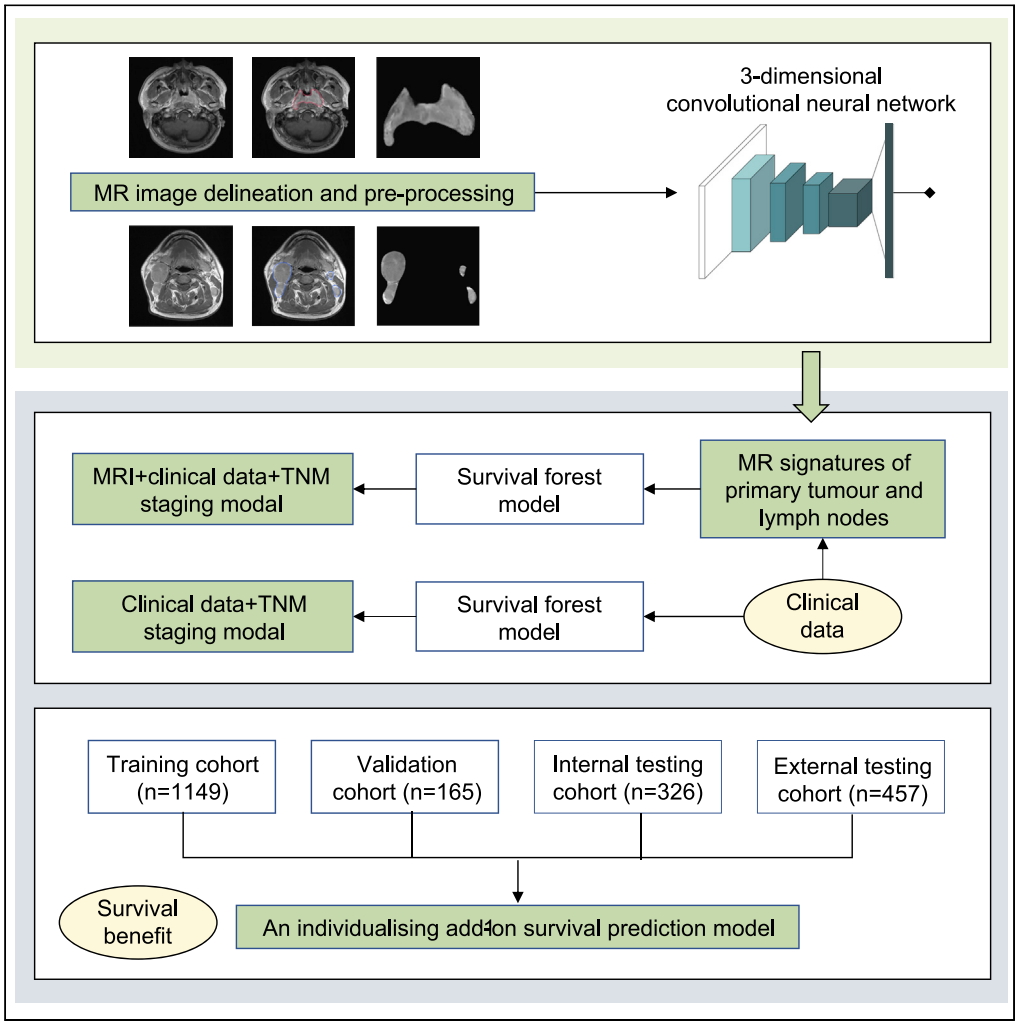
iScience

CellPress
OPEN ACCESS

Article

Add-on individualizing prediction of nasopharyngeal carcinoma using deep-learning based on MRI: A multicentre, validation study



Xun Cao, Xi Chen, Zhuo-Chen Lin, ..., Chao-Feng Li, Xiang Guo, Xing Lyu

lvxing@sysucc.org.cn

Highlights

3D-CNN was employed to extract the MRI signatures of nasopharyngeal carcinoma

The prediction model combined MRI signature, clinical data, TNM staging, and treatment

The model improved the prediction of progression-free survival and overall survival

The model can accurately predict individualized survival and decide treatment regimen

Cao et al., iScience 25, 104841
September 16, 2022
https://doi.org/10.1016/j.isci.2022.104841

iScience

CellPress
OPEN ACCESS

Article

Add-on individualizing prediction of nasopharyngeal carcinoma using deep-learning based on MRI: A multicentre, validation study

Xun Cao,[1,2] Xi Chen,[1] Zhuo-Chen Lin,[3] Chi-Xiong Liang,[1] Ying-Ying Huang,[1] Zhuo-Chen Cai,[1] Jian-Peng Li,[4] Ming-Yong Gao,[5] Hai-Qiang Mai,[1] Chao-Feng Li,[6,7] Xiang Guo,[1,7] and Xing Lyu[1,7,8,*]

SUMMARY

In nasopharyngeal carcinoma, deep-learning extracted signatures on MR images might be correlated with survival. In this study, we sought to develop an individualizing model using deep-learning MRI signatures and clinical data to predict survival and to estimate the benefit of induction chemotherapy on survivals of patients with nasopharyngeal carcinoma. Two thousand ninety-seven patients from three independent hospitals were identified and randomly assigned. When the deep-learning signatures of the primary tumor and clinically involved gross cervical lymph nodes extracted from MR images were added to the clinical data and TNM staging for the progression-free survival prediction model, the combined model achieved better prediction performance. Its application is among patients deciding on treatment regimens. Under the same conditions, with the increasing MRI signatures, the survival benefits achieved by induction chemotherapy are increased. In nasopharyngeal carcinoma, these prediction models are the first to provide an individualized estimation of survivals and model the benefit of induction chemotherapy on survivals.

INTRODUCTION

Nasopharyngeal carcinoma is a head and neck cancer arising from the mucosal lining of the nasopharynx (Chen et al., 2019). Its geographical global distribution is extremely specific; >70% of new cases occur in east Asia, southeast Asia, and northern Africa (Bray et al., 2018; Lv et al., 2021). Approximately 70-80% of patients have locoregionally advanced disease at diagnosis (Zhang et al., 2017). Fatal local, regional, and/or distant failure can occur in patients with this disease despite radical treatment (Chen et al., 2019). An early prediction of locoregional recurrence and distant metastasis could aid in the timely allocation of resources to patients with nasopharyngeal carcinoma who are appropriate targets for or who will benefit from aggressive treatment.

Magnetic resonance imaging (MRI) is a non-invasive tool routinely used for the diagnosis, staging, and response evaluation of treatment in patients with nasopharyngeal carcinoma owing to its sensitivity, specificity, and accessibility (Chen et al., 2019; Liao et al., 2008). Through radiomics analysis, a series of parameters and signatures have been linked with treatment response and outcome (Dong et al., 2019; Gillies et al., 2016; Wan et al., 2019; Zhang et al., 2017). Since 2015, deep-learning artificial intelligence methods have been shown to be efficient and superior in several predictive applications referring to medical imaging interpretation in patients with cancer (Lu et al., 2020; Peng et al., 2019; Qiang et al., 2021; Wan et al., 2019; Wu et al., 2018; Zhang et al., 2017, 2021; Zhong et al., 2020). However, in the context of nasopharyngeal carcinoma, there is a scarcity of studies that combine deep-learning methods into MR images analysis of both primary tumors and clinically involved gross cervical lymph nodes for prognosis and identify the incremental value of the signatures extracted from MR images on the prognostic model based on tumor-node-metastasis (TNM) staging and/or clinical data. Additionally, the current National Comprehensive Cancer Network (NCCN) Guidelines indicate that induction chemotherapy followed by concurrent chemoradiation (level 2A evidence) should be considered for patients with stage II-IVA nasopharyngeal carcinoma. Administration of induction chemotherapy to all patients is unnecessary and may be harmful to some, who can be cured with concurrent chemoradiation alone (Chen et al., 2019). Over-treatment will inevitably induce acute and late toxicities, and the latter can emerge months or even years after

