Artificial intelligence in lung cancer

Claudia Bardoni[1], Lorenzo Spaggiari[1,2], Luca Bertolaccini[1]^

[1]Department of Thoracic Surgery, IEO, European Institute of Oncology IRCCS, Milan, Italy; [2]Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy

Correspondence to: Luca Bertolaccini, MD, PhD, FCCP. Department of Thoracic Surgery, IEO, European Institute of Oncology IRCCS, Via Ripamonti 435, 20141 Milan, Italy. Email: luca.bertolaccini@gmail.com.

Keywords: Lung cancer; artificial intelligence (AI); diagnosis

Submitted Jun 02, 2022. Accepted for publication Jan 12, 2024. Published online Apr 07, 2024.
doi: 10.21037/atm-22-2918
View this article at: https://dx.doi.org/10.21037/atm-22-2918

In the United States, lung cancer is the most significant cause of cancer-related fatalities (1). The disease stage at diagnosis is the most important prognostic factor in lung cancer. In comparison, the 5-year survival rate for patients with early-stage cancer is 56%, and the 5-year survival rate for patients with advanced disease is less than 5%. Since most cases are diagnosed at an advanced stage, the prognosis for most diagnosed individuals is dismal. Early diagnosis is critical to reducing lung cancer-related mortality, which requires precise and quick identification of lung nodules and confirmation by histopathology (2). Artificial intelligence (AI) can assist clinicians with early diagnosis, decision-making, and prognosis forecasting. In the 1950s, Alan Turing used the term "Computing Machinery and Intelligence" (3) to characterise the AI concept. In 1956, the "Dartmouth Summer Research Project on AI" was the first worldwide conference on AI; here, researchers attempted to build intelligence comparable to that of the human brain by changing characteristics known as "reasoning as a search" (4).

We aim to give an updated view of AI's evolution and the current state-of-the-art AI relative to lung cancer up to 2024, discussing AI techniques in the radiological and histopathological diagnosis of lung cancer.

AI: the beginning of the screening process

In 2011, Beck *et al.* (5) utilised pathological image processing for breast cancer prognosis in seminal publications. The National Lung Screening Trial (NLST) was a unique trial in the same year. The study assesses whether screening with low-dose computed tomography (CT) can minimise lung cancer mortality relative to chest radiography. The patient population consisted of smokers with at least 30 pack-years of experience or former smokers who had quit within the previous 15 years, all between the ages of 55 and 74.

Edwards *et al.* (6) employed AI to predict whether lung nodules were benign. To accomplish this, they used clinical and radiological data from 165 consecutive patients who underwent surgery at their facility. Then, they develop an algorithm with a 96% sensitivity and an 89% specificity.

AI, supervised learning and convolutional neural networks (CNNs)

Recently, AI has succeeded in medical picture analysis using deep learning algorithms. Chiu *et al.* (7) provided a comprehensive description of AI models, broadly categorised as supervised, unsupervised, semi-supervised, and reinforcement learning. To train the algorithm, researchers must provide a labelled dataset in supervised learning, including inputs and expected outputs. It is appropriate for solving prediction issues like classification and regression. Using the algorithm's computed probability, researchers can convert a discrete response into a continuous variable to solve regression issues and vice versa. Numerous applications are derived from this, including survival prediction, cancer risk assessment, nodule identification,

^ ORCID: 0000-0002-1153-3334.

and nodule characterisation. In unsupervised learning, samples are partitioned by the algorithm corresponding to their inputs. Data without labels are optimal for clustering and dimensionality reduction. Even if the algorithm produced by supervised learning is more accurate, labelled data are uncommon. Unsupervised learning can use unlabelled data, but the method's accuracy is reduced. Thus, semi-supervised learning may have both the benefits of supervised learning and the ability to generate a labelling tool. Additionally, supervised learning may be able to build a massive dataset for additional training. So-called "reinforcement learning" is a reward-based process: the algorithm evolves through interaction with the data set. AI aids in earlier diagnosis and the improvement of diagnostic tools towards a more personalised therapy based on molecular, genetic, and histopathological characteristics. AI and deep learning aid pathology image analysis activities such as tumour region identification, prognosis prediction, tumour microenvironment characterisation and metastasis detection. Wang *et al.* (8) highlighted how deep learning approaches using an appropriate neural network design can enhance pathologists' performance by raising the accuracy of cancer detection and histology subtype classification. They discovered that a deep learning model, such as CNNs, can automatically extract features and oversee complex prediction problems.

