PERSPECTIVE OPEN

China Protocol for early screening, precise diagnosis, and individualized treatment of lung cancer

Chengdi Wang [1,2 ✉], Bojiang Chen[1], Shufan Liang[1], Jun Shao[1], Jingwei Li[1], Liuqing Yang[1], Pengwei Ren[1], Zhoufeng Wang[1], Wenxin Luo[1], Li Zhang[1], Dan Liu[1] and Weimin Li [1,2 ✉]

Early screening, diagnosis, and treatment of lung cancer are pivotal in clinical practice since the tumor stage remains the most dominant factor that affects patient survival. Previous initiatives have tried to develop new tools for decision-making of lung cancer. In this study, we proposed the China Protocol, a complete workflow of lung cancer tailored to the Chinese population, which is implemented by steps including early screening by evaluation of risk factors and three-dimensional thin-layer image reconstruction technique for low-dose computed tomography (Tre-LDCT), accurate diagnosis *via* artificial intelligence (AI) and novel biomarkers, and individualized treatment through non-invasive molecule visualization strategies. The application of this protocol has improved the early diagnosis and 5-year survival rates of lung cancer in China. The proportion of early-stage (stage I) lung cancer has increased from 46.3% to 65.6%, along with a 5-year survival rate of 90.4%. Moreover, especially for stage IA1 lung cancer, the diagnosis rate has improved from 16% to 27.9%; meanwhile, the 5-year survival rate of this group achieved 97.5%. Thus, here we defined stage IA1 lung cancer, which cohort benefits significantly from early diagnosis and treatment, as the "ultra-early stage lung cancer", aiming to provide an intuitive description for more precise management and survival improvement. In the future, we will promote our findings to multicenter remote areas through medical alliances and mobile health services with the desire to move forward the diagnosis and treatment of lung cancer.

INTRODUCTION

Lung cancer is the leading cause of incidence and mortality related to malignancy worldwide, resulting in over 2.48 million new cases and 1.8 million deaths per year.[1] In China, lung cancer has also emerged as the top formidable health challenge and it is estimated that there are 1.06 million people were diagnosed and 0.73 million people died from lung cancer annually.[2] The tumor stage serves as the most important factor related to patient survival, as the 5-year survival rate of stage IA patients is about 82%, while the ratio decreases to 7% in stage IVB patients.[3] Thus, to sufficiently improve the overall outcome of lung cancer, early effective screening, precise diagnosis, and optimal individualized treatment are vitally important.

In early-stage lung cancer, few obvious symptoms can be recognized, leading to diagnosis at advanced stages, which hampers prompt intervention.[4] Thus, identifying patients at high risk of lung cancer, especially screening in asymptomatic cases, is critical. Two randomized controlled trials, the National Lung Screening Trial (NLST) and Nederlands-Leuvens Longkanker Screenings Onderzoek (NELSON) trial, have provided evidence that low-dose computed tomography (LDCT) screening reduces the mortality of lung cancer.[5,6] However, along with the increased detection of pulmonary nodules, it's essential to accurately discriminate the malignant nodules from benign lesions to give timely intervention of lung cancer and reduce the

overdiagnosis and overtreatment of benign nodules.[7] Moreover, current lung cancer screening guidelines indeed decreases the death rate, but the problem of missed diagnosis in specific population caused by international criteria remains to be solved.[8] Thus, establishing a screening approach suitable for the Chinese population is warranted.[9] In recent years, artificial intelligence (AI) has developed rapidly, and the advent of large language models (LLMs) represents a new wave in this field, which inspires deeper human-computer interaction.[10–12] The sustaining enthusiasm of attempting to integrate AI and medicine also provides novel tools for lung cancer care, including screening, diagnosis, prognosis, and so on.[13–15] Nevertheless, the limited interpretability and generalization may hinder real-world practicality, requiring technological innovation and transformation to improve the usability of AI systems.[16]

With the purpose of improving the detection rate of early-stage lung cancer and assisting optimal intervention, we proposed the China Protocol with realistic applicable potential which is composed of four key steps: screening by early identification of high-risk population and three-dimensional reconstruction CT examination, diagnosis by AI and novel biomarkers, treatment through non-invasive molecular characterization approaches, and whole-process management (Fig. 1). Through the implementation of China Protocol, we aim to diagnose lung cancer in earlier

[1]Department of Pulmonary and Critical Care Medicine, State Key Laboratory of Respiratory Health and Multimorbidity, Targeted Tracer Research and Development Laboratory, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, West China School of Medicine, Sichuan University, Chengdu, China and [2]Frontiers Medical Center, Tianfu Jincheng Laboratory, Chengdu, China
Correspondence: Chengdi Wang (chengdi_wang@scu.edu.cn) or Weimin Li (weimi003@scu.edu.cn)
These authors contributed equally: Chengdi Wang, Bojiang Chen, Shufan Liang.