[1]Department of Nasopharyngeal Carcinoma, Sun Yat-sen University Cancer Centre, State Key Laboratory of Oncology in South China, Collaborative Innovation Centre for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangzhou, China

[2]Department of Critical Care Medicine, Sun Yat-sen University Cancer Centre, State Key Laboratory of Oncology in South China, Collaborative Innovation Centre for Cancer Medicine, Guangzhou, China

[3]Department of Medical Records, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China

[4]Department of Radiology, Dongguan People's Hospital, Dongguan, China

[5]Department of Medical Imaging, The First People's Hospital of Foshan, Foshan, China

[6]Department of Information Technology, Sun Yat-sen University Cancer Centre, State Key Laboratory of Oncology in South China, Collaborative Innovation Centre for Cancer Medicine, Guangzhou, China

[7]These authors contributed equally

[8]Lead contact

*Correspondence: lvxing@sysucc.org.cn

https://doi.org/10.1016/j.isci.2022.104841

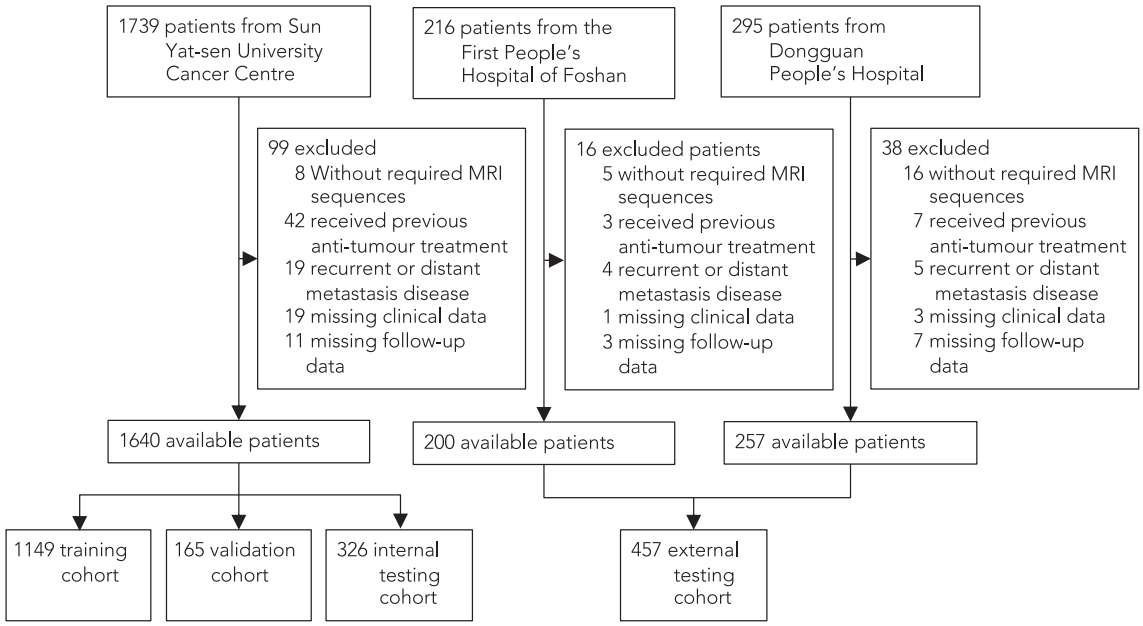


Figure 1. Study flow diagram

treatment completion. This therapeutic problem highlights and emphasizes the importance of identifying the appropriate patients who are more likely to benefit from induction chemotherapy. Models are needed to predict survival and guide individual treatment for patients with nasopharyngeal carcinoma. Unfortunately, the prognostic models reported by previous studies were mostly derived and extended using heterogeneous indicators (Tang et al., 2016, 2018), such as body mass index, lactate dehydrogenase, C-reactive protein, and miRNAs, rather than therapeutic regimen. As we know, a therapeutic regimen is a key variable in the development of the prognostic model.

Here, a deep-learning method named 3-dimensional convolutional neural network (3D-CNN) (Liang et al., 2019; Wang et al., 2019) was first used to extract and analyze both the primary tumor and clinically involved gross cervical lymph nodes in the MR images of nasopharyngeal carcinoma. We present a new approach—to employ these deep-learning signatures extracted from MR images in conjunction with a therapeutic regimen and other clinical data in a survival forest model to predict survival in nasopharyngeal carcinoma. Our aim was to develop a prognostic model that could contextualize the survival of an individual patient with nasopharyngeal carcinoma and allow modeling of the potential benefit of induction chemotherapy for survival.

RESULTS

Patient characteristics

A total of 2097 patients in three independent hospitals were recruited in this study. One thousand six hundred forty patients who presented via Sun Yat-sen University Cancer Center between November 1, 2011 and March 31, 2016 were identified and randomly divided into training (n = 1149), validation (n = 165), and internal testing (n = 326) cohorts. Two hundred patients who presented via the First People's Hospital of Foshan (between January 1, 2010 and December 31, 2015) and 257 patients who presented via Dongguan People's Hospital (between August 1, 2010 and November 30, 2016) were identified for the external testing cohort (n = 457) (Figure 1). The clinical and pathological characteristics of the patients are reported in Table 1.

The median follow-up of the Sun Yat-sen University Cancer Center cohort was 55.9 months (interquartile range [IQR] 41.4-68.7 months), and that of the First People's Hospital of Foshan cohort was 42.4 months (IQR 37.3-46.0 months), and that of the Dongguan People's Hospital cohort was 49.2 months (IQR 37.4-60.7 months).

Table 1. Patient characteristics

	Patients from Sun Yat-sen University Cancer Center (n = 1640)		Patients from The First People's Hospital of Foshan (n = 200)		Patients from Dongguan People's Hospital (n = 257)	
Age, years						
Median (IQR)	45	(38-52)	47	(40-58)	48	(40-58)
Sex						
Male	1248	(76.1%)	158	(79.0%)	195	(75.9%)
Female	392	(23.9%)	42	(21.0%)	62	(24.1%)
EBV-DNA, copies/mL						
Median (IQR)	1365	(0-11900)				
<4000	1030	(62.8%)	129	(64.5%)	163	(63.4%)
≥4000	610	(37.2%)	71	(35.5%)	94	(36.6%)
T category						
T1	187	(11.4%)	13	(6.5%)	6	(2.3%)
T2	301	(18.4%)	33	(16.5%)	94	(36.6%)
T3	870	(53.0%)	97	(48.5%)	120	(46.7%)
T4	282	(17.2%)	57	(28.5%)	37	(14.4%)
N category						
N0	274	(16.7%)	42	(21.0%)	27	(10.5%)
N1	768	(46.8%)	99	(49.5%)	68	(26.5%)
N2	348	(21.2%)	43	(21.5%)	115	(44.7%)
N3	250	(15.3%)	16	(8.0%)	47	(18.3%)
Stage						
I	0	(0%)	0	(0%)	0	(0%)
II	295	(18.0%)	29	(14.5%)	36	(14.0%)
III	768	(51.1%)	100	(50.0%)	140	(54.5%)
IV	507	(30.9%)	71	(34.5%)	81	(31.5%)
Treatment						
CCRT	808	(49.3%)	68	(34.0%)	120	(47.7%)
Induction chemotherapy plus CCRT	832	(50.7%)	132	(66.0%)	137	(53.3%)

Data are n (%) or median (IQR).
IQR, interquartile range; EBV-DNA, Epstein-Barr virus DNA; CCRT, concurrent chemoradiotherapy.