Another benefit of CNN is its flexible neural network construction, which consists of loss function selection—defined as the difference between the ground truth and the prediction—and structure designation. One of CNN's weaknesses is classification accuracy, such as distinguishing between malignant and non-malignant tissues. Wang *et al.* built a CNN model to categorise particular image patches from hematoxylin and eosin (H&E)-stained lung adenocarcinoma whole slide imaging (WSI) as malignant or non-malignant (8), with an accuracy of 89.89% (9). Using traditional image processing techniques, Wang *et al.* also retrieved small picture patches of cell nuclei (8). They then sorted these picture patches into several cell types using a CNN. A predictive model was developed using picture characteristics on proportion and analysing the distribution of identified cells. Using a well-trained AI model can reduce human workload and associated errors. It may outperform a pathologist in a limited time and has a higher detection rate for single-cell or micrometastases. Sakamoto *et al.* (10) presented an additional study that addressed the difficulty of applying AI in pathological diagnosis, recognising cancer regions, predicting histologic subtypes of cancer, finding lymph node metastases, and assessing tumour cellularity for genomic analysis. Sakamoto's team created an AI model to evaluate the percentage of tumour cells in lung adenocarcinoma to identify tumour regions and a model for detecting unique nuclei to count the number of cells in each location (10). Using this model, their prospective study examines 53 biopsies; the study reveals that pathologists adjusted their choice by comparing it to the findings of AI analysis and tended to overstate the proportion of cancer. Li *et al.* researched and analysed the performance of several CNNs for categorising specific image pixel patches as malignant versus non-malignant (11). As Munir *et al.* (12) reported, deep learning models are available and have been applied to various forms of cancer, including brain, skin, prostate, and breast cancer. Paul *et al.* (13,14) used a CNN pre-trained on large-scale data to diagnose lung cancer by extracting features from CT scans.

A new era: radiomics

According to Tunali *et al.* (15), "radiomics" has evolved as a scientific discipline. Radiomics primarily aims to transform standard-of-care images into quantitative image-based data. Therefore, AI techniques can analyse these data. Using frequently accessible standard-of-care photos, it is a non-invasive technology designed to enhance clinical decision-making. Another study recommended CNN as a diagnostic technique for lung cancer. Joy Mathew *et al.* (16) reported several AI models for lung cancer screening. They described how CNN improved classification accuracy and lung nodule classification detection rate. Winkels *et al.* (17) suggested a 3D CNN with group convolutions (3D-G-CNNs) classification approach for lung nodules. They enhanced CNN's data efficiency and offered a potential answer to the problem of insufficient labelled data while training a CNN model with satisfactory performance. Computer-aided detection (CAD) systems are algorithms used in diagnosing lung nodules, as described by Binczyk (18). CAD is based on lung segmentation, identification, and classification of pulmonary nodules. Currently, the AI-RAD Companion framework by Siemens Healthcare is used to do lung segmentation by removing the background and unwanted portions from the CT image input to narrow the image region for further analysis. Nodule candidate detection identifies structures within the lung suspected of harbouring cancer. A polygonal approximation technique was presented in 2019, followed in 2020 by a neuro-evolutionary strategy. Despite Binczyk's description of deep learning as the most accurate CAD tool, additional research is still required. AI can catalyse to increase the efficiency of lung cancer screening, particularly in primary care settings, for timely referral and adequate incidental pulmonary nodules (IPNs) management.

AI challenging healthcare cost-effectiveness

Goncalves *et al.* observed how AI could transform healthcare by increasing the accuracy and automatic identification of radiographic lung lesions on images captured during intentional screening programs (19). Moreover, they retrospectively studied IPN detection through AI pandemics, using local and national databases from Russia, Latin America, and Asia. They observed that AI is defined by enhanced accuracy, efficiency, and precision, reducing inter-individual variability and bias, reducing the burden on healthcare providers, and improving democratisation and healthcare cost-effectiveness, especially in countries where radiology expertise is mediocre or primary care is in an under-served area. New treatment strategies have held to increased complexity in the last decade, and immunotherapy, molecularly targeted therapy, or advanced radiotherapy (RT) techniques (e.g., proton therapy) have changed the clinical scenario.

