Frontiers in Bioscience Landmark

Front. Biosci. (Landmark Ed) **2022**; 27(7): 212
https://doi.org/10.31083/j.fbl2707212

Original Research

Deep Learning Empowers Lung Cancer Screening Based on Mobile Low-Dose Computed Tomography in Resource-Constrained Sites

Jun Shao[1,†], Gang Wang[2,†], Le Yi[3], Chengdi Wang[1], Tianzhong Lan[3], Xiuyuan Xu[3], Jixiang Guo[3], Taibing Deng[4], Dan Liu[1], Bojiang Chen[1], Zhang Yi[3,*], Weimin Li[1,*]

[1]Department of Respiratory and Critical Care Medicine, Frontiers Science Center for Disease-Related Molecular Network, West China Hospital, West China School of Medicine, Sichuan University, 610041 Chengdu, Sichuan, China
[2]Precision Medicine Research Center, West China Hospital, Sichuan University, 610041 Chengdu, Sichuan, China
[3]Machine Intelligence Laboratory, College of Computer Science, Sichuan University, 610065 Chengdu, Sichuan, China
[4]Department of Respiratory Disease, Guang'An Hospital, 638001 Guangan, Sichuan, China
*Correspondence: weimi003@scu.edu.cn (Weimin Li); zhangyi@scu.edu.cn (Zhang Yi)
[†]These authors contributed equally.
Academic Editor: Graham Pawelec
Submitted: 31 December 2021 Revised: 9 March 2022 Accepted: 15 March 2022 Published: 4 July 2022

Abstract

Background: Existing challenges of lung cancer screening included non-accessibility of computed tomography (CT) scanners and inter-reader variability, especially in resource-limited areas. The combination of mobile CT and deep learning technique has inspired innovations in the routine clinical practice. **Methods**: This study recruited participants prospectively in two rural sites of western China. A deep learning system was developed to assist clinicians to identify the nodules and evaluate the malignancy with state-of-the-art performance assessed by recall, free-response receiver operating characteristic curve (FROC), accuracy (ACC), area under the receiver operating characteristic curve (AUC). **Results**: This study enrolled 12,360 participants scanned by mobile CT vehicle, and detected 9511 (76.95%) patients with pulmonary nodules. Majority of participants were female (8169, 66.09%), and never-smokers (9784, 79.16%). After 1-year follow-up, 86 patients were diagnosed with lung cancer, with 80 (93.03%) of adenocarcinoma, and 73 (84.88%) at stage I. This deep learning system was developed to detect nodules (recall of 0.9507; FROC of 0.6470) and stratify the risk (ACC of 0.8696; macro-AUC of 0.8516) automatically. **Conclusions**: A novel model for lung cancer screening, the integration mobile CT with deep learning, was proposed. It enabled specialists to increase the accuracy and consistency of workflow and has potential to assist clinicians in detecting early-stage lung cancer effectively.

Keywords: pulmonary nodules; deep learning; CT images; lung cancer screening; malignancy risk

1. Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide in 2020 accounting for 18.4% of overall cancer deaths [1]. The 5-year survival of lung cancer was less than 20% in China due to delayed diagnosis [2]. Patients with early-stage lung cancer who received curative treatment would have a better prognosis substantially, compared with those with lung cancer at advanced stage [3]. Only 17.3% of patients were diagnosed with lung cancer at stage I, which was inferior to America (25.3%). Further, the proportion of stage I was higher in urban (19.5%) than in rural (11.1%) areas in China [4]. Challenges for early detection of lung cancer are warranted to address in China.

Low-dose computed tomography (LDCT) has been proved to significantly reduce the mortality of lung cancer [5]. National Lung Screening Trial (NLST) demonstrated a reduction in the lung-cancer mortality with LDCT screening of about 20% as compared with that in the chest radiography group [6]. Meanwhile, Nederlands-Leuvens Longkanker Screenings Onderzoek (NELSON) lung-cancer screening trial indicated a cumulative rate ratio of 0.76 for death from lung cancer in the computed tomography (CT) arm relative to the no screening control arm [7]. However, CT scanners were relatively unavailable in resource-constrained areas, which caused diagnosis at advanced stage. What's more, inter-reader variability among radiologists might lead to the missed diagnosis, clinical and financial cost waste [8–10]. It is imperative to improve the accessibility and consistency of lung cancer screening especially in the resource-limited sites.

The mobile low-dose whole body CT screening unit has the potential to promote the availability of lung cancer screening. It not only provides reliable imaging results, but also reduces geographic barriers, making lung cancer screening more extensively [11]. It had been implemented in Kenya, community in Yorkshire and UK [12,13]. On the other hand, the rapid development of deep learning has revaluated medical routine at a broad variety of screening and evaluation of lung cancer, diagnosis of COVID-19, detection of lymph node metastases in breast cancer [14–19]. Based on CT images, deep learning approach detects holis-

tic nodules to automate standard image analysis. Previous studies were limited to small and retrospective samples. The combination of mobile CT vehicle and deep learning model still warrants further investigation.

Herein, a prospective study of lung cancer screening based on mobile CT in rural area was declared, and a sophisticated deep learning system was constructed to assist clinicians to detect the nodules and predict the malignancy risk.