China Protocol for early screening, precise diagnosis, and individualized…
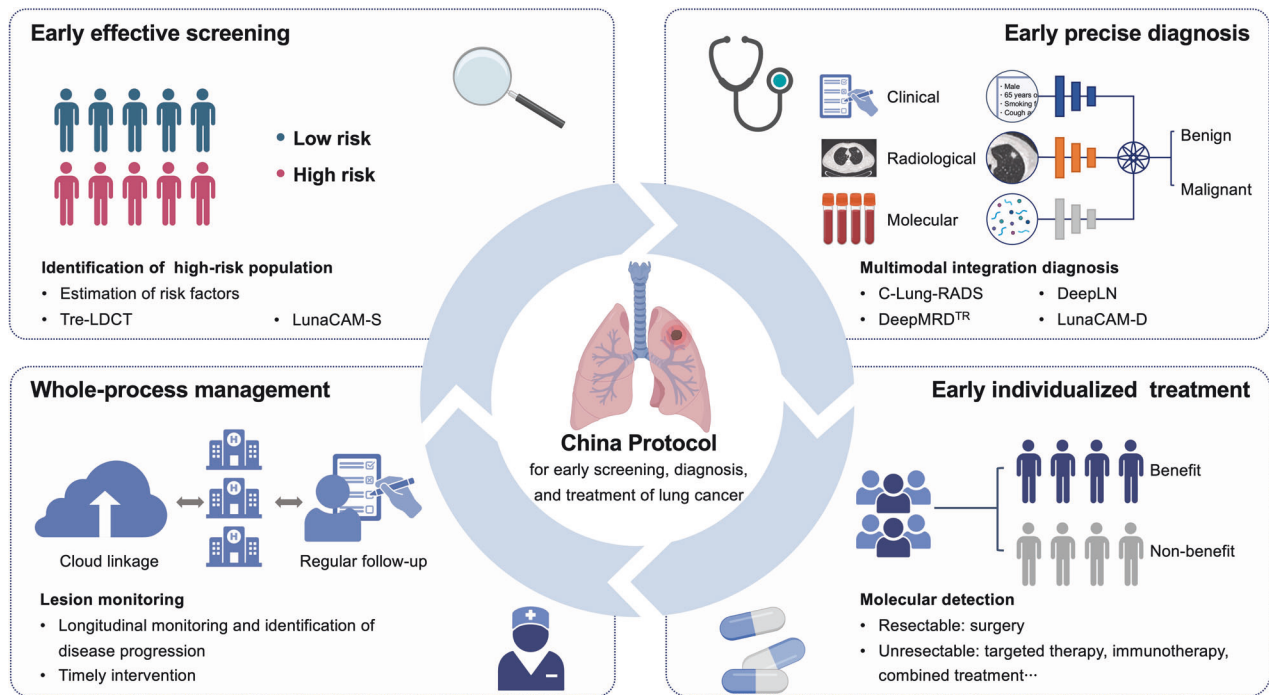Wang et al.

2

Fig. 1 Overview of China Protocol for lung cancer management. The China Protocol consists of screening, diagnosis, treatment, and whole-process management of lung cancer. Through estimation of risk factors, the high-risk cohort can be identified and undergo low-dose computed tomography (LDCT) examination or biomarker-based screening by lung cancer alertness by ctDNA methylation (LunaCAM)-S. Then, the artificial intelligence (AI)-supported systems assist in lung cancer diagnosis, and the diagnostic biomarkers can further improve the accuracy. After diagnosis, the molecular characterization helps to provide precise treatment plans. In addition, the regular follow-up can identify the disease progression and facilitate timely intervention. C-Lung-RADS Chinese Lung Nodules Reporting and Data System, Tre-LDCT three-dimensional thin-layer image reconstruction technique for low-dose computed tomography

stages, especially in the "ultra-early stage", which is described first and essentially refers to stage IA1 lung cancer, as well as improving the overall survival.

NOVEL EFFECTIVE SCREENING APPROACHES OF LUNG CANCER AND PRACTICE IN WESTERN CHINA

The determination of the population at high risk of lung cancer faces the challenge of varied definitions in terms of the same risk factors in different regions, leading to potential missed diagnosis by non-customized screening standards.[17] Therefore, we studied the risk factors in western China population to accurately identify the high-risk group and the results indicated that lung cancer more frequently occurs in the group of age over 40 years old, with smoking history, chronic obstructive pulmonary disease (COPD), diffuse pulmonary fibrosis, and previous history of malignancy.[18] Then, the selected high-risk patients can undergo the CT examinations. To solve the missed diagnosis of subtle lesions due to the coarse spatial resolution of thick-slice CT, we developed a novel three-dimensional thin-layer image reconstruction technique for LDCT (Tre-LDCT) to effectively detect suspicious lesions and help discover lung cancer. The Tre-LDCT achieved efficient detection of pulmonary nodules with a sensitivity around 96%, along with reconstruction and visualization of the lesions to assist in radiological interpretation.[19] Subsequently, to further determine the high-risk nodules detected on CT, we meticulously crafted an assessment tool consisting of nine factors, such as age, smoking, family history of malignancy, nodule diameter, spiculation, lobulation, and mixed ground-glass opacity (mGGO) density, were associated with lung cancer.[20] Moreover, to aid timely and precise lung cancer management, we first defined the stage IA1 lung cancer as "ultra-early stage lung cancer", a more intuitive description which refers to lung cancer that is in the initial developing stage, with

nodules with diameters no more than 1 cm, localized, without lymph node or distant metastasis, and possessing better prognosis. In real-world validation, the thin-slice LDCT has improved lung cancer screening efficiency especially for early-stage (stage I) and ultra-early stage lung cancer.[21]

DEVELOPMENT AND APPLICATION OF PRECISE DIAGNOSIS APPROACHES FOR LUNG CANCER

Although the screening approaches have facilitated the detection of pulmonary nodules, it is challenging to make accurate disease determination of the identified nodules, which may result in misdiagnosis. Moreover, the pathological diagnosis relies upon invasive tissue sampling, which is inconvenient and may cause complications.[22] Thus, previous research has tried to realize non-invasive and automated diagnosis by AI-enabled image decoding.[23,24] The efficiency of commercial computer-aided detection (CAD) systems has been validated.[25] We have explored the capability of AI in processing and analyzing multidimensional medical data for information extraction, image standardization, lesion segmentation, representation fusion, and so on.[26–29] And the technological innovation provides robust tools for lung cancer diagnosis, as elaborated as follows.

Firstly, we have established a respiratory disease database involving lung cancer and other pulmonary diseases, which serves as the foundation of model construction, consisting of radiological images, laboratory test results, and other clinical text data from 434,735 patients.[30,31] Then, we developed a two-stage segmentation framework to accurately delineate the lung regions and lesions using the backbone of DeepLabv3, obtaining clearer boundaries than human experts.[26] Subsequently, to capture the radiological characteristics of the segmented pulmonary nodules from CT images, a multi-scale cost-sensitive neural network which

China Protocol for early screening, precise diagnosis, and individualized…
Wang et al.