Progression-free survival prediction model

For progression-free survival, the TNM (tumor node metastasis) staging prognostic model had a C-index of 0.620 (95% CI 0.560-0.679) and an F score of 0.444 (0.393-0.494) in the internal testing cohort and a C-index of 0.607 (0.557-0.657) and an F score of 0.400 (0.358-0.443) on the external testing cohort. The clinical data-TNM staging prognostic model had a C-index of 0.719 (0.661-0.777) and an F score of 0.553 (0.502-0.603) on the internal testing cohort and a C-index of 0.622 (0.564-0.680) and an F score of 0.477 (0.436-0.519) on the external testing cohort. When the deep-learning signatures of the primary tumor and clinically involved gross cervical lymph nodes extracted from MR images ($MRIS_{pt}$ and $MRIS_{ln}$) were combined with clinical data and TNM staging, the survival prognostic model performance improved to a C-index of 0.800 (0.745-0.847), an F score of 0.628 (0.578-0.679) on the internal testing cohort and a C-index of 0.702 (0.648-0.756) and an F score of 0.542 (0.499-0.586) on the external testing cohort. The MRI-clinical data-TNM staging prognostic model exhibits a significant improvement compared with the clinical data-TNN staging-based prediction (internal testing cohort, $p = 0.0004$; external testing cohort: $p = 0.01$) and TNM staging-based prediction (internal testing cohort, $p < 0.0001$; external testing cohort: $p = 0.0032$) (Table 2). Detailed evaluation of the progression-free survival prediction model performance is provided in Table 3, Figures S1 and S2.

Table 2. Performance of progression-free survival prediction model

	C-index (95% CI)	F score (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Binomial p value[a]
Internal testing cohort					
TNM staging model	0.620 (0.560–0.679)	0.444 (0.393–0.494)	0.485 (0.367–0.604)	0.740 (0.687–0.794)	<0.0001
Clinical data-TNM staging model	0.719 (0.661–0.777)	0.553 (0.502–0.603)	0.647 (0.533–0.761)	0.694 (0.638–0.750)	0.0004
MRI- clinical data-TNM staging model	0.800 (0.745–0.847)	0.628 (0.578–0.679)	0.765 (0.664–0.866)	0.764 (0.712–0.815)	reference
External testing cohort					
TNM staging model	0.607 (0.557–0.657)	0.400 (0.358–0.443)	0.456 (0.353–0.558)	0.698 (0.651–0.745)	0.0032
Clinical data-TNM staging model	0.622 (0.564–0.680)	0.477 (0.436–0.519)	0.589 (0.487–0.691)	0.613 (0.563–0.663)	0.01
MRI- clinical data-TNM staging model	0.702 (0.648–0.756)	0.542 (0.499–0.586)	0.622 (0.522–0.722)	0.717 (0.671–0.763)	reference

C-index for right-censored data measures the model performance by comparing the survival information with predicted risk scores; a larger C-index correlates with better survival prediction performance.

C-index, concordance index; TNM staging, tumor-node-metastasis staging; MRI, magnetic resonance imaging.

[a]Measures the difference in performance between the MRI-clinical data-TNM staging model and other prediction models; a smaller p value represents greater likelihood of a difference between the MRI-clinical data-TNM staging model and other models.

Overall survival prediction model

For overall survival, the TNM staging prognostic model had a C-index of 0.666 (95% CI 0.571-0.760) and an F score of 0.305 (0.234-0.375) on the internal testing cohort and a C-index of 0.604 (0.538-0.671) and an F score of 0.389 (0.343-0.436) on the external testing cohort. The clinical data-TNM staging prognostic model had a C-index of 0.753 (0.641-0.866) and an F score of 0.367 (0.290-0.444) on the internal testing cohort and a C-index of 0.645 (0.578-0.713) and an F score of 0.428 (0.380-0.475) on the external testing cohort. When the deep-learning signatures of the primary tumor and clinically involved gross cervical lymph nodes extracted from MR images ($MRIS_{pt}$ and $MRIS_{ln}$) were combined with clinical data and TNM staging, the survival prognostic model performance improved to a C-index of 0.815 (0.739-0.892), an F score of 0.369 (0.298-0.449) on the internal testing cohort and a C-index of 0.702 (0.632-0.772) and an F score of 0.506 (0.457-0.556) on the external testing cohort. Compared with the TNM staging prognostic model, the combined prognostic model using the deep-learning signatures of the primary tumor and clinically involved gross cervical lymph nodes extracted from MR images, clinical data, and TNM staging achieved significantly better performance in both the internal testing cohort ($p = 0.038$) and the external testing cohort ($p = 0.042$) (Table 4). A detailed evaluation of the overall survival prognostic model performance is provided in Table 5, Figure S3, and S4.

Proposed clinical utility of the model

We expect that the primary utility of this prognostic model will be among patients deciding between induction chemotherapy plus concurrent chemoradiation and concurrent chemoradiation alone. Example outputs from the prognostic models for nine hypothetical vignettes are presented in Figure 2. The statuses of all clinical data, TNM staging, and MRI signatures ($MRIS_{pt}$ and $MRIS_{ln}$) are altered within each patient to demonstrate the potential benefit of induction chemotherapy. Under the same conditions of clinical data and TNM staging, with the increasing $MRIS_{pt}$ and $MRIS_{ln}$, the survival benefits achieved by induction chemotherapy are increased. For example, a 58-year-old man patient with the disease characteristics shown in Figure 2 has an estimated 1.14% 5-year progression-free survival benefit from induction chemotherapy when $MRIS_{pt}$ is 22 and $MRIS_{ln}$ is 51. Although $MRIS_{pt}$ rose to 42 and $MRIS_{ln}$ rose to 96, the 5-year progression-free survival benefits improved by induction chemotherapy are increased to 6.82%. The increased $MRIS_{pt}$ and $MRIS_{ln}$ partly yield treatment benefits of induction chemotherapy. Figure 3 shows visualization for the prediction of 5-year progression-free survival benefit improved by induction chemotherapy in representative cases.

DISCUSSION

The prediction of progression-free survival and overall survival in patients with nasopharyngeal carcinoma is important, because early aggressive intervention has been shown to improve mortality (Chen et al., 2019). MRI is a versatile imaging modality that has shown efficiency in diagnosing, staging, and prognosis of

Table 3. Comparison of area under the receiver operating characteristic curve of progression-free survival prediction model on internal and external testing cohorts

	Internal testing cohort	External testing cohort
	ROC-AUC (95% CI)	ROC-AUC (95% CI)
TNM staging model	0.637 (0.571–0.702)	0.606 (0.551–0.662)
Clinical data-TNM staging model	0.747 (0.679–0.816)	0.630 (0.564–0.696)
MRI-clinical data-TNM staging model	0.829 (0.772–0.887)	0.718 (0.659–0.776)

A larger ROC-AUC represents better prediction performance.
ROC-AUC, area under the receiver operating characteristic curve; MRI, magnetic resonance imaging; TNM staging, tumor-node-metastasis staging.

nasopharyngeal carcinoma (Chen et al., 2019; Liao et al., 2008). In this study, using a new deep-learning method, 3D-CNN, we first extracted the signatures of both the primary tumor and clinically involved gross cervical lymph nodes from MR images of nasopharyngeal carcinoma. In addition, the combined survival prognostic models based on the MRI signatures of the primary tumor and clinically involved gross cervical lymph nodes, clinical data, and TNM staging were implemented and shown to incrementally improve the prognostic ability as compared with prediction using clinical data and/or TNM staging alone. More importantly, to our knowledge, these survival prognostic models first incorporate the impact of induction chemotherapy, facilitating comparison against the option of concurrent chemoradiation alone within the context of a nasopharyngeal carcinoma patient's survival benefit.

Deep-learning is a data-driven application of artificial intelligence in which systems automatically learn and analyze without explicit programming. Therefore, deep-learning can autonomously exploit datasets to identify new variables and more complex relationships between them. Its application in medical imaging has increasingly been used to develop novel prognostic models in nasopharyngeal carcinoma (Peng et al., 2019; Qiang et al., 2021; Wan et al., 2019; Zhang et al., 2017).