2. Materials and Methods

2.1 Screening Trial Oversight and Population Recruitment

Natural Population Cohort Study of West China Hospital of Sichuan University, a prospective community-based trial of screening in rural areas such as Mianzhu (Site 1), Longquan (Site 2), Pidu, and Seda, Sichuan Province, China, aimed to build a cohort of 80,000 natural population. The lung cancer screening trial was approved by the ethics committee of West China Hospital of Sichuan University.

During January 2020 to June 2020, recruitment of physical examination population was carried out in the corresponding community. The cohort invited permanent residents who were over 20 years old. Each participant provided written informed consent and completed the questionnaire including demographic information, behavior habits such as smoking, disease history like chronic obstructive pulmonary disease (COPD), and family history of cancer. People over 40 years old were eligible for lung cancer screening. Exclusion criteria were as follows: previously diagnosed with lung cancer, pulmonary surgery, chest CT scans within the past 6 months, unwilling to receive mobile CT examination. The screening program in Site 1 started from Jul 2020 to Sep 2020, and the program in Site 2 started from Oct 2020 to Nov 2020. All data had passed the strict quality-control and were stored in Electronic Data Capture (EDC) system.

Chest CT images were acquired from mobile LDCT scanners (Neusoft Corporation). Scanning parameters were standardized as follows: tube voltage of 120 kV of, 512 $\times$ 512 of pixel matrix. Images were reconstructed to high-resolution images (window width 1800, level –400, slice thickness 1 mm), then sent to Picture Archiving and Communication Systems (PACS) to West China Hospital of Sichuan University through cloud network.

2.2 Deep Learning Network Architecture

We built a two-stage deep learning model to identify high-risk patients from the population-based cohort (Fig. 1). Specifically, the first-stage model detected suspicious nodule candidates from LDCT images of the participants. In the second stage, then, the model input each nodule's image patches and predicted the corresponding probability of malignancy. Accordingly, the malignancy risk for each patient was depended on the detected nodule which had the highest probability of being lung cancer.

To detect lung nodules from LDCT images, we constructed the detection model and trained the framework followed the steps by our preceding work [20]. Given a LDCT scan, the detection pipeline first went through a series of pre-processing steps (containing lung segmentation, cropping the lung field, resizing, dividing into patches, and normalizing) to adapt to the input space of the model. At the end of the forward propagation, the region proposal network output predictions including the probability of the region being a lung nodule, the 3-dimensional coordinates, and the diameter. Then, non-maximum suppression algorithm was used to remove overlapping predictions. In order to derive a well detector, we iteratively optimized the parameters of the model under the guidance of a loss function on the training dataset. The detection loss function consisted of two parts including the cross-entropy loss (CEL) for classifying whether the region being a nodule and the smoothed L1-norm loss for regressing location and diameter of lung nodules. More details about the training setup could be referred to our previous studies [20]. In this study, we set smaller anchor sizes compared with the consideration that natural population tends to have smaller lung nodules, and their sizes were 5.0 mm, 10.0 mm, and 20.0 mm, respectively [20].

We further constructed a classification model, which predicted probability of each nodule candidate being malignant. The risk of each participant was determined by the most-likely cancerous nodule. For each nodule candidate (detected by the first stage model), the resolution was first resampled to $1.0 \times 1.0 \times 1.0$ mm for each voxel, and a patch with the size of $32 \times 64 \times 64$ located at the nodule's center was cropped from the resampled images. These image patches were then inputted into the classification model, a 3D ResNet-18 network as the backbone, to forward compute the predicted probability of malignancy risk. We did not use deeper networks since the deep counterparts had more trainable parameters and were more likely to be overfitted when trained on small datasets. CEL was employed to evaluate the distance from predictions to ground-truth labels, so that to guide the model to update its parameters. Stochastic gradient descent algorithm was employed as the optimizer, with initial learning rate setting to 0.01, momentum setting to 0.9, and weight decay setting to 0.001. The model was trained with 200 epochs, and learning rate would decrease by multiplying 0.1 every 50 epochs. Moreover, data augmentation operations such as random flip, rotation, and center perturbation were employed to alleviate overfitting risk of the model. The model parameters obtaining the best accuracy on the validation set were selected, and then the results on the testing set were evaluated.

The performance of the proposed method was evaluated according to available targets from two aspects: (1) evaluating two models respectively by giving the nodule-wise location and the corresponding risk; (2) evaluating the overall performance of the framework by giving the patient-

