation of the 10 CNNs were repeated 10 times.

Average results and standard deviations for the accuracy, >5 -year (defined as true positive rate), and ≤ 5 -year (defined as true negative rate) prediction and prognosis (see Table S4 for definitions in the Supporting Information) obtained from selected top-performance CNNs, whose average accuracy $\geq 93\%$, are shown in Table 1. As a case for using DenseNet201 for biopsies without pRT, the average prediction for >5 years = 90%, ≤ 5 years = 98%, with average accuracy = 96%; and for biopsies with pRT, the average prediction for >5 years = 97%, ≤ 5 years = 80%,

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics

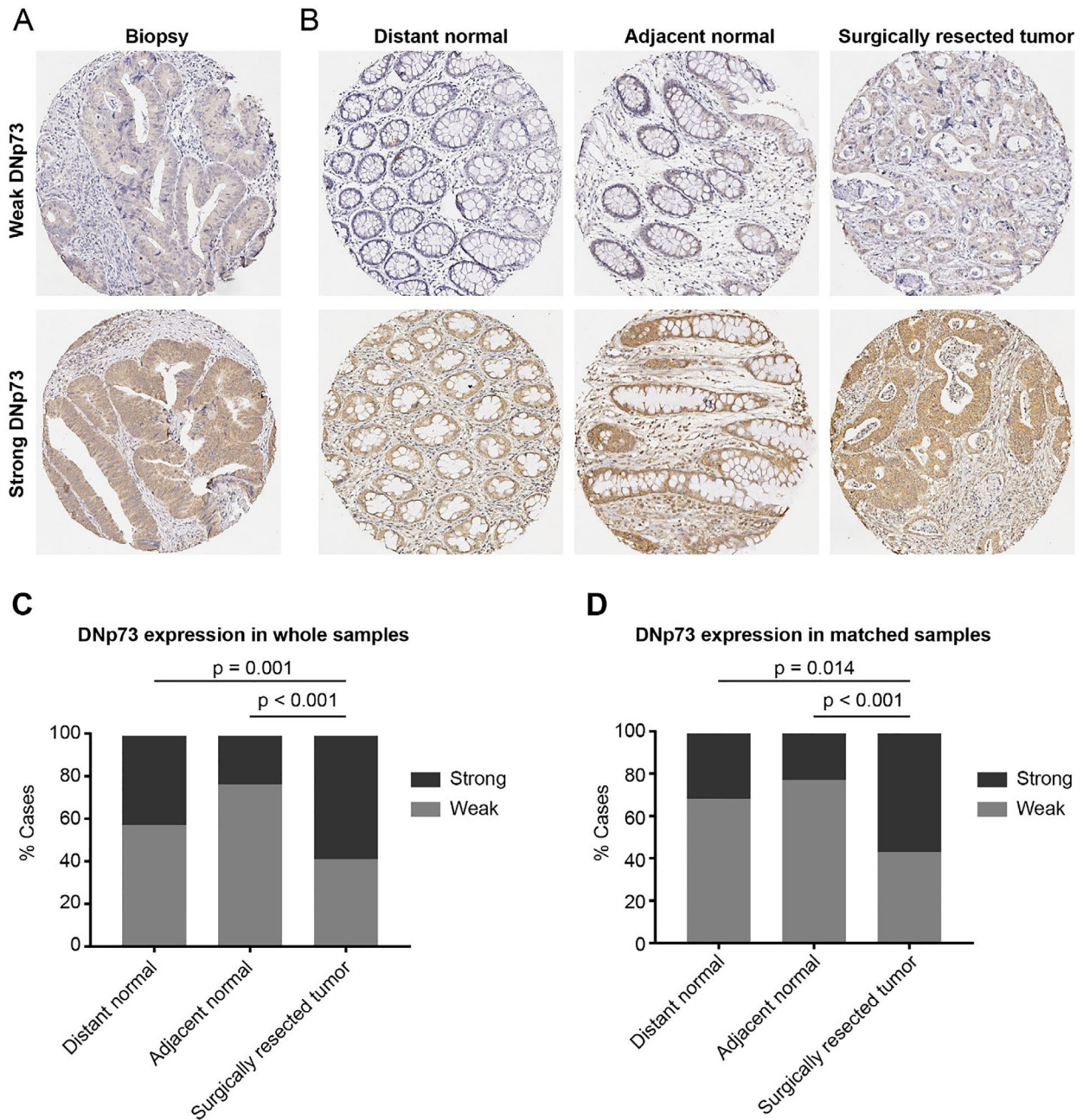


FIGURE 1 DNP73 expression by IHC staining in tumor samples from rectal cancer patients. A representative IHC image of DNP73 expression in biopsies (A) and surgically resected samples, including distant normal mucosa, adjacent normal mucosa, and surgically resected tumor (B); DNP73 expression in distant normal mucosa, adjacent normal mucosa, and surgical tumor obtained from whole samples (C), and matched samples (D). Whole samples indicated all surgically resected samples. Matched samples included surgically resected samples (including distant normal, adjacent normal, and primary tumor samples) from the same patient

with average accuracy = 93%. Using ResNet101 for surgically resected tumors without pRT, the average prediction for >5 years = 98%, ≤5 years = 90%, with average accuracy = 96%. Using DenseNet201 for the tumors with pRT, the average prediction for >5 years = 93%, ≤5 years = 95%, with average accuracy = 93%.