3

ensembled three light models training on CT patches cropped in different sizes was proposed to extract the key features that were of diagnostic values, such as signs of lobulation, spicule, and cavity.[32] Also, we provided a Semi-Supervised Medical image Detector (SSMD) for effective supervision of unlabeled data, which could reduce the labor of manual annotating and generate standardized labels for model training.[33] In addition, we innovatively proposed a unified feature fusion method which integrated multimodal data, including demographics, chief complaints, laboratory results, and medical images, to improve the model performance. By leveraging transformer as the backbone and bringing in the bidirectional multimodal attention, the performance of this approach surpassed the image-only and the non-unified methods.[28]

Based on the abovementioned algorithm innovation, we then developed lung cancer diagnosis models. For sub-centimeter pulmonary nodules, baseline plus follow-up lesion analysis achieved a diagnostic area under the curve (AUC) of 0.942 and was comparable to the senior clinician.[34] Then, for pulmonary nodules detected with more diverse characteristics, we established a multi-dimensional stepwise lung cancer risk assessment model, Chinese Lung Nodules Reporting and Data System (C-Lung-RADS), which stratified the nodules into low-, mid-, high-, and extremely high-risk. And the system demonstrated a higher sensitivity in lung cancer diagnosis, surpassing the current international benchmark Lung CT Screening Reporting and Data System (Lung-RADS) v2022 (87.1% vs. 63.3%). Then, personalized plans are made to avoid overdiagnosis of low-risk cases, as well as alleviating missed diagnosis of high-risk population.[35] We also designed a semi-automated framework, named DeepLN, to identify lung cancer by deep neural networks on CT images, showing an accuracy of 92.46% in discrimination between benign and malignant lesions.[36] Moreover, to address the diagnostic challenge resulted by the shared radiological abnormalities in different pulmonary diseases, DeepMRD[TR] was designed to precisely discriminate lung cancer from pneumonia, tuberculosis, interstitial lung disease, and so on.[30] Importantly, for clinical application, we have developed an AI-powered product for malignancy risk evaluation and diagnosis of pulmonary nodules, and the AI system-assisted approach missed less ultra-early stage lung cancer than manual reading (1.5% vs. 10.6%).[37]

DISCOVERY AND VALIDATION OF BIOMARKERS FOR EARLY-STAGE LUNG CANCER

Currently, traditional biomarkers for lung cancer identification, such as carcinoembryonic antigen (CEA), cancer antigen 125 (CA125), and cytokeratin-19 fragment (CYFRA 21-1) are commonly utilized.[38,39] However, the limited performance in the early-stage lung cancer with sensitivity around 0.6 hindered accurate detection, consequently raising the necessity of exploring novel biomarkers to offer new insights into the early diagnosis of lung cancer.[40]

Comprehensive deciphering the mechanisms of lung cancer carcinogenesis is crucial to find potential diagnostic targets. The
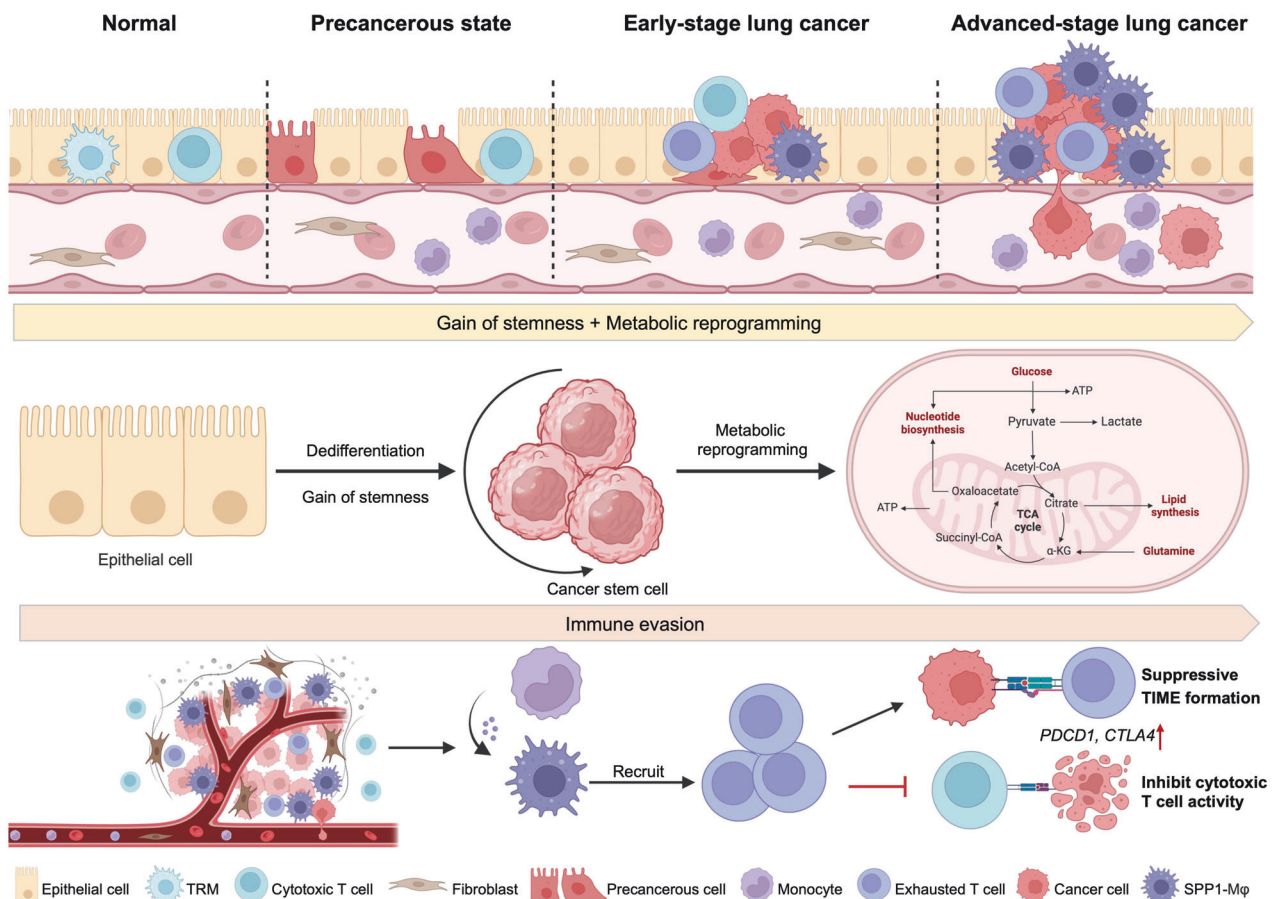