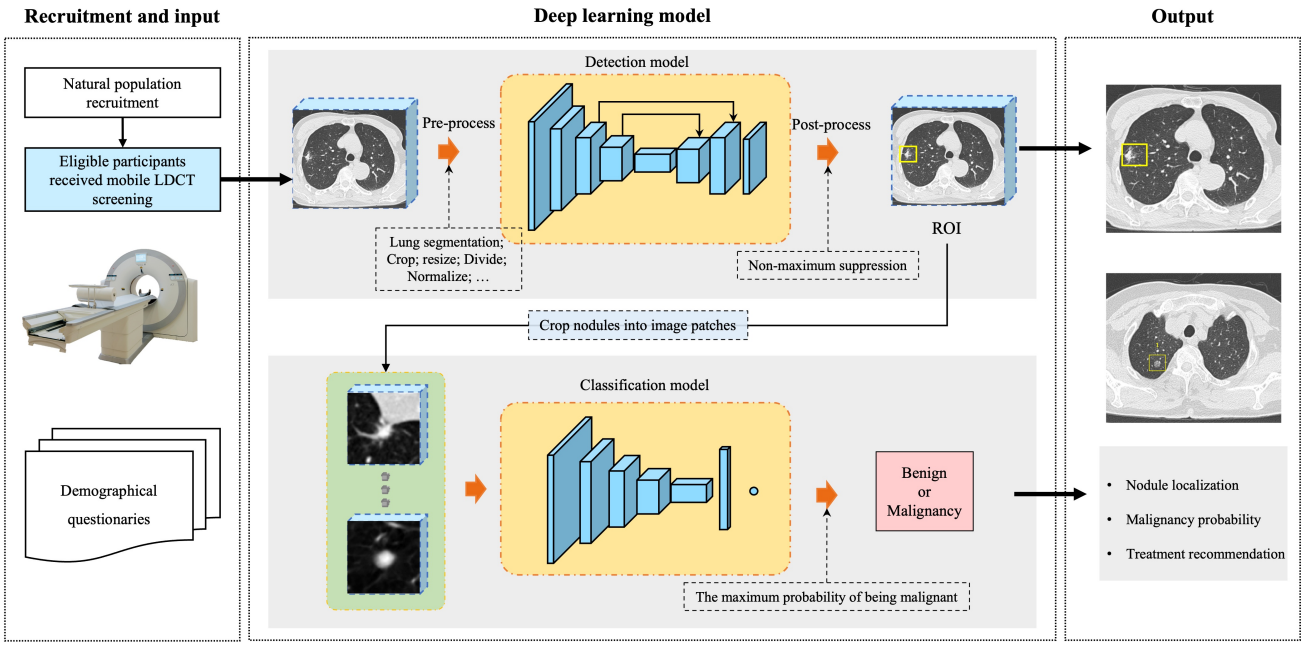

Fig. 1. Overall modeling framework. Each enrolled participant was required mobile CT scanning and demographic questionaries. The deep learning model inputs CT volume, analyzes volumetric ROIs, and outputs nodule localization and its malignancy. Abbreviations: LDCT, low-dose computed tomography; ROI, region of interest.

wise cancerous risk. We split the dataset into 7:1.5:1.5 for training,validation and testing regardless of which evaluating approach was used. It could ensure no prior knowledge of nodule location was available on the testing set when evaluates the performance of the proposed framework. In other word, the overall framework was only provided with LDCT images, and the detection model provided the location of nodule candidates for the classification model. All in all, the key index of our framework was depended on results of the latter evaluation methods since it imitated the real diagnostic procedure in clinical practice.

2.3 Eligibility Criteria and Nodule Management Protocol

The questionnaires were strictly defined. A former or current smoker was defined as a person who had smoked more than one cigarette per day lasting more than 6 months. Occupational exposure history focused on exposure to radon,asbestos, arsenic, dust and oil fume at work for more than 12 months. Tumor history was expected lung cancer. Family included parents, brothers, sisters, grandparents, grandsons, and other immediate family members.

Remote reading of CT images was managed by experienced radiologists and nodules were classified into four existing Lung-RADS risk buckets [21]. Lung-RADS 1 was defined low-risk without nodules or with definitely benign nodules. Since Lung-RADS 2 and 3 have relatively mediate risk of malignancy, the above both were grouped in mid-risk bucket for this experiment. Lung-RADS 4A/B/X suspected or confirmed malignancy was defined as high-risk. Clinicians combined with results of deep learning model and imaging reports to judge whether the nodule was at high risk, and further recommend patients for treatment and follow-up. Patients diagnosed with lung cancer pathologically by surgery or biopsy would further receive standardized medical treatment according to lung cancer guideline of China [22] and NCCN guidelines [23].

2.4 Follow-Up Visits

The follow-up to December 2021 was conducted by the whole process management center of West China Hospital of Sichuan University. The clinical information of patients with lung cancer were retrieved from hospital information system (HIS) or telephone visit. Patients with low-risk nodules or non-abnormalities were recommend to screen annually.

2.5 Statistical Analysis

Demographic information was analyzed utilizing Student's *t*-test for continuous variables and chi-square test for categorical variables. Two-side $p < 0.05$ was recognized statistically significant. Only an inference procedure was required for the population-based cohort. The experimental code and deep learning model were implemented on an Ubuntu 20.04 server equipped with Python (version 3.7.0, https://www.python.org/) and PyTorch (version 1.7.0, https://pytorch.org/). As the two-stage model can be split to two binary classification tasks, we used free-response receiver operating characteristic curve (FROC), accuracy (ACC), area under the receiver operating characteristic curve (AUC), F1 score and 95% confidence interval

Fig. 2. Participant flowchart through the screening pilot in two sites. There were four stage including recruitment, CT screening, nodules detection, pathological diagnosis for lung cancer patients. Site 1, Mianzhu; Site 2, Longquan, Sichuan province, China.

(CI) to evaluate the performance of the model.