AI: what is new?

AI may lead to the further development of personalised immunotherapy and molecular target therapy treatment techniques. Tanaka *et al.* (20) emphasised the promise of AI in merging omics data and clinical information by analysing new biomarkers for the future of personalised medicine. Targeting oncogenic driver mutations with inhibitors may become a new norm in molecular targeted therapy. AI will play several roles in developing and revising innovative therapeutic regimens. Resistance to targeted therapies for oncogenic driver aberrations has been associated with secondary oncogenic driver aberrations, exhibiting several resistance molecular pathways. Due to the virtual screening of these particles, specialised treatment strategies can be developed to combat the mechanism for overcoming mixed resistance. According to Zhang *et al.* (21), targeted treatment is a novel method for treating cancer, and AI may facilitate the reapplication of old targeted medications to new indications. Wang *et al.* proposed an end-to-end deep learning model for predicting the chance of an EGFR mutation. They analysed over 15,000 CT scans of 844 individuals for EGFR mutation-related characteristics (8). Their model yielded positive findings in the primary cohort [area under the curve (AUC) =0.85; 95% confidence interval (CI): 0.83–0.88] and independent validation cohort (AUC =0.81; 95% CI: 0.79–0.83).

According to Etienne *et al.* (22), the quality of AI analysis depends on the quantity and quality of available data. Pathologists and radiologists must oversee data in the diagnosis of lung cancer. It is essential to continually modify and update data-driven algorithms so that decision-making can be changed and improved. More research and trials are required to ensure safe and secure utilisation in daily clinical practice. Their implementation would lead to a cost-effective benefit for the healthcare system. Notably, AI application in lung cancer has a bright and promising future.

A study presents a compelling fusion of deep learning and radiomics to predict EGFR mutations in non-small cell lung cancer (NSCLC) patients using CT imaging. Integrating 1,280 patient datasets, encompassing radiomics features and clinical data, resulting in a robust predictive model with high AUC values during internal and external validation. The potential for CT-image-based genetic testing to streamline EGFR mutation predictions highlights its simplicity and underscores its promising role in refining personalised treatment strategies for NSCLC patients (23).

After one or more prospective multicenter randomised controlled trials (phase III studies) have demonstrated advantages in the target population, new medications are approved for sale in the medical field. If we apply this to AI algorithms, prospective multicenter randomised controlled trials would be required to gather the necessary data. Although numerous studies have emphasised the significance of comprehensive testing before the application of AI, in reality, randomised controlled trials (randomised controlled trials) for AI software are rare and are optional for regulatory approval. Proving the efficacy of an AI system is difficult since integrating AI systems into health systems is contingent on numerous elements that are difficult to examine concurrently. In addition, the continual development of AI algorithms as they are fed increasing amounts of training data complicates the design of randomised controlled trials for AI (14). Although randomised controlled trials are an essential tool for establishing causality, there needs to be a consensus on their function in directing the deployment of AI in health care (24).

AI would lead to a synergistic collaboration between surgeons, pathologists, and radiologists with advanced technology (25). Still, impose well-accepted knowledge into deep learning models to incorporate prior knowledge into deep learning models. Histopathological diagnosis is a crucial and definitive tool for confirming diagnosis, disease grading, and progression. Thus, deep learning-based pathological image analysis can improve image analysis, helping pathologists and radiologists diagnose more accurately and earlier. Moreover, AI personalises patient management, increasing accuracy and learning by adaptation. Current models and research depend on retrospective studies, which causes selection and information bias. Also, some studies included a small sample size group for algorithm validation. Thus, future research should consist of both retrospective and prospective data to verify and optimise AI algorithms.

Acknowledgments

Funding: This work was partially supported by the Italian Ministry of Health with *Ricerca Corrente* and *5x1000* funds.

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editor (Calvin S. H. Ng) for the series "Lung Cancer Management—The Next Decade" published in *Annals of Translational Medicine*. The article has undergone external peer review.