Our study has several differences compared with previous studies. The drawbacks of earlier works in this area include the use of a small cohort and inadequate variables, absence of estimation of sensitivity and specificity, and direct categorization of patients into risk-stratified groups (Peng et al., 2019; Qiang et al., 2021; Wan et al., 2019; Zhang et al., 2017, 2021). Additionally, in the earlier literature, therapeutic modalities (eg, concurrent chemoradiotherapy with or without induction chemotherapy) have not been input and developed for the prediction models (Qiang et al., 2021; Zhang et al., 2017, 2021; Zhong et al., 2020), and we acknowledged that lacking such important information might have introduced bias. A previous

Table 4. Performance of overall survival prediction model

	C-index (95% CI)	F score (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Binomial p value[a]
Internal testing cohort					
TNM staging model	0.666 (0.571–0.760)	0.305 (0.234–0.375)	0.625 (0.388–0.862)	0.710 (0.659–0.760)	0.038
Clinical data-TNM staging model	0.753 (0.641–0.866)	0.367 (0.290–0.444)	0.750 (0.537–0.960)	0.709 (0.649–0.751)	0.058
MRI- clinical data-TNM staging model	0.815 (0.739–0.892)	0.369 (0.298–0.449)	0.751 (0.538–0.962)	0.752 (0.704–0.800)	reference
External testing cohort					
TNM staging model	0.604 (0.538–0.671)	0.389 (0.343–0.436)	0.485 (0.367–0.604)	0.694 (0.648–0.740)	0.042
Clinical data-TNM staging model	0.645 (0.578–0.713)	0.428 (0.380–0.475)	0.603 (0.487–0.719)	0.612 (0.563–0.660)	0.087
MRI- clinical data-TNM staging model	0.702 (0.632–0.772)	0.506 (0.457–0.556)	0.662 (0.549–0.774)	0.699 (0.654–0.745)	reference

C-index for right-censored data measures the model performance by comparing the survival information with predicted risk scores; a larger C-index correlates with better survival prediction performance.
C-index, concordance index; TNM staging, tumor-node-metastasis staging; MRI, magnetic resonance imaging.
[a]Measures the difference in performance between the MRI-clinical data-TNM staging model and other prediction models; a smaller p value represents greater likelihood of a difference between the MRI-clinical data-TNM staging model and other models.

Table 5. Comparison of area under the receiver operating characteristic curve of overall survival prediction model on internal and external testing cohorts

	Internal testing cohort	External testing cohort
	ROC-AUC (95% CI)	ROC-AUC (95% CI)
TNM staging model	0.675 (0.547–0.802)	0.615 (0.553–0.678)
Clinical data-TNM staging model	0.771 (0.654–0.889)	0.641 (0.569–0.712)
MRI-clinical data-TNM staging model	0.818 (0.737–0.899)	0.708 (0.637–0.779)

A larger ROC-AUC represents better prediction performance.
ROC-AUC, area under the receiver operating characteristic curve; MRI, magnetic resonance imaging; TNM staging, tumor-node-metastasis staging.

study on MRI features found MRI scores of the primary tumor as a significant prognosticator of failure-free survival, but it did not include the MRI scores of clinically involved gross cervical lymph nodes, nor did it discuss the added value of MRI scores to clinical indicators in developing the prognostic model (Wan et al., 2019). Zhong and colleagues identified the value of deep learning on MR-based radiomics for prediction in patients with stage T3N1M0 nasopharyngeal carcinoma treated with induction chemotherapy plus concurrent chemotherapy (Zhong et al., 2020). Peng and colleagues identified that deep-learning positron emission tomography with computed tomography (PET-CT)-based radiomics serves as an indicator for predicting outcomes in patients with advanced nasopharyngeal carcinoma (Peng et al., 2019). However, studies have shown that using MRI rather than PET-CT is more sensitive and accurate when detecting and evaluating primary tumor extension, retropharyngeal lymph node metastasis, and cervical lymph node metastasis (Chen et al., 2019; Liao et al., 2008). Therefore, MRI might be preferred for clinical decision-making. Another relevant study that only included 118 patients found that multiparametric MRI-based radiomics nomograms could provide improved prognostic ability in advanced nasopharyngeal carcinoma, but it did not consider MRI of clinically involved gross cervical lymph nodes for its prediction models (Zhang et al., 2017). In addition, as the study did not collect treatment data, it could not consider and model the effects of treatments on outcomes, which are key variables in prediction models. Moreover, the models provided by the two studies have not yet been externally tested (Peng et al., 2019; Zhang et al., 2017). The largest study to date using deep-learning to examine nasopharyngeal carcinoma included 3444 patients. A prognostic system incorporated MRI scores with clinical data and achieved a statistically improved C-index in the prediction of disease-free survival, overall survival, and distant metastasis-free survival (Qiang et al., 2021). Prognostic models are often limited by which variables investigators have traditionally considered important and consequently collected (eg, clinical variables, pathological variables, and treatment). Our study indicates that model performance improves with greater granularity, and prognostic models could have an advantage when incorporating MRI signatures. Our study is unique and clinically relevant because it showed that using deep-learning signatures extracted from MR images can incrementally increase the strength of survival predictions and outperform predictions by clinical data and/or TNM staging models.

Existing studies mainly focus on the segmentation of primary tumors, eliding the recognition of cervical lymph nodes, and thus fail to comprehensively provide a landscape for tumor identification. Li and colleagues proposed an automatic segmentation network, named by NPCNet, to achieve segmentation of primary tumors and lymph nodes simultaneously. The NPCNet was conducted on the dataset of 9124 samples collected from 754 patients. The results demonstrated that NPCNet achieved state-of-the-art performance (Li et al., 2022). Similarly, Tao and colleagues developed a sequential method (SeqSeg) to achieve accurate nasopharyngeal carcinoma segmentation. Next, the performance of SeqSeg was evaluated on the large nasopharyngeal carcinoma dataset containing 1101 patients. The experimental results demonstrated that the proposed SeqSeg not only outperforms several state-of-the-art methods but also achieves better performance in multi-device and multi-centre datasets (Tao et al., 2022). The application of NPCnet and Seqseg to segment primary tumor and/or lymph nodes as markers for nasopharyngeal carcinoma patient staging is beneficial in prediction and radiotherapy planning.

Currently, the NCCN Guidelines recommend induction chemotherapy followed by concurrent chemoradiation as level 2A evidence for stage II-IIA nasopharyngeal carcinoma. However, randomized trials have