3. Results

3.1 Baseline Characteristics of the Participants

After recruiting 19,281 people, a total of 12,360 participants were enrolled in this lung cancer screening cohort including 4730 people of Site 1 (Mianzhu) and 7630 people of Site 2 (Longquan) (Fig. 2, Table 1). The median age was 58 years old. 8169 (66.09%) participants were female, and 9784 (79.16%) participants were never smokers. The vast majority of people had no history of occupational exposure (11,075, 89.60%). A few people had a history of COPD (85, 0.69%), previous tumor history (129, 1.04%), family history of cancer (1169, 9.46%), or family history of lung cancer (329, 2.66%). Baseline characteristics differed significantly between the two sites such as age ($p < 0.001$), sex ($p < 0.001$), smoking history ($p < 0.001$), occupational exposure history ($p < 0.001$), family history of cancer ($p < 0.001$), family history of lung cancer ($p < 0.001$), except for COPD ($p = 0.14$) and tumor history ($p = 0.87$). 9511 (76.95%) patients were detected with pulmonary nodules and detection rate of nodules was similar between the two locations (Site 1, 3649/4,730, 77.15%; Site 2, 5862/7630, 76.83%; $p = 0.70$).

3.2 Results of 1-Year Follow-up

In the 1-year follow-up round, 86 patients (86/12,360, 0.70%) were test positive of lung cancer including 35 (35/4730, 0.74%) in Site 1 and 51 (51/7630, 0.67%) in Site 2 (Table 2). Most patients were female (59, 68.60%), never smoking (68, 79.07%) and had no family history of lung cancer (83, 96.51%). There was almost no statistical difference in the clinical features of the two sites: sex ($p = 0.47$), smoking history ($p = 0.93$), family history of lung cancer ($p = 1$), except age ($p < 0.001$). Numerous patients underwent surgery (81, 94.19%), and several patients received chemotherapy (9, 10.47%), radiotherapy (2, 2.32%), and targeted therapy (5, 5.81%). Adenocarcinoma accounted for the majority (80, 93.03%) regardless of Site 1 or Site 2 (32, 91.43%; 48, 94.12%; $p = 0.43$). Lung cancer patients were diagnosed mainly at stage I (73, 84.88%).

3.3 Performance on Utilization of Deep Learning Model

Deep learning algorithm assisted clinicians to detect nodules and stratify risk effectively (Fig. 3, Table 3). With regard to nodules detection, the recall yielded 0.8953 (95% CI: 0.8693–0.9207) and FROC score was 0.6359 (95% CI: 0.6354–0.6384) in validation dataset; the recall reached 0.9507 (95% CI: 0.9342–0.9679) and FROC score was 0.6470 (95% CI: 0.6467–0.6495) in testing set. In term of risk stratification, the performance of deep learning model was promising with ACC of 0.9036 (95% CI: 0.8715–0.9317), macro-AUC of 0.8798 (95% CI: 0.8304–0.9248) in validation set, and ACC of 0.8696 (95% CI: 0.8370–0.9022), macro-AUC of 0.8516 (95% CI: 0.7934–0.9051) in testing dataset. Especially for the identification of high-risk groups, the deep learning model achieved superior performance with recall of 0.8571 (95% CI: 0.6667–1.0000), precision of 1.0000, F1 of 0.9630 (95% CI: 0.8889–1.0000), and AUC of 0.9894 (95% CI: 0.9702–1.0000) in validation set, and recall of 0.6923 (95% CI: 0.5000–0.8947), precision of 1.0000, F1 of 0.7619 (95% CI: 0.5600–0.9231), and AUC of 0.9634 (95% CI: 0.9328–0.9906) in testing dataset.

This AI system might be a versatile tool for physicians. As an example of clinical deployment shown in Fig. 4, it could automatically locate pulmonary nodules, predict the degree of risk and recommend treatment for the next step.

4. Discussion

A prospective lung cancer screening cohort was conducted in two rural areas of West China. A total of 12,360

Table 1. Demographic characteristic of lung cancer screening cohort.

Characteristic	Total N (%)	Site 1 N (%)	Site 2 N (%)	*p*
	N = 12360	N = 4730	N = 7630	
Age (SD)	58.19 (9.65)	61.13 (9.83)	56.36 (9.07)	<0.001
Sex				<0.001
Male	4191 (33.91)	1729 (36.55)	2462 (32.27)	
Female	8169 (66.09)	3001 (63.45)	5168 (67.73)	
Smoking history				<0.001
Current or former	2576 (20.84)	1181 (24.97)	1395 (18.28)	
Never	9784 (79.16)	3549 (75.03)	6235 (81.72)	
Occupational exposure history				<0.001
Yes	1285 (10.40)	266 (5.62)	1019 (13.36)	
No	11075 (89.60)	4464 (94.38)	6611 (86.64)	
COPD				0.14
Yes	85 (0.69)	26 (0.55)	59 (0.77)	
No	12275 (99.31)	4704 (99.45)	7571 (99.23)	
Tumor history				0.87
Yes	129 (1.04)	48 (1.01)	81 (1.06)	
No	12231 (98.96)	4682 (98.99)	7549 (98.94)	
Family history of cancer				<0.001
Yes	1169 (9.46)	275 (5.81)	894 (11.72)	
No	11191 (90.54)	4455 (94.19)	6736 (88.28)	
Family history of lung cancer				<0.001
Yes	329 (2.66)	63 (1.33)	266 (3.49)	
No	12031 (97.34)	4667 (98.67)	7364 (96.51)	
Patients with pulmonary nodule				0.70
Yes	9511 (76.95)	3649 (77.15)	5862 (76.83)	
No	2849 (23.05)	1081 (22.85)	1768 (23.17)	

Abbreviation: COPD, chronic obstructive pulmonary disease.