Fig. 2 Multi-stage tumorigenesis and progression of lung cancer. The progression of lung cancer is a multifaceted biological process including various stages and steps. During the progression, epithelial cells acquire stemness and undergo metabolic reprogramming. Concurrently, there is a progressive increase in exhausted T cells and SPP1-macrophages. Furthermore, our findings indicate that the multidimensional communication between lung stem cells and immune cells plays a crucial role in sustaining cellular stemness and facilitating tumor immune evasion, thereby promoting the development of lung cancer. TIME, tumor immune microenvironment; TRM, tissue-resident macrophage

4

occurrence and development of lung cancer are complex multi-stage processes that may be affected by molecular aberrations of cancer cells and dynamic regulation of tumor microenvironment (TME).[41] We proposed the "seed-and-soil" hypothesis underpinning lung cancer pathogenesis and found that epithelial cells, such as AT2 and basal cells, serve as the "seeds". Gradual dedifferentiation and enhanced stemness of these cells are critical steps during lung cancer evolution. And the marker genes, including midkine (*MDK*) and tissue inhibitor matrix metalloproteinase 1 (*TIMP1*), which represent the process of metabolism regulation, are overexpressed, serving as potential biomarkers.[42,43] Through high-throughput multiomics analysis, we then found that hypomethylation and high expression of *PRAME* can promote the progression of early-stage lung cancer by regulating epithelial-to-mesenchymal transition (EMT)-related genes.[44] Furthermore, the tumor immune microenvironment (TIME) forms the "soil" for lung cancer development. Specifically, macrophages play a crucial role in the early transformation of lung cancer. The hypomethylated and upregulated *PRAME* can also promote the recruitment of monocytes to tumor tissues and differentiation into tumor-associated macrophages (TAMs), including SPP1 macrophages. Subsequently, SPP1 macrophages secrete migration inhibitory factor (MIF) to inhibit the cytotoxic effect of CD8 + T cells, resulting in immunosuppression. These findings reveal a novel mechanism concerning immune escape of lung cancer.[45,46] Moreover, we proposed the "epithelium-immunocyte multidimensional cell-cell communication" theory and found that SPP1 macrophages interact with lung stem cells and T cells through SPP1-CD44 and MIF-CD74, thereby maintaining cell stemness and leading to a suppressed TIME, thus promoting the progression of lung cancer (Fig. 2).[45–47]

Novel liquid biopsies have attracted much attention in lung cancer detection benefiting from the less invasiveness and feasible sampling, especially better accuracies than conventional biomarkers.[48,49] Based on various molecules such as circulating tumor DNA (ctDNA) and cell-free DNA (cfDNA), the liquid biopsy methods have shown potential in lung cancer management from early diagnosis, treatment monitoring, to prognosis prediction.[50–52] We selected 11 DNA methylation markers, including *CDO1*, *GSHR*, *HOXA11*, etc., and validated their diagnostic abilities of lung cancer through bronchoalveolar lavage fluid (BALF) samples.[53] Then, to further explore the detection capability of DNA methylation for early-stage lung cancer, two lung cancer alertness by ctDNA methylation (LunaCAM) models, LunaCAM-S and LunaCAM-D, were separately built on 7 and 6 methylation markers, such as *HOXD9* and *SHOX2*, respectively targeting in the discrimination of lung cancer from healthy population and patients with benign pulmonary diseases. The two models finally achieved AUCs of 0.90 and 0.81 for lung cancer screening and diagnosis. In particular, the LunaCAM-S showed stable performance in different stages of lung cancer and detected 86% tumors less than 1.2 cm, which implied that the model may be capable of catching the malignant information in early stage. Moreover, the LunaCAM-D substantially surpassed the CEA test (AUC: 0.56) for diagnosis.[54]

As for non-coding RNAs (ncRNAs), in lung adenocarcinoma (LUAD), the hsa_circ_0072309 (circLIFR) was downregulated and exhibited tumor inhibition function by impressing the cell migration.[55] Furthermore, we implemented circRNA sequencing in both LUAD and lung squamous cell carcinoma (LUSC), as well as matched normal tissues, then selected four circRNAs of high abundance, namely hsa_circ_0001073, hsa_circ_0001495, hsa_-circ_0077837, and hsa_circ_0001821. The hsa_circ_0001073 and hsa_circ_0001495 showed altered expressions in LUAD and LUSC, and the two molecules presented subtyping potential with AUCs of 0.919 and 0.965 for LUAD or LUSC identification.[56] To further reduce overdiagnosis, it is warranted to establish the validated thresholds to distinguish the indolent lesions from those demonstrate aggressive biological behavior and require immediate intervention.