The results obtained from other CNNs for the prediction and prognosis using the biopsies and tumors are shown in

Table S1. Figures S1 and S2 (in the Supporting Information) show a training process and features learned by DenseNet-201 for classifying the biopsies without pRT, respectively. Using the maximum number of epochs = 6 for training, the accuracy could reach 100% (Figure S1).

These present results have a useful implication that DNP73 expression, by examining either biopsies or surgical tumors, can determine the prediction or prognosis of

TABLE 1 AI-based prediction and prognosis of DNP73 expression

CNN model	Accuracy (%)	>Five years (%)	≤Five years (%)
Biopsies without preoperative radiotherapy			
ResNet50	94.00 ± 9.66	70.00 ± 48.30	100.00 ± 0.00
VGG16	94.00 ± 9.66	80.00 ± 42.16	97.50 ± 7.91
DenseNet201	96.00 ± 8.43	90.00 ± 31.62	97.50 ± 7.91
Biopsies with preoperative radiotherapy			
ResNet101	92.50 ± 12.08	100.00 ± 0.00	70.00 ± 48.30
DenseNet201	92.50 ± 16.87	96.67 ± 10.54	80.00 ± 42.16
Tumors without preoperative radiotherapy			
GoogleNet	94.29 ± 18.07	96.00 ± 12.65	90.00 ± 31.63
ResNet50	94.29 ± 12.05	96.00 ± 8.43	90.00 ± 21.08
DenseNet201	94.29 ± 13.80	96.00 ± 8.43	90.00 ± 31.62
InceptionV3	94.29 ± 12.05	100.00 ± 0.00	80.00 ± 42.16
ResNet101	95.71 ± 9.64	98.00 ± 6.32	90.00 ± 21.08
Tumors with preoperative radiotherapy			
ResNet101	90.00 ± 16.10	92.50 ± 12.08	85.00 ± 33.74
InceptionV3	93.33 ± 11.65	95.00 ± 10.54	90.00 ± 31.62
DenseNet201	93.33 ± 16.10	92.50 ± 16.87	95.00 ± 15.81
NasNetLarge	93.33 ± 16.10	95.00 ± 15.81	90.00 ± 21.08

the patients without pRT or with pRT. More interestingly, for the first time, we report an accurate AI-based classification of the biopsy IHC-staining images and its correlation of 5-year prognosis, which is expected to be of benefit for clinical treatment decision, rather than traditional IHC assay.

AUTHOR CONTRIBUTIONS


TDP, CWF, HZ, and XFS designed the research; TDP conceptualized and performed the study of AI; CWF, HZ, and XFS provided the data; TDP, CWF, HZ, and XFS contributed to the analysis of the results; and TDP, CWF, and XFS wrote the manuscript.

CONFLICT OF INTEREST

The authors declare no competing interest.

DATA AND CODE AVAILABILITY

The IHC data and Matlab code used in this study are deposited at <https://sites.google.com/view/tuan-d-pham/codes>.

Tuan D. Pham¹ 
 Chuanwen Fan²
 Hong Zhang³
 Xiao-Feng Sun²

¹ Center for Artificial Intelligence, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia

² Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

³ Department of Medical Sciences, Örebro University, Örebro, Sweden

Correspondence

Tuan D. Pham, Center for Artificial Intelligence, Prince Mohammad Bin Fahd University, Khobar 31952, Saudi Arabia; Xiao-Feng Sun, Department of Biomedical and Clinical Sciences, Linköping University, 58183 Linköping, Sweden.

Email: tpham@pmu.edu.sa; xiao-feng.sun@liu.se

ORCID

Tuan D. Pham  <https://orcid.org/0000-0002-4255-5130>

REFERENCES

- Camma C, Giunta M, Fiorica F, Pagliaro L, Craxi A, et al. Preoperative radiotherapy for resectable rectal cancer: a meta-analysis. *JAMA*. 2000;284:1008-1015.
- Swedish Rectal Cancer Trial. Improved survival with preoperative radiotherapy in resectable rectal cancer. *N Engl J Med*. 1997;8:980-987.
- van den Brink M, van den Hout WB, Stiggelbout AM, Kranenburg EK, Marijnen CAM, et al. Cost-utility analysis of preoperative radiotherapy in patients with rectal cancer undergoing total mesorectal excision: a study of the Dutch Colorectal Cancer Group. *J Clin Oncol*. 2004;22:244-253.
- Pham TD, Fan C, Pfeifer D, Zhang H, Sun XF. Image-based network analysis of DNP73 expression by immunohistochemistry in rectal cancer patients. *Front Physiol*. 2019;10:1551.
- Niehaus K, Wan N, White B, Kannan A, Gafni E, et al. Early stage colorectal cancer detection using artificial intelligence and whole-genome sequencing of cell-free DNA in a retrospective cohort of 1,040 patients. *Am J Gastroenterol*. 2018;113:S169.
- Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, et al. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci*. 2018;115:E2970-E2979.
- Takamatsu M, Yamamoto N, Kawachi H, Chino A, Saito S, et al. Prediction of early colorectal cancer metastasis by machine learning using digital slide images. *Comput Methods Programs Biomed*. 2019;178:155-161.
- Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: opportunities and challenges. *Cancer Lett*. 2020;471:61-71.
- Ho D. Artificial intelligence in cancer therapy. *Science*. 2020;367:982-983.
- Savage N. How AI is improving cancer diagnostics. *Nature*. 2020;579:S14-S16.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.