Table 2. Clinical features of lung cancer patients with 1-year follow-up.

Characteristic	Total N (%)	Site 1 N (%)	Site 2 N (%)	*p*
	N = 86	N = 35	N = 51	
Age (SD)	59.92 (9.26)	64.00 (8.69)	57.12 (8.65)	<0.001
Sex				0.47
Male	27 (31.40)	13 (37.14)	14 (27.45)	
Female	59 (68.60)	22 (62.86)	37 (72.55)	
Smoking history				0.93
Current or former	18 (20.93)	8 (22.86)	10 (19.61)	
Never	68 (79.07)	27 (77.14)	41 (80.39)	
Family history of lung cancer				1
Yes	3 (3.49)	1 (2.86)	2 (3.92)	
No	83 (96.51)	34 (97.14)	49 (96.08)	
Treatment				
Surgery	81 (94.19)	33 (94.29)	48 (94.12)	
Chemotherapy	9 (10.47)	2 (5.71)	7 (13.73)	
Radiotherapy	2 (2.32)	0 (0)	2 (3.92)	
Targeted therapy	5 (5.81)	0 (0)	5 (9.80)	
Histopathology				0.43
Adenocarcinoma	80 (93.03)	32 (91.43)	48 (94.12)	
Squamous cancer	4 (4.65)	2 (5.71)	2 (3.92)	
Small cell lung cancer	1 (1.16)	0 (0)	1 (1.96)	
Sarcomatoid carcinoma	1 (1.16)	1 (2.86)	0 (0)	
Stage				0.30
I	73 (84.88)	31 (88.57)	42 (82.36)	
II	6 (6.98)	1 (2.86)	5 (9.80)	
III	5 (5.81)	3 (8.57)	2 (3.92)	
IV	2 (2.33)	0 (0)	2 (3.92)	

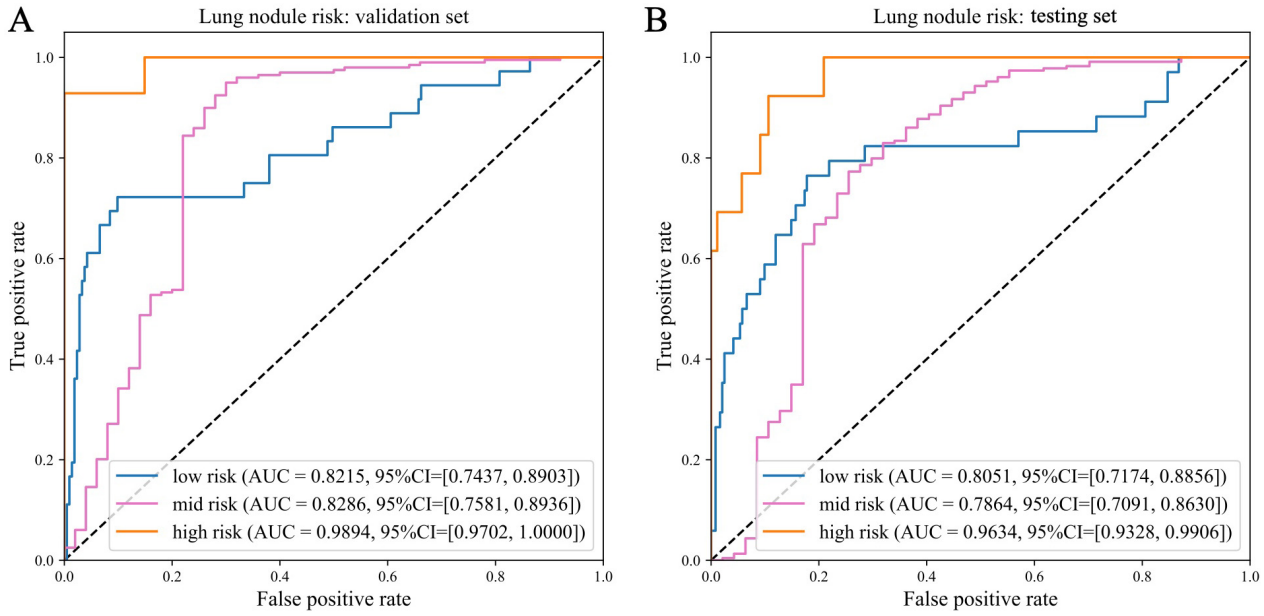

Fig. 3. Deep learning performance to identify nodule and predict its malignancy. Receiver operating characteristics (ROC) curves to identify nodule in (A) validation set and (B) testing set. Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval.

Table 3. Performance of deep learning model to detect nodules and predict malignancy.