MOLECULAR CHARACTERIZATION TO GUIDE INDIVIDUALIZED TREATMENT

For resectable early-stage lung cancer, surgical operation serves as the key local treatment.[57] Nevertheless, for patients with unresectable lung cancer and positive molecular events such as driver gene mutations or immune checkpoint expressions, the treatment has entered the era of precision therapy based on molecular detection; however, there are drawbacks such as invasive procedures and tumor heterogeneity.[58–61] Thus, to tackle these issues, we explored the non-invasive molecular characterization approaches to guide the optimal therapy scheduling. We utilized whole-lung features on CT scans to predict epidermal growth factor receptor (*EGFR*) mutation, which method reached an AUC of 0.756.[62] In the same task, another intelligent system was subsequently constructed *via* a self-supervised multitask deep learning strategy, achieving an improved AUC of 0.824.[63] For discrimination of *EGFR* mutation subtypes, including 19Del, L858R, and others, combination of deep learning features, radiomics features, and clinical features obtained an AUC of 0.841 in this three-way classification task. And the heatmaps highlighted the radiology foundations behind the final outputs.[64] Furthermore, we quantitatively visualized the activation levels of EGFR in lung cancer patients through a specific molecular probe, HX103, to accurately identify the cohort who are likely to benefit from EGFR-tyrosine kinase inhibitor (EGFR-TKI) targeted therapy.[65] In addition, to overcome the challenge of EGFR-TKI resistance, we adopted next-generation sequencing (NGS)-based liquid biopsy to identify patients with T790M relative allele frequency (RAF) < 20% who were more likely to develop resistance after the first-generation EGFR-TKI treatment, providing a potential strategy for therapy planning and prognosis prediction.[66] On the basis of the yes-associated protein (YAP) which is confirmed as a critical driver of EGFR-TKI resistance and the emerging photodynamic therapy (PDT), we proposed a nanococktail therapy implemented by conjugates that were able to deliver gefitinib, YAP-siRNA, and a photosensitizer (Ppa) simultaneously, thus promoting apoptosis of cancer cells and showing a tumor inhibition rate of 86.7%.[67] Moreover, in terms of another driver gene of lung cancer, anaplastic lymphoma kinase (*ALK*), a capture-based targeted sequencing panel was applied to detect the rearrangement events in plasma, reaching a sensitivity of 79.2%, assisting more precise targeted treatment.[68]

Besides targeted therapy, for immunotherapy, a programmed death-ligand 1 (PD-L1) expression signature (PD-L1ES) was constructed by a deep convolutional neural network to predict the high PD-L1 expression (≥ 50%) by decoding the pretreatment CT volumes.[69] Then, the combination of deep learning, radiomics, and clinical features predicted the different expressions of PD-L1, namely, <1%, 1-49%, and ≥ 50%, with AUCs of 0.950, 0.934, and 0.946.[70] Furthermore, a multi-molecular assessment model was further developed to predict multiple events such as *EGFR* and Kirsten rat sarcoma viral oncogene homolog (*KRAS*) mutations, as well as PD-L1 expression, better aligning with the real-world clinical application scenarios and enabling the appropriate therapy planning.[71,72] Therefore, we defined the "medical imaging ecology" which refers to mining the radiological characteristics of the lesions and surrounded environment and subsequently predicting the biological events, as well as highlighting the interpretability to visualize the correlation between the macroscopic manifestation and underlying microscopic information. Therefore, candidates who may benefit from specific molecule-based treatment can be pre-emptively selected to better assist clinical decision-making.
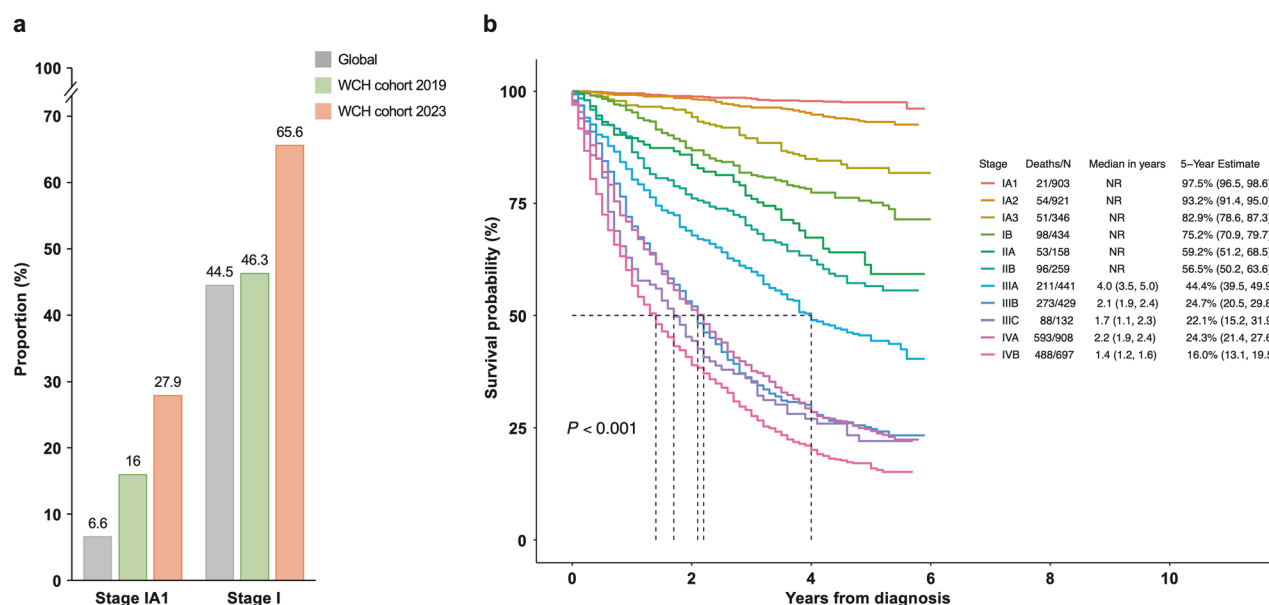
China Protocol for early screening, precise diagnosis, and individualized...
Wang et al.