Peer Review File: Available at https://atm.amegroups.com/article/view/10.21037/atm-22-2918/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://atm.amegroups.com/article/view/10.21037/atm-22-2918/coif). The series "Lung Cancer Management—The Next Decade" was commissioned by the editorial office without any funding or sponsorship. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66:7-30.
2. Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput Med Imaging Graph 2007;31:198-211.
3. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017;2:230-43.
4. McCarthy J, Minsky ML, Rochester N, et al. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. AI Magazine 2006;27:12.
5. Beck AH, Sangoi AR, Leung S, et al. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. Sci Transl Med 2011;3:108ra113.
6. Edwards FH, Schaefer PS, Cohen AJ, et al. Use of artificial intelligence for the preoperative diagnosis of pulmonary lesions. Ann Thorac Surg 1989;48:556-9.
7. Chiu HY, Chao HS, Chen YM. Application of Artificial Intelligence in Lung Cancer. Cancers (Basel) 2022;14:1370.
8. Wang S, Yang DM, Rong R, et al. Artificial Intelligence in Lung Cancer Pathology Image Analysis. Cancers (Basel) 2019;11:1673.
9. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017;60:84-90.
10. Sakamoto T, Furukawa T, Lami K, et al. A narrative review of digital pathology and artificial intelligence: focusing on lung cancer. Transl Lung Cancer Res 2020;9:2255-76.
11. Li J, Wu J, Zhao Z, et al. Artificial intelligence-assisted decision making for prognosis and drug efficacy prediction in lung cancer patients: a narrative review. J Thorac Dis 2021;13:7021-33.
12. Munir K, Elahi H, Ayub A, et al. Cancer Diagnosis Using Deep Learning: A Bibliographic Review. Cancers (Basel) 2019;11:1235.
13. Paul TK, Iba H. Prediction of cancer class with majority voting genetic programming classifier using gene expression data. IEEE/ACM Trans Comput Biol

Bioinform 2009;6:353-67.
14. Paul R, Hawkins SH, Hall LO, et al. Combining deep neural network and traditional image features to improve survival prediction accuracy for lung cancer patients from diagnostic CT. 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Budapest, Hungary: IEEE, 2016:2570-5.
15. Tunali I, Gillies RJ, Schabath MB. Application of Radiomics and Artificial Intelligence for Lung Cancer Precision Medicine. Cold Spring Harb Perspect Med 2021;11:a039537.
16. Joy Mathew C, David AM, Joy Mathew CM. Artificial Intelligence and its future potential in lung cancer screening. EXCLI J 2020;19:1552-62.
17. Winkels M, Cohen TS. 3D G-CNNs for Pulmonary Nodule Detection. arXiv 2018. arXiv:1804.04656.
18. Binczyk F, Prazuch W, Bozek P, et al. Radiomics and artificial intelligence in lung cancer screening. Transl Lung Cancer Res 2021;10:1186-99.
19. Goncalves S, Fong PC, Blokhina M. Artificial intelligence for early diagnosis of lung cancer through incidental nodule detection in low- and middle-income countries-acceleration during the COVID-19 pandemic but here to stay. Am J Cancer Res 2022;12:1-16.
20. Tanaka I, Furukawa T, Morise M. The current issues and future perspective of artificial intelligence for developing new treatment strategy in non-small cell lung cancer: harmonization of molecular cancer biology and artificial intelligence. Cancer Cell Int 2021;21:454.
21. Zhang H, Meng D, Cai S, et al. The application of artificial intelligence in lung cancer: a narrative review. Transl Cancer Res 2021;10:2478-87.
22. Etienne H, Hamdi S, Le Roux M, et al. Artificial intelligence in thoracic surgery: past, present, perspective and limits. Eur Respir Rev 2020;29:200010.
23. Kim S, Lim JH, Kim CH, et al. Deep learning-radiomics integrated noninvasive detection of epidermal growth factor receptor mutations in non-small cell lung cancer patients. Sci Rep 2024;14:922.
24. Schreuder A, Scholten ET, van Ginneken B, et al. Artificial intelligence for detection and characterization of pulmonary nodules in lung cancer CT screening: ready for practice? Transl Lung Cancer Res 2021;10:2378-88.
25. Bertolaccini L, Solli P, Pardolesi A, et al. An overview of the use of artificial neural networks in lung cancer research. J Thorac Dis 2017;9:924-31.

Cite this article as: Bardoni C, Spaggiari L, Bertolaccini L. Artificial intelligence in lung cancer. Ann Transl Med 2024;12(4):79. doi: 10.21037/atm-22-2918