Nodule Detection	ACC	Recall (95% CI)	Precision (95% CI)	F1 (95% CI)	FROC score (95% CI)
Validation set		**0.8953 (0.8693–0.9207)**	–	–	0.6359 (0.6354–0.6384)
Testing set		**0.9507 (0.9342–0.9679)**	–	–	0.6470 (0.6467–0.6495)
Risk Classification	**ACC**	**Recall (95% CI)**	**Precision (95% CI)**	**F1 (95% CI)**	**AUC (95% CI)**
Validation set					
Low risk		0.5833 (0.4483–0.7143)	0.7241 (0.5833–0.8571)	0.6462 (0.5172–0.7463)	**0.8215 (0.7437–0.8903)**
Mid risk	0.9036 (0.8715–0.9317)	0.9598 (0.9344–0.9805)	0.9227 (0.8910–0.9517)	0.9409 (0.9196–0.9590)	**0.8286 (0.7581–0.8936)**
High risk		0.8571 (0.6667–1.0000)	1 (1.0000–1.0000)	0.9630 (0.8889–1.0000)	**0.9894 (0.9702–1.0000)**
Testing set					
Low risk		0.4706 (0.3243–0.6129)	0.5517 (0.4000–0.7037)	0.5079 (0.3704–0.6349)	**0.8051 (0.7174–0.8856)**
Mid risk	0.8696 (0.8370–0.9022)	0.9432 (0.9188–0.9661)	0.9038 (0.8710–0.9367)	0.9231 (0.9006–0.9441)	**0.7864 (0.7091–0.8630)**
High risk		0.6923 (0.5000–0.8947)	1 (1.0000–1.0000)	0.7619 (0.5600–0.9231)	**0.9634 (0.9328–0.9906)**

Abbreviations: ACC, accuracy; FROC, free-response receiver operating characteristic curve; AUC, area under the receiver operating characteristic curve; CI, confidence interval.

participants were enrolled undergoing mobile CT vehicle and 86 patients were diagnosed with lung cancer after one year follow-up. What's more, a deep learning model was constructed to aid clinicians to recognize high-risk nodules. This novel model made lung cancer screening possible in resource-deficient areas.

Participants in this screening group reported a high rate of pulmonary nodules (9511/12,360, 76.95%). In previous screening trails, the proportion of patients with lung nodule was 22–59% at baseline [5,24–26]. The incidence of lung cancer (86/12,360, 0.70%) was inferior than NLST (first screening, 270/26,309, 1.03%) but higher than the crude incidence of lung cancer reported by National Cancer Center in 2015 (57.26 per 100,000) [6,27]. This difference might be due to the different inclusion criteria of patients. Eligible participants in NLST were between 55 and 74 years old, who had a heavy cigarette smoking history of at least 30 pack-years, or had quit within the last 15 years. National Cancer Center collected cancer data from 368 cancer registries from across China covering rural and urban areas. Furthermore, the majority of patients were female (59/86, 68.60%) and non-smokers (68/86, 79.07%). Existing screening standards emphasized male, heavy smoking history and over 55 years old, but now female, non-smoking and young lung cancer should also be taken seriously [6,7]. Future precise screening should focus on a subset of individuals at high risk of this particular cancer within the general population.

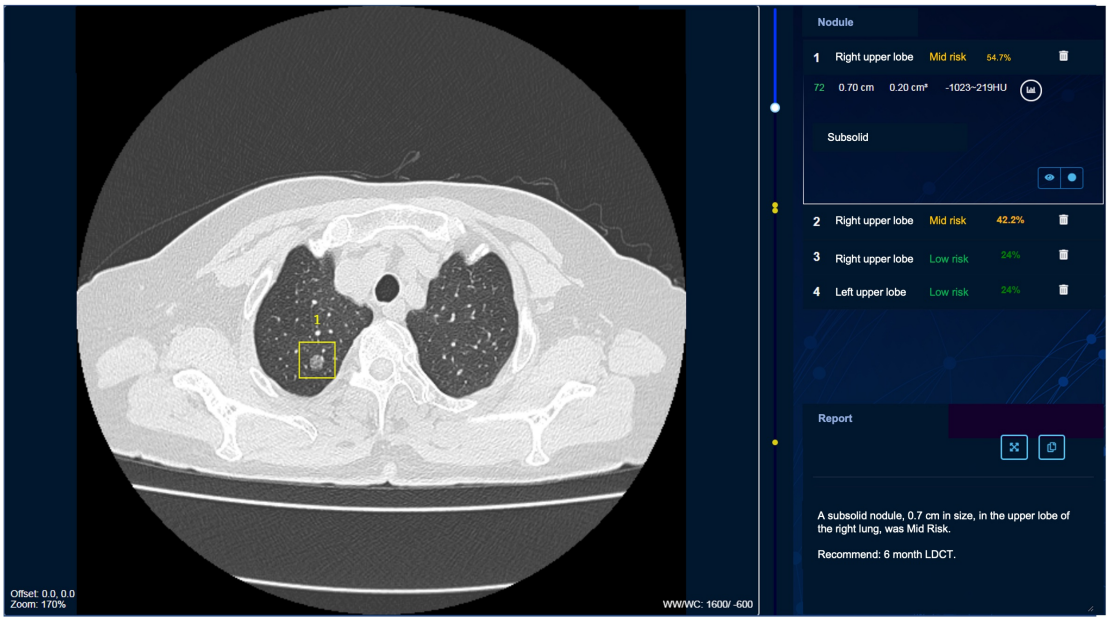

Fig. 4. Illustration of deep learning system for the detection and diagnosis of lung cancer in clinical application. This system provided localization and risk analysis of nodules on CT images.

As far as we know, this was by far the largest prospective screening program using mobile CT involving more than 10,000 people. Although mobile CT solved the geographical restrictions in remote areas, someone preferred hospital-based CT during this program. Residents' health awareness and acceptance of lung cancer screening should be strengthened by primary-care physicians and specialists. Another strength of this study was the application of deep learning model in clinical routine work. This system could assess localization and malignancy risk calculations of prior imaging, which enabled specialists to gain the efficiency and consistency of workflow. However, the deep learning system recognized no more than 3 mm nodules, resulting in high false positive of detection task. The cut-off value of positive nodules needed to be further optimized. Whatever, the combination of mobile CT and deep learning model might be helpful in alleviating the weakness of facilities and experience in distant regions for lung cancer screening.