5

Fig. 3 Application results of China Protocol in early diagnosis and survival improvement of lung cancer. **a** Proportion of stage IA1 and stage I lung cancer globally and in West China Hospital of Sichuan University (WCH) cohorts; **b** Kaplan-Meier survival curves of patients diagnosed with lung cancer in 2019. NR, not reached

APPLICATION OF CHINA PROTOCOL FOR EARLY DIAGNOSIS AND PRECISE TREATMENT OF LUNG CANCER

Through the continuous effort, we established the China Protocol, a complete pipeline containing screening, diagnosis, treatment, along with the whole-process management of lung cancer. It is especially suitable for people over 40 years old who are at high risk of lung cancer and are detected with pulmonary nodules on CT scans. Specifically, the high-risk population can be identified *via* estimation of the risk factors; then, Tre-LDCT is able to detect the existing pulmonary nodules from this group by CT examination. Once the nodules are found, different risk levels are determined using C-Lung-RADS. In addition, through individual-specific follow-up based on the trait evolution of pulmonary nodules, lung cancer can be identified as early as possible. Moreover, for the uncertain nodules that are challenging to diagnose by CT images, the novel biomarkers are capable of further improving diagnostic accuracy through combined analysis of the laboratory results and AI predictions. After lung cancer diagnosis, the non-invasive molecular characterization methods are utilized to guide individualized therapy.

The implementation of this protocol improved the early diagnosis and 5-year survival rates of lung cancer. And to detailedly quantify the application effect, we enrolled a total of 11,844 patients pathologically confirmed with lung cancer in 2019 and 2023 as examples (Supplementary Table 1). For patients diagnosed in 2019, follow-up data were obtained until December 2024 and the median follow-up time was 46.8 months (interquartile range [IQR], 12-60 months). Detailed clinical characteristics of this cohort were provided in Supplementary Table 2. As a result, the detection rate of early-stage lung cancer achieved 46.3% in 2019 and 65.6% in 2023, surpassing the global proportion of 44.5%. In ultra-early stage lung cancer, this rate has increased from 16% to 27.9%, higher than the international ratio of 6.6% (Fig. 3a).[3] As for survival, the timely diagnosis has improved the 5-year survival rate of lung cancer to 59%. For early-stage and ultra-early stage lung cancer, the rates reached 90.4% and 97.5%, respectively (Supplementary Table 3, Fig. 3b).

For further application and promotion, our dynamic database of patients with lung cancer not only provides solid data support for the continuous optimization of China Protocol but also enhances the generalization of this strategy in various populations.[30,35] Furthermore, the wide coverage and efficient implementation of China Protocol are facilitated by the application of advanced technologies, such as mobile CT, 5G communication technology, and Internet of Things (IoT). For instance, in resource-limited areas, patients can undergo lung cancer screening through vehicle-mounted CT and receive diagnosis from experts at tertiary hospitals through cloud computing platforms, improving the accessibility of China Protocol and providing patients in remote locations with equal access to advanced healthcare services.[73,74] And we have developed the intelligent product for lung cancer diagnosis, which can be applied in different medical institutions once assembled. Moreover, the biomarkers we developed can be detected by low-cost quantitative polymerase chain reaction (qPCR) assay which is practicable and convenient. These advantages make it possible to generalize this protocol widely.

Compared to other lung cancer management strategies, China Protocol was established on a large-scale database, and we redefined the high-risk population based on Chinese cohort and developed the novel Tre-LDCT technology to realize accurate stepwise screening. Then, we innovated AI algorithms for image processing and diverse feature fusion, as well as mining novel biomarkers, for precise diagnosis. Moreover, we explored non-invasive molecular characterization methods to guide the individual-specific treatment, more advanced and convenient than traditional approaches. Despite certain achievements that have been made, there are several limitations of this research. Since participants involved in our investigation were primarily from western China, prospective multi-ethnic multi-center validation is needed. Then, intelligent diagnosis based on the multi-dimensional fusion of medical images and molecules to generate more accurate clinical decisions is worthy of deep research. In addition, as the LLMs have become the frontier of AI, developing generative models is useful to enable human-computer interactive clinical workflow of lung cancer.

In summary, China Protocol has yielded achievements in early screening, precise diagnosis, and individualized treatment of lung cancer, opening a new chapter for this disease management. In the future, more technological innovation and strategy optimization suitable for real-world settings are promising to revolutionize the clinical paradigm of lung cancer, thus improving the patient care.

China Protocol for early screening, precise diagnosis, and individualized…
Wang et al.

6

DATA AVAILABILITY

All the data are available from the corresponding authors upon reasonable request.

AUTHOR CONTRIBUTIONS

W.L. and C.W. conceived the idea and designed the experiments. C.W., B.C., S.L., J.S., J.L., L.Y., P.R., Z.W., W.L., L.Z., and D.L. contributed to data collection and statistics and drafted the original version of the manuscript. C.W., B.C., S.L., J.S., and J.L. designed the tables and figures. C.W. and W.L. revised the final submission. All the authors reviewed, edited, and approved the manuscript.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41392-025-02256-1.

Competing interests: Weimin Li is one of the Associate Editors of Signal Transduction and Targeted Therapy, but he has not been involved in the process of the manuscript handling. Other authors declare no conflict of interest.