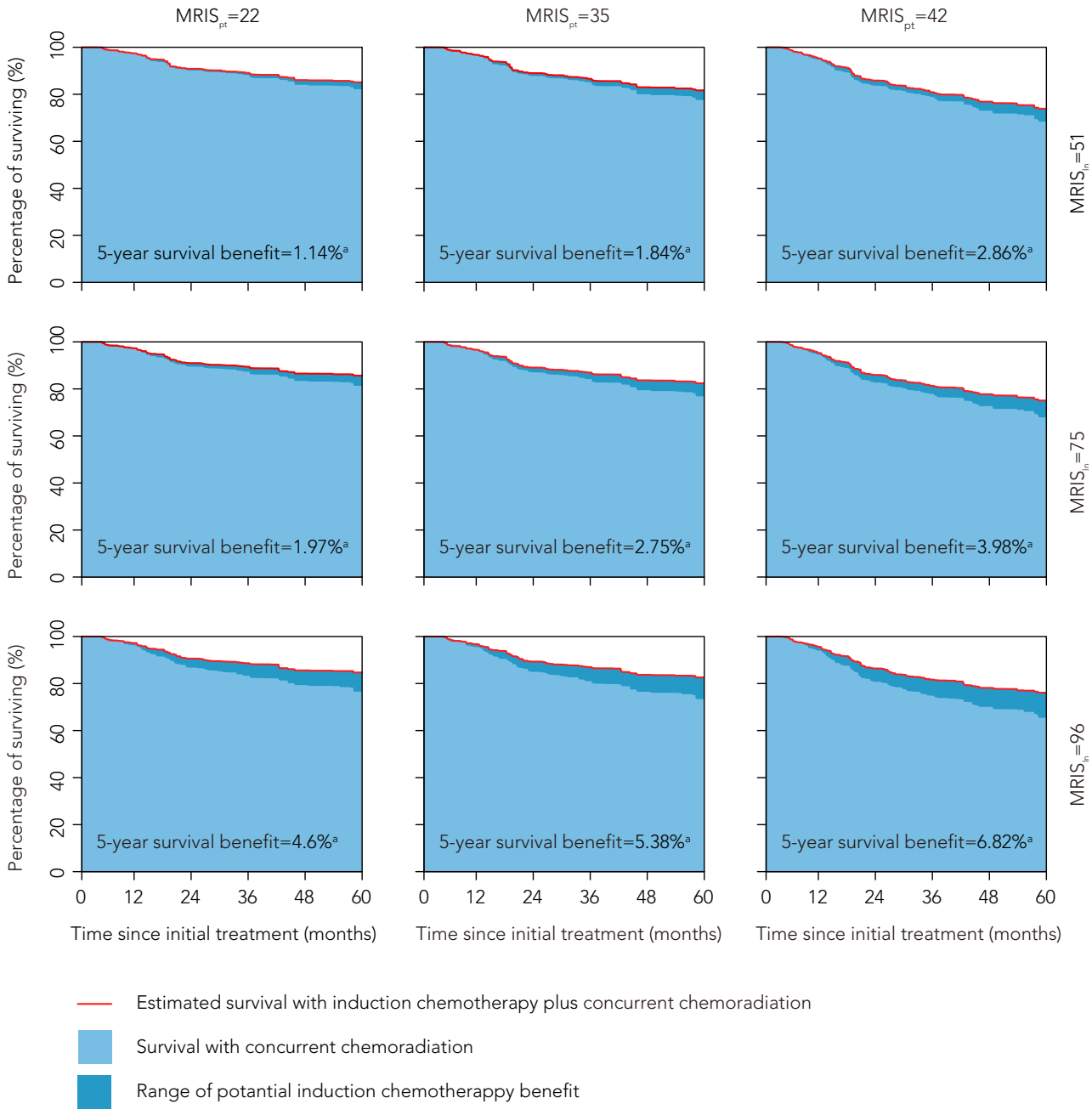


Figure 2. Example prognostic model outputs using 5-year progression-free survival for nine hypothetical vignettes

Only the MRIS$_{pt}$ and MRIS$_{ln}$ status changed between each column and row to demonstrate the growth in 5-year progression-free survival benefits from induction chemotherapy. [a]Prediction of 5-year progression-free survival benefit improved by induction chemotherapy. MRIS$_{pt}$, MRI signature of primary tumor; MRIS$_{ln}$, MRI signature of clinically involved gross cervical lymph nodes.

shown inconsistent results regarding the efficacy of additional induction chemotherapy (Fountzilas et al., 2012; Hui et al., 2009; Tan et al., 2015), probably suggesting that some patients will not benefit from it. Adding induction chemotherapy to concurrent chemoradiation is associated with risks of significant toxic effects (Chen et al., 2019; Lv et al., 2021). Identifying patients appropriate for induction chemotherapy and conveying this information to their own survivals is a currently crucial exercise. Most prognostic models are established by simple categorization of patients into risk-stratified groups and are potentially conflicted by a bias toward the treatments they received. Our study incorporates treatment in developing

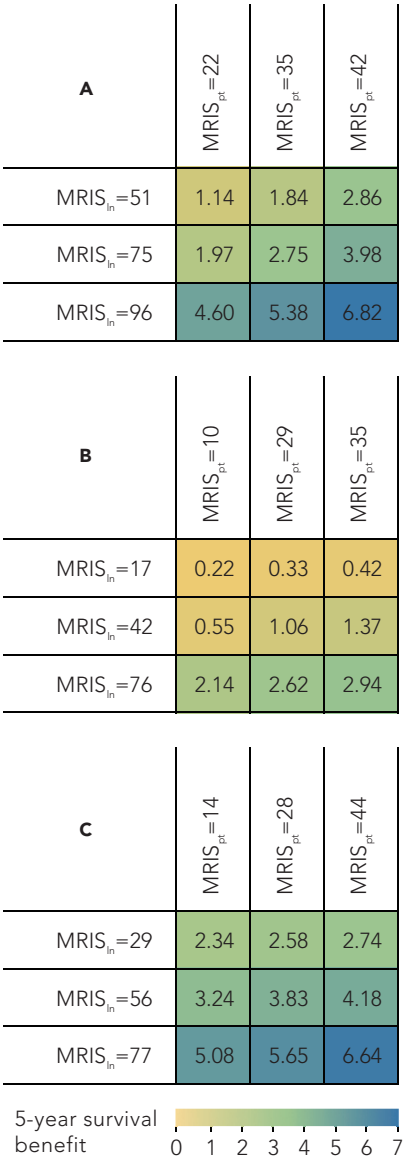


Figure 3. Visualization for the prediction of 5-year progression-free survival benefit improved by induction chemotherapy in representative cases

(A) Representative case 1, a 58-year-old man, EBV-BNA = 12580 copies/mL, stage III disease, with altered MRI signatures ($MRIS_{pt}$ and $MRIS_{In}$) is shown.

(B) Representative case 2, a 50-year-old woman, EBV-BNA = 9690 copies/mL, stage II disease, with altered MRI signatures ($MRIS_{pt}$ and $MRIS_{In}$) is shown.

(C) Representative case 3, a 60-year-old man, EBV-BNA = 16,800 copies/mL, stage IV disease, with altered MRI signatures ($MRIS_{pt}$ and $MRIS_{In}$) is shown. $MRIS_{pt}$, MRI signature of primary tumor; $MRIS_{In}$, MRI signature of clinically involved gross cervical lymph nodes; EBV-DNA, Epstein-Barr virus DNA.

prognostic models, which will help address a critical gap in clinical practice and improve multidisciplinary treatment decision-making. In conclusion, to the best of our knowledge, our survival prediction models were the first based on deep-learning MRI signatures of both primary tumor and clinically involved cervical lymph nodes, treatment regimens, and clinical data (including TNM staging) that augmented the performances of predicting survival in patients with nasopharyngeal carcinoma. These prediction models uniquely provide individualized estimation of survival and model the impact of induction chemotherapy on these outcomes.

Limitations of the study

First, compared with the internal testing cohort, the survival prognostic models showed a decreased performance in the external testing cohort, which suggested that generalization might be less possible. These results could be caused by the heterogeneous MR scanning and image acquisition among the three different hospitals, as lower performance on an external testing cohort is a common phenomenon in deep-learning studies using multicentre data. Of note, adding MRI signatures to clinical data and TNM staging improved the model performance both in the internal and external testing cohorts. Second, in the prognostic models, the sensitivity exhibits greater improvements than the specificity. The clinicians judge sensitivity to be the more clinically important measure of prognostic accuracy in patients with nasopharyngeal carcinoma patients owing to the clinical consequences of missing the aggressive disease. Third, although our study had a large cohort of patients from three independent hospitals, the sample size might still be smaller than for deep-learning models. Further improvements in model performance can be achieved by using a larger population, together with further external testing by other researchers using our model codes. Additionally, as with genomic tests, any prognostic molecular markers can be integrated into future refinement of our prognostic models.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Study design and patients
  - Clinical and treatment information collection
  - MRI acquisition
  - MR image delineation and pre-processing
  - The detail of 3D-CNN
  - 3D-CNN for MRI signatures
  - Survival forest model
  - Follow-up and outcomes
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104841.

ACKNOWLEDGMENTS

This study was supported by grants from the National Natural Science Foundation of China (No. 81872375, No. 82172863), the Natural Science Foundation of Guangdong Province (No. 2021A1515010118), and the Science and Technology Planning Project Foundation of Guangdong Province (No. 2019B110233004). We thank Jian Wang (Tiehan Environment Protection Group Co., Ltd) for his assistance with data management and logistic support. The article has been edited by Elsevier Language Editing Services (Serial number: LE-219398-0260DB70C8CD).