Deep learning algorithm has the potential to alter the clinical workflow of lung cancer [28,29]. At present, increasing number of studies have demonstrated the excellent application of deep learning in screening, diagnosis, and prognosis prediction of lung cancer. Previous study conducted an end-to-end deep learning algorithm to predict cancerous nodules on the basis of 6716 NLST cases, the performance of which was on-par with the radiologists [15]. Our study achieved a state-of-the-art performance on nodule detection with recall of 0.9507 and risk classification with ACC of 0.8696 in prospective large populations. Furthermore, it was possible to determine adenocarcinoma subtypes, gene mutation status and prognosis based on non-invasive CT images, reforming the selection of treatment strategies [30–32]. Beyond gains in consistency and accuracy, the capacity of deep learning to leverage diverse information has become of prime importance in improving efficiency of lung cancer management. Importantly, clinical adoption of these tools required further verification in external dataset to improve generalizability and effectiveness [33].

There were still several limitations in this study. First, some patients with high-risk pulmonary nodules lost visits or were not diagnosed pathologically as of the follow-up time, bringing unavoidable biases. Secondly, the features of nodules such as size, morphology, and growth were of paramount for risk stratification, but lacked in the predictive model. Last but not least, this study was the result of 1-year follow-up of cohort, 2-year or longer follow-up will make the results more convincing. We would continue to manage these participants, and validate accurate lung cancer screening model.

5. Conclusions

In conclusion, these results represented a large-scale prospectively screening study on mobile CT and an automated system to evaluate pulmonary nodules and lung cancer malignancy through deep learning. The novel approach in medical applications may assist clinicians to facilitate early diagnosis of lung cancer effectively, especially in resource-constrained sites.

Author Contributions

WML, ZY, DL and BJC designed the research study. JS, CDW, and TBD performed the research. LY, TZL, XYX, GW and JXG analyzed the data. JS, LY and CDW

wrote the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

Ethics Approval and Consent to Participate

Ethics approval was obtained from the ethics committee of West China Hospital (Project.2020(232)).

Acknowledgment

We thank all of the volunteers who participated in the study; and staff in Mianzhu people's Hospital, the First People's Hospital of Longquan District, Chengdu, as well as West China Hospital of Sichuan University, who were active in the whole process of the screening.

Funding

This research was funded by the National Natural Science Foundation of China (92159302, 82100119), the Science and Technology Project of Sichuan (2022ZDZX0018, 2020YFG0473), Science and Technology Innovation Project of Guang'an (2020SYF03), and the 1·3·5 Project for Disciplines of Excellence, West China Hospital. Sichuan University (ZYJC18001).

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al*. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians. 2021; 71: 209–249.

[2] Zeng H, Chen W, Zheng R, Zhang S, Ji JS, Zou X, *et al*. Changing Cancer Survival in China During 2003-15: A Pooled Analysis of 17 Population-Based Cancer Registries. The Lancet. Global Health. 2018; 6: e555–e567.

[3] Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, *et al*. Cancer Treatment and Survivorship Statistics, 2019. CA: A Cancer Journal for Clinicians. 2019; 69: 363–385.

[4] Zeng H, Ran X, An L, Zheng R, Zhang S, Ji JS, *et al*. Disparities in Stage at Diagnosis for Five Common Cancers in China: A Multicentre, Hospital-Based, Observational Study. The Lancet Public Health. 2021; 6: e877–e887.

[5] Oudkerk M, Liu S, Heuvelmans MA, Walter JE, Field JK. Lung Cancer LDCT Screening and Mortality Reduction — Evidence, Pitfalls and Future Perspectives. Nature Reviews Clinical Oncology. 2021; 18: 135–151.

[6] Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, *et al*. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. The New England Journal of Medicine. 2011; 365: 395–409.

[7] de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, *et al*. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. New England Journal of Medicine. 2020; 382: 503–513.

[8] Black WC, Gareen IF, Soneji SS, Sicks JD, Keeler EB, Aberle DR, *et al*. Cost-Effectiveness of CT Screening in the National Lung Screening Trial. The New England Journal of Medicine. 2014; 371: 1793–1802.

[9] Criss SD, Cao P, Bastani M, ten Haaf K, Chen Y, Sheehan DF, *et al*. Cost-Effectiveness Analysis of Lung Cancer Screening in the United States. Annals of Internal Medicine. 2019; 171: 796.

[10] Mehta HJ, Mohammed T, Jantz MA. The American College of Radiology Lung Imaging Reporting and Data System: Potential Drawbacks and need for Revision. Chest. 2017; 151: 539–543.

[11] Rivera MP, Katki HA, Tanner NT, Triplette M, Sakoda LC, Wiener RS, *et al*. Addressing Disparities in Lung Cancer Screening Eligibility and Healthcare Access. an Official American Thoracic Society Statement. American Journal of Respiratory and Critical Care Medicine. 2020; 202: e95–e112.