Ethics declarations: The study was approved by the Ethics Committee of West China Hospital, Sichuan University (Approval Nos. 2020.232 and 2023.2287) and the requirement for informed consent was waived.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **74**, 229–263 (2024).
2. Han, B. et al. Cancer incidence and mortality in China, 2022. *J. Natl. Cancer Cent.* **4**, 47–53 (2024).
3. Rami-Porta, R. et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: proposals for revision of the TNM stage groups in the forthcoming (ninth) edition of the TNM classification for lung cancer. *J. Thorac. Oncol.* **19**, 1007–1027 (2024).
4. Crosby, D. et al. Early detection of cancer. *Science* **375**, eaay9040 (2022).
5. Aberle, D. R. et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N. Engl. J. Med.* **365**, 395–409 (2011).
6. de Koning, H. J. et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N. Engl. J. Med.* **382**, 503–513 (2020).
7. Mazzone, P. J. & Lam, L. Evaluating the patient with a pulmonary nodule: a review. *JAMA* **327**, 264–273 (2022).
8. Adams, S. J. et al. Lung cancer screening. *Lancet* **401**, 390–408 (2023).
9. Cheng, Y. et al. Implementation planning for lung cancer screening in China. *Precis Clin. Med.* **2**, 13–44 (2019).
10. Lu, M. Y. et al. AI-based pathology predicts origins for cancers of unknown primary. *Nature* **594**, 106–110 (2021).
11. Lu, M. Y. et al. A multimodal generative AI copilot for human pathology. *Nature* **634**, 466–473 (2024).
12. Chen, K., Li, J. & Li, L. Artificial intelligence for disease X: progress and challenges. *J. Transl. Int. Med.* **12**, 534–536 (2024).
13. Ardila, D. et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* **25**, 954–961 (2019).
14. Wang, X. et al. A pathology foundation model for cancer diagnosis and prognosis prediction. *Nature* **634**, 970–978 (2024).
15. Shao, J. et al. Novel tools for early diagnosis and precision treatment based on artificial intelligence. *Chin. Med. J. Pulm. Crit. Care Med.* **1**, 148–160 (2023).
16. Huang, S. et al. Artificial intelligence in lung cancer diagnosis and prognosis: current application and future perspective. *Semin Cancer Biol.* **89**, 30–37 (2023).
17. Toumazis, I., Bastani, M., Han, S. S. & Plevritis, S. K. Risk-based lung cancer screening: a systematic review. *Lung Cancer* **147**, 154–186 (2020).
18. Ji, G. et al. Current lung cancer screening guidelines may miss high-risk population: a real-world study. *BMC Cancer* **21**, 50 (2021).
19. Xu, X. et al. DeepLN: a framework for automatic lung nodule detection using multi-resolution CT screening images. *Knowl. Based Syst.* **189**, 105128 (2020).
20. Zhu, Y. et al. Factors associated with concurrent malignancy risk among patients with incidental solitary pulmonary nodule: a systematic review taskforce for developing rapid recommendations. *J. Evid. Based Med.* **15**, 106–122 (2022).
21. Wu, J. et al. Screening for lung cancer using thin-slice low-dose computed tomography in southwestern China: a population-based real-world study. *Thorac. Cancer* **15**, 1522–1532 (2024).
22. Jonas, D. E. et al. Screening for lung cancer with low-dose computed tomography: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA* **325**, 971–987 (2021).
23. Mikhael, P. G. et al. Sybil: a validated deep learning model to predict future lung cancer risk from a single low-dose chest computed tomography. *J. Clin. Oncol.* **41**, 2191–2200 (2023).
24. Huang, P. et al. Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study. *Radiology* **286**, 286–295 (2018).
25. Paramasamy, J. et al. Validation of a commercially available CAD-system for lung nodule detection and characterization using CT-scans. *Eur. Radio.* **35**, 1076–1088 (2025).
26. Zhang, K. et al. Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography. *Cell* **181**, 1423–1433.e1411 (2020).
27. Wang, G. et al. A deep-learning pipeline for the diagnosis and discrimination of viral, non-viral and COVID-19 pneumonia from chest X-ray images. *Nat. Biomed. Eng.* **5**, 509–521 (2021).
28. Zhou, H. Y. et al. A transformer-based representation-learning model with unified processing of multimodal input for clinical diagnostics. *Nat. Biomed. Eng.* **7**, 743–755 (2023).
29. Shao, J. et al. A multimodal integration pipeline for accurate diagnosis, pathogen identification, and prognosis prediction of pulmonary infections. *Innovation* **5**, 100648 (2024).
30. Wang, C. et al. Development and validation of an abnormality-derived deep-learning diagnostic system for major respiratory diseases. *npj Digit Med.* **5**, 124 (2022).
31. Wang, C. et al. Persistent increase and improved survival of stage I lung cancer based on a large-scale real-world sample of 26,226 cases. *Chin. Med. J.* **136**, 1937–1948 (2023).
32. Xu, X. et al. MSCS-DeepLN: evaluating lung nodule malignancy using multi-scale cost-sensitive neural networks. *Med Image Anal.* **65**, 101772 (2020).
33. Zhou, H. Y. et al. SSMD: Semi-Supervised medical image detection with adaptive consistency and heterogeneous perturbation. *Med Image Anal.* **72**, 102117 (2021).
34. Zhang, R. et al. Deep learning for malignancy risk estimation of incidental sub-centimeter pulmonary nodules on CT images. *Eur. Radio.* **34**, 4218–4229 (2024).
35. Wang, C. et al. Data-driven risk stratification and precision management of pulmonary nodules detected on chest computed tomography. *Nat. Med.* **30**, 3184–3195 (2024).
36. Guo, J. et al. DeepLN: an artificial intelligence-based automated system for lung cancer screening. *Ann. Transl. Med.* **8**, 1126 (2020).
37. Wu, J. et al. Application of artificial intelligence in lung cancer screening: a real-world study in a Chinese physical examination population. *Thorac. Cancer* **15**, 2061–2072 (2024).
38. Patz, E. F. Jr. et al. Panel of serum biomarkers for the diagnosis of lung cancer. *J. Clin. Oncol.* **25**, 5578–5583 (2007).
39. Bigbee, W. L. et al. A multiplexed serum biomarker immunoassay panel discriminates clinical lung cancer patients from high-risk individuals found to be cancer-free by CT screening. *J. Thorac. Oncol.* **7**, 698–708 (2012).
40. Guida, F. et al. Assessment of lung cancer risk on the basis of a biomarker panel of circulating proteins. *JAMA Oncol.* **4**, e182078 (2018).
41. Weeden, C. E. et al. Impact of risk factors on early cancer evolution. *Cell* **186**, 1541–1563 (2023).
42. Wang, Z. et al. Deciphering cell lineage specification of human lung adenocarcinoma with single-cell RNA sequencing. *Nat. Commun.* **12**, 6500 (2021).
43. Zhang, L. et al. Integrated single-cell RNA sequencing analysis reveals distinct cellular and transcriptional modules associated with survival in lung cancer. *Signal Transduct. Target Ther.* **7**, 9 (2022).
44. Wang, C. et al. Multi-omics analyses reveal biological and clinical insights in recurrent stage I non-small cell lung cancer. *Nat. Commun.* **16**, 1477 (2025).
45. Luo, W. et al. Distinct immune microenvironment of lung adenocarcinoma in never-smokers from smokers. *Cell Rep. Med.* **4**, 101078 (2023).