AUTHOR CONTRIBUTIONS

Xu.C., X.G., and X.L. conceived and designed the study. Xu.C., Xi.C., C.X.L., Y.Y.H., Z.C.C., J.P.L., M.Y.G., and X.L. acquired the data. Xu.C., Xi.C., Z.C.L., H.Q.M., X.G., and X.L. implemented quality control of data and the algorithms. Xu.C., Xi.C., Z.C.L., X.G., and X.L. verified the data. Xu.C., Z.C.L., X.G., and X.L. did the statistical analyses. Xu.C., Z.C.L., and X.L. developed, trained, and applied the deep-learning model. All authors prepared the first draft of the article and generated the figures. Xu.C., Xi.C., Z.C.L., H.Q.M., X.G., and X.L. revised the article. All authors read and reviewed the final article. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 15, 2022
Revised: July 8, 2022
Accepted: July 21, 2022
Published: September 16, 2022

REFERENCES

Amin, M.B.; American Joint Committee on Cancer (2017). In AJCC cancer staging manual, Eighth edition (Springer).

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA-Cancer J Clin *68*, 394–424.

Cai, T., Pepe, M.S., Zheng, Y., Lumley, T., and Jenny, N.S. (2006). The sensitivity and specificity of markers for event times. Biostatistics *7*, 182–197.

Chen, Y.P., Chan, A.T.C., Le, Q.T., Blanchard, P., Sun, Y., and Ma, J. (2019). Nasopharyngeal carcinoma. Lancet *394*, 64–80.

Chua, D.T., Sham, J.S., Kwong, D.L., Tai, K.S., Wu, P.M., Lo, M., Yung, A., Choy, D., and Leong, L. (1997). Volumetric analysis of tumor extent in nasopharyngeal carcinoma and correlation with treatment outcome. Int. J. Radiat. Oncol. Biol. Phys. *39*, 711–719.

Dong, D., Zhang, F., Zhong, L.Z., Fang, M.J., Huang, C.L., Yao, J.J., Sun, Y., Tian, J., Ma, J., and Tang, L.L. (2019). Development and validation of a novel MR imaging predictor of response to induction chemotherapy in locoregionally advanced nasopharyngeal cancer: a randomized controlled trial substudy (NCT01245959). BMC Med. *17*.

Fountzilas, G., Ciuleanu, E., Bobos, M., Kalogera-Fountzila, A., Eleftheraki, A.G., Karayannopoulou, G., Zaramboukas, T., Nikolaou, A., Markou, K., Resiga, L., et al. (2012). Induction chemotherapy followed by concomitant radiotherapy and weekly cisplatin versus the same concomitant chemoradiotherapy in patients with nasopharyngeal carcinoma: a randomized phase II study conducted by the Hellenic Cooperative Oncology Group (HeCOG) with biomarker evaluation. Ann. Oncol. *23*, 427–435.

Gillies, R.J., Kinahan, P.E., and Hricak, H.H. (2016). Radiomics: images are more than pictures, they are data. Radiology *278*, 563–577.

Harrell, F.E., Jr., Lee, K.L., and Mark, D.B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat. Med. *15*, 361–387.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In Paper presented at: IEEE International Conference on Computer Vision.

Huang, G., Liu, Z., Laurens, V., and Weinberger, K.Q. (2016). Densely connected convolutional networks. In Paper presented at: IEEE Computer Society.

Hui, E.P., Ma, B.B., Leung, S.F., King, A.D., Mo, F., Kam, M.K., Yu, B.K., Chiu, S.K., Kwan, W.H., Ho, R., et al. (2009). Randomized phase II trial of concurrent cisplatin-radiotherapy with or without neoadjuvant docetaxel and cisplatin in advanced nasopharyngeal carcinoma. J. Clin. Oncol. *27*, 242–249.

Ishwaran, H., Kogalur, U.B., Blackstone, E.H., and Lauer, M.S. (2008). Random survival forests. Ann. Appl. Stat. *2*, 841–860.

Kingma, D., and Ba, J. (2014). Adam: a method for stochastic optimization. Computer Science.

Li, Y., Dan, T., Li, H., Chen, J., Peng, H., Liu, L., and Cai, H. (2022). NPCNet: jointly segment primary nasopharyngeal carcinoma tumors and metastatic lymph nodes in MR images. IEEE Trans. Med. Imag. *41*, 1639–1650.

Liang, S.J., Tang, F., Huang, X., Yang, K.F., Zhong, T., Hu, R.Y., Liu, S.Q., Yuan, X.R., and Zhang, Y. (2019). Deep-learning-based detection and segmentation of organs at risk in nasopharyngeal carcinoma computed tomographic images for radiotherapy planning. Eur. Radiol. *29*, 1961–1967.

Liao, X.B., Mao, Y.P., Liti, L.Z., Tang, L.L., Sun, Y., Wang, Y., Lin, A.H., Cui, C.Y., Li, L., and Ma, J. (2008). How does magnetic resonance imaging influence staging according to Ajcc staging system for nasopharyngeal carcinoma compared with computed tomography? Int. J. Radiat. Oncol. Biol. Phys. *72*, 1368–1377.

Lin, M., Chen, Q., and Yan, S. (2014). Network in network. In Paper presented at: ICLR.

Llorca, J., and Delgadorodríguez, M. (1972). Cox DR. Regression models and life-tables. J. Roy. Stat. Soc. *34*, 187–220.

Lu, C., Bera, K., Wang, X.X., Prasanna, P., Xu, J., Janowczyk, A., Beig, N., Yang, M., Fu, P.F., Lewis, J., et al. (2020). A prognostic model for overall survival of patients with early-stage non-small cell lung cancer: a multicentre, retrospective study. Lancet Digit Health *2*, E594–E606.

Lv, X., Cao, X., Xia, W.X., Liu, K.Y., Qiang, M.Y., Guo, L., Qian, C.N., Cao, K.J., Mo, H.Y., Li, X.M., et al. (2021). Induction chemotherapy with lobaplatin and fluorouracil versus cisplatin and fluorouracil followed by chemoradiotherapy in patients with stage III-IVB nasopharyngeal carcinoma: an open-label, non-inferiority, randomised, controlled, phase 3 trial. Lancet Oncol. *22*, 716–726.

Nair, V., and Hinton, G.E. (2010). Rectified linear units improve restricted Boltzmann machines vinod nair. In Paper presented at: International Conference on International Conference on Machine Learning.

Peng, H., Dong, D., Fang, M.J., Li, L., Tang, L.L., Chen, L., Li, W.F., Mao, Y.P., Fan, W., Liu, L.Z., et al. (2019). Prognostic value of deep learning PET/CT-Based radiomics: potential role for future individual induction chemotherapy in advanced nasopharyngeal carcinoma. Clin. Cancer Res. *25*, 4271–4279.

Qiang, M.Y., Li, C.F., Sun, Y.Y., Sun, Y., Ke, L.R., Xie, C.A.M., Zhang, T., Zou, Y.J., Qiu, W.Z., Gao, M.Y., et al. (2021). A prognostic predictive system based on deep learning for locoregionally advanced nasopharyngeal carcinoma. J. Natl. Cancer Inst. *113*, 606–615.

Saito, T., and Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One *10*, e0118432.

Tan, T., Lim, W.T., Fong, K.W., Cheah, S.L., Soong, Y.L., Ang, M.K., Ng, Q.S., Tan, D., Ong, W.S., Tan, S.H., et al. (2015). Concurrent chemo-radiation with or without induction gemcitabine, carboplatin, and paclitaxel: a randomized, phase 2/3 trial in locally advanced nasopharyngeal carcinoma. Int. J. Radiat. Oncol. Biol. Phys. *91*, 952–960.