[12] Dhoot R, Humphrey JM, O'Meara P, Gardner A, McDonald CJ, Ogot K, *et al*. Implementing a Mobile Diagnostic Unit to Increase Access to Imaging and Laboratory Services in Western Kenya. BMJ Global Health. 2018; 3: e000947.

[13] Crosbie PA, Gabe R, Simmonds I, Kennedy M, Rogerson S, Ahmed N, *et al*. Yorkshire Lung Screening Trial (YLST): Protocol for a Randomised Controlled Trial to Evaluate Invitation to Community-Based Low-Dose CT Screening for Lung Cancer Versus Usual Care in A Targeted Population at Risk. BMJ Open. 2020; 10: e037075.

[14] Zhou Y, Xu X, Song L, Wang C, Guo J, Yi Z, *et al*. The Application of Artificial Intelligence and Radiomics in Lung Cancer. Precision Clinical Medicine. 2020; 3: 214–227.

[15] Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, *et al*. End-to-End Lung Cancer Screening with Three-Dimensional Deep Learning on Low-Dose Chest Computed Tomography. Nature Medicine. 2019; 25: 954–961.

[16] Zhang K, Liu X, Shen J, Li Z, Sang Y, Wu X, *et al*. Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. Cell. 2020; 181: 1423–1433.e11.

[17] Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, *et al*. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. The Journal of the American Medical Association. 2017; 318: 2199–2210.

[18] Wang C, Ma J, Shao J, Zhang S, Liu Z, Yu Y, *et al*. Predicting EGFR and PD-L1 Status in NSCLC Patients Using Multitask AI System Based on CT Images. Frontiers in Immunology. 2022; 13: 813072

[19] Wang C, Xu X, Shao J, Zhou K, Zhao K, He Y, *et al*. Deep Learning to Predict EGFR Mutation and PD-L1 Expression Status in Non-Small-Cell Lung Cancer on Computed Tomography Images. Journal of Oncology. 2021; 2021: 5499385.

[20] Xu X, Wang C, Guo J, Yang L, Bai H, Li W, *et al*. DeepLN: A Framework for Automatic Lung Nodule Detection Using Multi-Resolution CT Screening Images. Knowledge-Based Systems. 2020; 189: 105128.

[21] Pinsky PF, Gierada DS, Black W, Munden R, Nath H, Aberle D, *et al*. Performance of Lung-RADS in the National Lung Screening Trial: a retrospective assessment. Annals of Internal Medicine. 2015; 162: 485–491.

[22] Gao S, Li N, Wang S, Zhang F, Wei W, Li N, *et al*. Lung Cancer in People's Republic of China. Journal of Thoracic Oncology. 2020; 15: 1567–1576.

[23] Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman JR, Bharat A, *et al*. NCCN Guidelines Insights: Non–Small Cell Lung Cancer, Version 2.2021. Journal of the National Comprehensive Cancer Network. 2021; 19: 254–266.

[24] Field JK, Duffy SW, Baldwin DR, Whynes DK, Devaraj A, Brain KE, *et al*. UK Lung Cancer RCT Pilot Screening Trial: Baseline Findings from The Screening Arm Provide Evidence for the Potential Implementation of Lung Cancer Screening. Thorax. 2016; 71: 161–170.

[25] Walter JE, Heuvelmans MA, de Jong PA, Vliegenthart R, van

Ooijen PMA, Peters RB, *et al*. Occurrence and Lung Cancer Probability of New Solid Nodules at Incidence Screening with Low-Dose CT: Analysis of Data from The Randomised, Controlled NELSON Trial. Oncology. 2016; 17: 907–916.

[26] Cui S, Ming S, Lin Y, Chen F, Shen Q, Li H, *et al*. Development and Clinical Application of Deep Learning Model for Lung Nodules Screening on CT Images. Scientific Reports. 2020; 10: 13657.

[27] Zhang S, Sun K, Zheng R, Zeng H, Wang S, Chen R, *et al*. Cancer Incidence and Mortality in China, 2015. Journal of the National Cancer Center. 2021; 1: 2–11.

[28] Kann BH, Hosny A, Aerts HJWL. Artificial intelligence for clinical oncology. Cancer Cell. 2021; 39: 916–927.

[29] Huynh E, Hosny A, Guthier C, Bitterman DS, Petit SF, Haas-Kogan DA, *et al*. Artificial Intelligence in Radiation Oncology. Nature Reviews Clinical Oncology. 2020; 17: 771–781.

[30] Wang C, Shao J, Lv J, Cao Y, Zhu C, Li J, *et al*. Deep Learning for Predicting Subtype Classification and Survival of Lung Adenocarcinoma on Computed Tomography. Translational Oncology. 2021; 14: 101141.

[31] Arbour KC, Luu AT, Luo J, Rizvi H, Plodkowski AJ, Sakhi M, *et al*. Deep Learning to Estimate RECIST in Patients with NSCLC Treated with PD-1 Blockade. Cancer Discovery. 2021; 11: 59–67.

[32] Tian P, He B, Mu W, Liu K, Liu L, Zeng H, *et al*. Assessing PD-L1 Expression in Non-Small Cell Lung Cancer and Predicting Responses to Immune Checkpoint Inhibitors Using Deep Learning on Computed Tomography Images. Theranostics. 2021; 11: 2098–2107.

[33] Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, *et al*. A Guide to Deep Learning in Healthcare. Nature Medicine. 2019; 25: 24–29.