China Protocol for early screening, precise diagnosis, and individualized...
Wang et al.

7

46. Wang, C. et al. The heterogeneous immune landscape between lung adeno-carcinoma and squamous carcinoma revealed by single-cell RNA sequencing. *Signal Transduct. Target Ther.* **7**, 289 (2022).

47. Wang, D. et al. Single-cell transcriptome analysis deciphers the CD74-mediated immune evasion and tumour growth in lung squamous cell carci-noma with chronic obstructive pulmonary disease. *Clin. Transl. Med.* **14**, e1786 (2024).

48. Nikanjam, M., Kato, S. & Kurzrock, R. Liquid biopsy: current technology and clinical applications. *J. Hematol. Oncol.* **15**, 131 (2022).

49. Bibikova, M. & Fan, J. Liquid biopsy for early detection of lung cancer. *Chin. Med J. Pulm. Crit. Care Med.* **1**, 200–206 (2023).

50. Li, C. et al. Circulating tumor DNA as liquid biopsy in lung cancer: biological characteristics and clinical integration. *Cancer Lett.* **577**, 216365 (2023).

51. Chabon, J. J. et al. Integrating genomic features for non-invasive early lung cancer detection. *Nature* **580**, 245–251 (2020).

52. Ren, F. et al. Liquid biopsy techniques and lung cancer: diagnosis, monitoring and evaluation. *J. Exp. Clin. Cancer Res.* **43**, 96 (2024).

53. Li, L. et al. Diagnosis of pulmonary nodules by DNA methylation analysis in bronchoalveolar lavage fluids. *Clin. Epigenet.* **13**, 185 (2021).

54. Wang, Z. et al. Early detection and stratification of lung cancer aided by a cost-effective assay targeting circulating tumor DNA (ctDNA) methylation. *Respir. Res.* **24**, 163 (2023).

55. Wang, C. et al. Characterization of distinct circular RNA signatures in solid tumors. *Mol. Cancer* **21**, 63 (2022).

56. Wang, C. et al. RNA-seq profiling of circular RNA in human lung adenocarcinoma and squamous cell carcinoma. *Mol. Cancer* **18**, 134 (2019).

57. Riely, G. J. et al. Non-small cell lung cancer, version 4.2024, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Cancer Netw.* **22**, 249–274 (2024).

58. Wang, M., Herbst, R. S. & Boshoff, C. Toward personalized treatment approaches for non-small-cell lung cancer. *Nat. Med.* **27**, 1345–1356 (2021).

59. Tan, A. C. & Tan, D. S. W. Targeted therapies for lung cancer patients with oncogenic driver molecular alterations. *J. Clin. Oncol.* **40**, 611–625 (2022).

60. Wang, C. et al. The landscape of immune checkpoint inhibitor therapy in advanced lung cancer. *BMC Cancer* **21**, 968 (2021).

61. Hu, J., Zhang, J., Wan, S. & Zhang, P. Neoadjuvant immunotherapy for non-small cell lung cancer: Opportunities and challenges. *Chin. Med J. Pulm. Crit. Care Med.* **2**, 224–239 (2024).

62. Wang, S. et al. Mining whole-lung information by artificial intelligence for pre-dicting EGFR genotype and targeted therapy response in lung cancer: a multi-cohort study. *Lancet Digit Health* **4**, e309–e319 (2022).

63. Zhang, S. et al. Intelligent prognosis evaluation system for stage I-III resected non-small-cell lung cancer patients on CT images: a multi-center study. *eClini-calMedicine* **65**, 102270 (2023).

64. Wang, C. et al. Predicting EGFR and PD-L1 status in NSCLC patients using mul-titask AI system based on CT images. *Front. Immunol.* **13**, 813072 (2022).

65. Deng, H. et al. A fluorogenic probe for predicting treatment response in non-small cell lung cancer with EGFR-activating mutations. *Nat. Commun.* **13**, 6944 (2022).

66. Wang, Y. et al. Low T790M relative allele frequency indicates concurrent resis-tance mechanisms and poor responsiveness to osimertinib. *Transl. Lung Cancer Res.* **9**, 1952–1962 (2020).

67. Huang, J. et al. Targeted drug/gene/photodynamic therapy via a stimuli-responsive dendritic-polymer-based nanococktail for treatment of EGFR-TKI-resistant non-small-cell lung cancer. *Adv. Mater.* **34**, e2201516 (2022).

68. Wang, Y. et al. Noninvasive genotyping and monitoring of anaplastic lymphoma kinase (ALK) rearranged non-small cell lung cancer by capture-based next-gen-eration sequencing. *Oncotarget* **7**, 65208–65217 (2016).

69. Tian, P. 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* **11**, 2098–2107 (2021).

70. Wang, C. et al. Non-invasive measurement using deep learning algorithm based on multi-source features fusion to predict PD-L1 expression and survival in NSCLC. *Front. Immunol.* **13**, 828560 (2022).

71. Shao, J. et al. Predicting gene mutation status via artificial intelligence technol-ogies based on multimodal integration (MMI) to advance precision oncology. *Semin. Cancer Biol.* **91**, 1–15 (2023).

72. Shao, J. et al. Radiogenomic system for non-invasive identification of multiple actionable mutations and PD-L1 expression in non-small cell lung cancer based on CT images. *Cancers* **14**, 4823 (2022).

73. Shao, J. et al. Deep learning empowers lung cancer screening based on mobile low-dose computed tomography in resource-constrained sites. *Front. Biosci.* **27**, 212 (2022).

74. Tao, W. et al. Telemedicine-enhanced lung cancer screening using mobile com-puted tomography unit with remote artificial intelligence assistance in under-served communities: initial results of a population cohort study in western China. *Telemed. J. E Health* **30**, e1695–e1704 (2024).