Tang, L.Q., Li, C.F., Li, J., Chen, W.H., Chen, Q.Y., Yuan, L.X., Lai, X.P., He, Y., Xu, Y.X.X., Hu, D.P., et al. (2016). Establishment and validation of prognostic nomograms for endemic nasopharyngeal carcinoma. J. Natl. Cancer Inst. *108*, djv291.

Tang, X.R., Li, Y.Q., Liang, S.B., Jiang, W., Liu, F., Ge, W.X., Tang, L.L., Mao, Y.P., He, Q.M., Yang, X.J., et al. (2018). Development and validation of a gene expression-based signature to predict distant metastasis in locoregionally advanced nasopharyngeal carcinoma: a retrospective, multicentre, cohort study. Lancet Oncol. *19*, 382–393.

Tao, G., Li, H., Huang, J., Han, C., Chen, J., Ruan, G., Huang, W., Hu, Y., Dan, T., Zhang, B., et al. (2022). SeqSeg: a sequential method to achieve nasopharyngeal carcinoma segmentation free from background dominance. Med. Image Anal. *78*, 102381.

Wan, Y., Tian, L., Zhang, G., Xin, H., Li, H., Dong, A., Liang, Y., Jing, B., Zhou, J., Cui, C., et al. (2019). The value of detailed MR imaging report of primary tumor and lymph nodes on prognostic nomograms for nasopharyngeal carcinoma after intensity-modulated radiotherapy. Radiother. Oncol. *131*, 35–44.

Wang, Y.N., Liu, C.B., Zhang, X., and Deng, W.W. (2019). Synthetic CT generation based on T2 weighted MRI of nasopharyngeal carcinoma (NPC) using a deep convolutional neural network (DCNN). Front. Oncol. *9*, 1333.

Wu, S.X., Zheng, J.J., Li, Y., Wu, Z., Shi, S.Y., Huang, M., Yu, H., Dong, W., Huang, J., and Lin, T.X. (2018). Development and validation of an MRI-based radiomics signature for the preoperative prediction of lymph node metastasis in bladder cancer. EBioMedicine *34*, 76–84.

Zhang, B., Tian, J., Dong, D., Gu, D.S., Dong, Y.H., Zhang, L., Lian, Z.Y., Liu, J., Luo, X.N., Pei, S.F., et al. (2017). Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. Clin. Cancer Res. *23*, 4259–4269.

Zhang, L., Wu, X., Liu, J., Zhang, B., Mo, X., Chen, Q., Fang, J., Wang, F., Li, M., Chen, Z., et al. (2021). MRI-based deep-learning model for distant metastasis-free survival in locoregionally advanced nasopharyngeal carcinoma. J. Magn. Reson. Imag. *53*, 167–178.

Zhong, L.Z., Fang, X.L., Dong, D., Peng, H., Fang, M.J., Huang, C.L., He, B.X., Lin, L., Ma, J., Tang, L.L., et al. (2020). A deep learning MR-based radiomic nomogram may predict survival for nasopharyngeal carcinoma patients with stage T3N1M0. Radiother. Oncol. *151*, 1–9.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
R (version 4.0.3)	R software	https://www.r-project.org
SPSS	IBM corporation	https://www.ibm.com/analytics
3D-CNN	(Huang et al., 2016)	https://github.com/liuzhuang13/DenseNet
	(He et al., 2015)	https://ieeexplore.ieee.org/document/7410480
	(Nair and Hinton, 2010)	https://icml.cc/Conferences/2010/
Survival prediction model	(Ishwaran et al., 2008)	https://cran.r-project.org/package=randomSurvivalForest
Other		
AJCC TNM staging system	(Amin et al., 2017)	https://link.springer.com/book/9783319406176
Research Data Deposit	This paper	https://www.researchdata.org.cn

RESOURCE AVAILABILITY

Lead contact

Further information about the methods and requests for data or scripts should be directed to and will be fulfilled by the lead contact, Dr. Xing Lyu (lvxing@sysucc.org.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

De-identified data have been deposited at the Research Data Deposit public platform (No. RDDA2022773130) and are publicly available as of the date of publication.

The source code is available for public use at: https://github.com/Anchorage-c/XC-npc-individualising-prediction-model.git.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Study design and patients

In this multicentre, retrospective study, we identified patients with nasopharyngeal carcinoma who presented to Sun Yat-sen University Cancer Center in Guangzhou, China. The patients were randomly assigned to training, validation, and internal testing cohorts (7:1:2). Patients who presented to the First People's Hospital of Foshan, Foshan, China, and Dongguan People's Hospital, Dongguan, China, were assigned to external testing cohort.

Eligible individuals were those who had histologically confirmed nasopharyngeal carcinoma, pre-treatment contrast-enhanced MRI of nasopharynx and neck available, received concurrent chemoradiation with or without induction chemotherapy, and complete clinical and pathological characteristics and follow-up data. The exclusion criteria for the study were a history of cancer, receipt of previous anti-tumour treatment, and presence of recurrent or distant metastasis. Patients were also excluded if the primary tumor could not be identified via MRI. Of note, if a patient had multiple MRIs from the admission, then the MRI taken closest to the time of initial treatment (within two weeks) were collected and included in the study.

All patients underwent contrast-enhanced MRI of the nasopharynx and neck before initial treatment (concurrent chemoradiation or induction chemotherapy plus current chemoradiation). Two investigators (X.L.

and H.Q,M.) independently re-evaluated all patients' MR images. Disagreements were reviewed a second time, followed by a conclusive judgment by both investigators. We re-staged all patients according to the American Joint Committee on Cancer staging system (eighth edition) (Amin et al., 2017). The institutional ethical review boards of all participating hospitals included in this study provided approval. The requirement for informed consent was waived for this retrospective analysis.

Clinical and treatment information collection

For each patient, clinical data including age, sex, Epstein-Barr virus (EBV) DNA, tumor (T) and lymph node (N) status, stage, and treatment regimens were collected.

All patients underwent radiotherapy. Patients from SunYat-sen University Cancer Center (Guangzhou, China) and the First People's Hospital of Foshan (Foshan, China) received intensity-modulated radiotherapy (IMRT). A total of 6 (2.3%) of 257 patients from Dongguan People's Hospital (Dongguan, China) received two-dimensional radiotherapy (2D-RT), and 251(97.7%) of 257 patients received IMRT. The cumulative radiation doses were 66 Gy or greater to the primary tumor and 60-70 Gy to the involved neck area. All potential sites of local infiltration and bilateral cervical lymphatics were irradiated to 50 Gy or greater. All patients were treated with 30–35 fractions with five daily fractions per week for 6–7 weeks 1101 (52.5%) of 2097 patients received induction chemotherapy. The induction chemotherapy regimens included cisplatin plus 5-fluorouracil, taxanes plus cisplatin, and taxanes plus cisplatin plus 5-fluorouracil, which was administrated triweekly for two or three cycles. All patients received concurrent chemotherapy. The regimen is cisplatin on weeks 1, 4 and 7 of radiotherapy, or weekly during radiotherapy.

MRI acquisition

Covering a region from the frontal sinuses to the lower margin of the sternal collarbone, MRI scans were performed on a 1.5 T imaging system (SIGNA™ EXCITE, GE Healthcare System, USA; SIGNA™ HDx, GE Healthcare System, USA) or 3.0 T imaging system (Discovery™ MR750, GE Healthcare, USA; MAGNETOM Trio Tim, SIEMENS, Germany). Before injecting Gd-DTPA (Magnevist, Bayer Schering Pharma AG, Germany) via the superficial vein, T1-weighted imaging (T1w) was performed in the axial, coronal and sagittal planes, and T2-weighted imaging (T2w) was performed in the axial plane. After Gd-DTPA injection with a dose of 0.3–0.4 mmol per kg, contrast-enhanced T1w (T1c) in the axial, coronal and sagittal planes and contrast-enhanced T2w (T2c) in the axial plane were executed. Additional information about MRI scanner is provided in Table S1.

MR image delineation and pre-processing

All MR images were retrieved from the picture archiving and communication system (PACS, Carestream). The primary tumor and clinically involved gross cervical lymph nodes were delineated on the MR images (twenty-five-layer of three sequences) by two investigators (Xi.C. and X.L.) using the Medical Imaging Interaction Toolkit software (version MITK-2016.11.0). Retropharyngeal lymph node involvement was included as part of the primary tumor because a clear distinction between the retropharyngeal node and primary tumor remains difficult in nasopharyngeal carcinoma (Chua et al., 1997). For primary tumor that was directly contiguous with the regional nodes, a cut-off level at the mid-C2 vertebra was used to separate the primary tumor from the positive lymph nodes (Chua et al., 1997). Both investigators were blinded to the clinical data but were aware that the patients had nasopharyngeal carcinoma. Any disagreement was resolved by reaching consensus regarding delineation. To focus analysis on the primary tumor and the clinically involved gross cervical lymph nodes, the image information outside the tumor and the clinically involved gross cervical lymph nodes was removed by applying the binary mask (Figure S5). Finally, an image patch with a size of 384 × 384 that centered around the primary tumor and the clinically involved gross cervical lymph nodes was respectively fed into the network for subsequent analysis.

Standardization was conducted as follow step:

We calculated the pixel mean value "u" of the delineated region.

$$u = \frac{1}{N}\sum_{i=0}^{N} X_i$$

X was pixel set in delineated region, N was total number of pixels in delineated region, X_i was pixel value of the number i pixel.

$$m = \frac{1}{\sum_{i=0}^{N} 1_{(X_i > u)}} \sum_{i=0}^{N} X_{i(X_i > u)}$$

X was pixel set in delineated region, u was mean pixel value of delineated region, X_i was pixel value of the number i pixel.

$$\widehat{X} = \frac{X - m}{m}$$

$\widehat{X}$ was standardized X.

The detail of 3D-CNN

The 3-dimensional convolutional neural network (3D-CNN) (Liang et al., 2019; Wang et al., 2019) consisted of two main modules: (1) 3D-DenseNet; and (2) DeepSurvivalNet. In 3D-DenseNet, each DenseNet learned deep representations and extracted image features from three magnetic resonance imaging (MRI) sequences (T1w, T2w and T1c), respectively. In DeepSurvivalNet, the low-dimensional features, extracted by 3D-DenseNet, were concatenated and fed into the fully connected layers, then MRI signatures (including MRI signature of primary tumor [$MRIS_{pt}$] and MRI signature of clinically involved gross cervical lymph nodes [$MRIS_{ln}$]) were outputted. Detailed description of 3D-CNN as follows:

The three MRI sequence (axial T1w, T2w, and T1c, respectively) are fed into the 3D convolution layer and the max pooling layer. The 3D-DenseNet has three dense blocks and three transition layers. The algorithm structure of the dense blocks and transition layers is the same as that described by Huang and colleagues (Huang et al., 2016). Each block obtains additional inputs from all preceding blocks and passes on its own feature maps to all subsequent blocks. The layers between two adjacent blocks are referred to as transition layers and change feature map sizes via convolution and pooling. The feature maps generated from the last block are fed into a batch normalization (BN) (He et al., 2015) layer and then activated by a rectified linear unit (ReLU) (Nair and Hinton, 2010). Finally, the last feature maps are subjected to global average pooling (GAP) (Lin et al., 2014) and then flattened as MRI features. The subnetwork, consisting of multiple fully connected layers, is referred to as DeepSurvivalNet. In each fully connected layer, we employed ReLU (pReLU) (He et al., 2015) as the nonlinear activation function.

To address right-censored survival outcomes and predict risk, the loss of the network was calculated using a Cox negative logarithm partial likelihood loss function (Cox loss), an extended application of the Cox proportional hazards model (Llorca and Delgadorodríguez, 1972).

Data augmentation is an important step in training deep networks. During training, we performed data augmentation as follows: rotation $\pm 5°$, shift $\pm 10\%$, and horizontal flip. The batch size was 32. The times needed for DenseNet-T1w, DenseNet-T2w, and DenseNet-T1c to train for 100 epochs were 8 h 24 m, 8 h 25 m, and 8 h 17 m, respectively. The three 3D-DenseNets (DenseNet-T1w, DenseNet-T2w, and DenseNet-T1c) and DeepSurvivalNet were optimized using initial learning rates of 4e-5, 2e-5, 2e-5, and 6e-6, respectively. The training process of three 3D-DenseNets and DeepSurvivalNet was achieved by minimizing the Cox loss using adaptive moment estimation (Adam) (Kingma and Ba, 2014) (Figure S6).

3D-CNN for MRI signatures

The MR images were analyzed by a deep-learning method named 3D-CNN (Liang et al., 2019; Wang et al., 2019), which consisted of two main modules: 3-dimensional DenseNet (3D-DenseNet) and DeepSurvivalNet (See Figure S6). We trained the deep-learning model to predict survival based on MR images and corresponding survival information. By uploading the delineated and masked MR images of primary tumor and clinically involved gross cervical lymph nodes of each patient, separately, we analyzed each image with the 3D-CNN. The 3D-CNN interpreted MR images without any clinical data (eg, T category, N category, and stage). After analysing all of one patient's MR images, the 3D-CNN output the signatures of the primary tumor ($MRIS_{pt}$) and clinically involved gross cervical lymph nodes ($MRIS_{ln}$), respectively (Figure S7).

Survival forest model

The survival forest is a variant of the random forest adapted for right-censored survival (Ishwaran et al., 2008). It is optimized by assigning risk scores to patients according to their survival. In each of the survival forest models, a collection of decision trees was used to model the complex relationships between input parameter vectors and the survival prediction.

The survival forest model (Ishwaran et al., 2008) was utilized to predict the survival of patients with nasopharyngeal carcinoma. For the clinical data-TNM staging prognostic model, the input variables were the four clinical characteristics (age, sex, EBV-DNA, and stage) and treatment. The variables were input to a survival forest model to derive clinical data-TNM staging-based risk scores. For the MRI-clinical data-TNM staging prognostic model, the input variables were the same four clinical characteristics, treatment, and the signatures (including $MRIS_{pt}$, and $MRIS_{ln}$) extracted from MR images by the 3D-CNN. The variables were input into the survival forest model to derive MRI-clinical data-TNM staging-based score, which acts as the combined risk for each patient (Figure S7).

Follow-up and outcomes

Follow-ups were conducted every three months over the first two years after primary treatment, once semi-annually in the third to fifth years, and once yearly thereafter. We chose progression-free survival as our primary endpoint, which was defined as the time from the first day of initial treatment to documented disease progression (locoregional recurrence and/or distant metastasis), death from any cause, or date of last follow-up visit, whichever occurred first. The secondary endpoint was overall survival. We calculated overall survival as the interval between the first day of initial treatment to the date of death from any cause or the date of the last follow-up visit.

QUANTIFICATION AND STATISTICAL ANALYSIS

The concordance index (C-index) for right-censored data was employed to evaluate the performance of survival prognostic models by comparing the survival with the ranks of predicted risk scores (Harrell et al., 1996). The stratification performances of different prognostic models were evaluated using a log rank test based on the predicted risk scores. The area under the receiver operating characteristic curve (ROC-AUC) was calculated for the survival prognostic model (Cai et al., 2006). The precision-recall curves and F-scores were used to evaluate the survival prognostic models (Saito and Rehmsmeier, 2015). Different survival prognostic models were compared using a binomial test to show the differences in performance. A two-sided p value of less than 0.05 was considered statistically significant. All statistical analyses were performed using R (version 4.0.3) and SPSS (version 26.0).

ADDITIONAL RESOURCES

This study was conducted with approval of the Institutional Ethics Review Board of Sun Yat-sen University Cancer Center.