



# Progress and challenges of artificial intelligence in lung cancer clinical translation



Erjia Zhu<sup>1,2,7</sup>, Amgad Muneer<sup>2,7</sup>, Jianjun Zhang<sup>1</sup>, Yang Xia<sup>3</sup>, Xiaomeng Li<sup>4</sup>, Caicun Zhou<sup>5</sup>, John V. Heymach<sup>1</sup>, Jia Wu<sup>1,2,6</sup> ✉ & Xiuning Le<sup>1</sup> ✉

Artificial intelligence (AI) algorithms, such as convolutional neural networks and transformers, have significantly impacted cancer care. For lung cancer, AI holds great potential in addressing smoking cessation, personalized screening, and imaging genomics. And these data could be incorporated to optimize treatment selection. This review highlights the transformative impact of AI in lung cancer management, discusses crucial barriers such as model bias and fairness, and outlines future directions for clinical application.

## Advancement of AI in oncology

Artificial Intelligence (AI) refers to computational systems that can perform tasks typically requiring human intelligence. AI has existed for decades, with deep learning now far ahead of the competition. Deep learning excels at discerning intricate patterns, especially in big data, and can deliver quantitative assessment automatically<sup>1</sup>. Supervised learning methods, such as convolutional neural networks (CNNs), have been widely applied in image analysis while they rely on intensive annotations. Weakly supervised methods, including multiple-instance learning and vision transformers (ViTs), address annotation limitations while enabling large-scale data analysis. Unsupervised methods, like self-supervised learning, further uncover hidden patterns in unlabeled data<sup>2</sup>. Deep learning has enabled AI to transition from theoretical concepts to practical applications in medicine<sup>3–6</sup>. In the 2010s, the remarkable performance of CNNs in dermoscopic diagnosis of skin cancer established them as a benchmark in medical image analysis<sup>7</sup>. In the 2020s, transformers, an architecture that revolutionized natural language processing<sup>8</sup>, has also enhanced the performance in medical image analysis when integrated with CNN<sup>9</sup>. It also facilitates the integration of multimodal data types, which was previously difficult to achieve only with CNNs. Recently, multimodal AI has been delivering more comprehensive and accurate insights across various application. As the backbone of generalist AI, this technology enables the integration of multi-modal medical data and supports multitasking capabilities. Generalist medical AI holds significant potential to revolutionize cancer treatment and research<sup>10</sup>.

AI is progressively being integrated into various facets of oncology. Radiology plays a crucial role in diagnosis and in monitoring tumors throughout the treatment process. The rapidly increasing volume of scans has significantly increased the workload for radiologists, leading to burnout and reduced interpretive accuracy<sup>11</sup>. However, the vast amount of data creates an ideal environment for AI development. Radiomics refers to the extraction of large amounts of quantitative features from radiological image using advanced computational algorithms. Deep learning methods such as ResNet, U-Net, and YOLO offer significant advantages for feature extraction by automatically learning hierarchical, high-dimensional representations from raw imaging data, enabling more robust and informative characterization of tumor phenotype compared to handcrafted features<sup>12</sup>. It ushers in a new era of “virtual biopsy”, enabling the extraction of clinically relevant pathology and genomic information from radiology scans across different lesion sites and at multiple time points<sup>13</sup>. As the development of GPU computational speed, the clinical integration of digital pathology platforms has enabled the real-time implementation of AI on whole-slide imaging. Pathological AI has tremendous potentials to reshape pathology practice by streamlining diagnostic tasks, enhancing diagnostic precision, improving workflow efficiency. Pathologists are freed to focus more on intellectually challenging diagnoses<sup>14</sup>. Large language models (LLMs) are being tested on summarizing and processing medical text, including radiology reports<sup>15</sup>, and pathology reports<sup>16</sup>, as well as serving as a medical chatbot to provide personalized treatment recommendations<sup>17</sup>. Although

<sup>1</sup>Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>2</sup>Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>3</sup>Department of Respiratory and Critical Care Medicine, Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. <sup>4</sup>Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China. <sup>5</sup>Department of Medical Oncology, Shanghai East Hospital, Tongji University School of Medicine, Shanghai, China. <sup>6</sup>Institute of Data Science in Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>7</sup>These authors contributed equally: Erjia Zhu, Amgad Muneer.

✉ e-mail: [JWu11@mdanderson.org](mailto:JWu11@mdanderson.org); [XLe1@mdanderson.org](mailto:XLe1@mdanderson.org)

the clinical application of AI remains in its early stages due to regulatory and validation challenges, it offers a promising future for oncology.

### AI in lung cancer

Lung cancer continues to be the leading cause of cancer-related deaths worldwide, with an estimated 1.8 million deaths per year<sup>18</sup>. Advances in understanding biological mechanisms and evolving treatment strategies have significantly improved patient survival rates. However, the vast amount of clinical and research data remains insufficiently integrated and analyzed. For instance, although various non-smoking-related risk factors for lung cancer have been identified, they have not yet been integrated into a reliable model for estimating risk in healthy individuals. This gap poses critical challenges to identify populations that would benefit most from lung cancer screening. Predicting the malignancy risk of intermediate pulmonary nodules on CT scans is also a difficult issue. Data-driven methods are needed to guide treatment selection for each patient based on biomarkers. With high-quality real-world data, these clinical challenges could be addressed by deep learning. AI has the potential to revolutionize lung cancer management across multiple domains, including prevention, screening, diagnosis, prognosis, treatment, and monitoring. (Fig. 1) Representative studies were summarized in Table 1 with tasks, number of cases, data modality, algorithm, and performance.

In this narrative review, we explored the translational potential of AI in lung cancer, focusing on urgent clinical challenges from the perspective of oncologists. Rather than detailing technical methodologies, we emphasized the functional capabilities of AI—what it can currently do and what it may enable in the future—in a manner accessible to clinicians. We further outlined key barriers to clinical translation, specifically for readers who are not data scientists. References for this review were identified through searches of MEDLINE, PubMed, and citations from relevant articles using the term “lung cancer” combined with keywords such as “Artificial Intelligence,”

“AI,” “Machine Learning,” “Deep Learning,” “Radiomics,” and “Large Language Model.” The search primarily focused on English-language publications from January 1, 2014, to December 21, 2024.

### Prevention

Tobacco remains the primary etiological factor for lung cancer. Smoking cessation or control is the most effective strategy for reducing its risk. However, it is estimated that global population of cigarette smokers remains close to 1 billion<sup>19</sup>. To support smoking cessation, AI can analyze images of a smoker’s daily environment to identify contexts associated with smoking cravings<sup>20,21</sup>, and can monitor smoking behavior by studying signals of wearable sensors and assess puff topography<sup>22</sup>. However, addressing smoking at the individual level alone is insufficient. Tobacco control is a major public health challenge as smoking prevalence is influenced by various factors, including being male, having a lower socioeconomic status, experiencing significant physical or mental health conditions, identifying as lesbian, gay, bisexual, or transgender, or belonging to certain racial and ethnic groups<sup>23</sup>. The underlying causes of smoking cravings and effective interventions for these populations are not yet fully understood. To tackle this issue, public health researchers must work closely with AI experts to harness the potential of advanced AI tools in supporting smoking cessation efforts.

### Screening

The U.S. Preventive Services Task Force recommends annual low-dose CT (LDCT) screening for people aged 50–80 years and a smoking history of 20 pack-years. However, this is an imprecise criterion. To date, few studies have directly compared the use of pack-year smoking history and age versus other measures. Some evidence suggested that 20-year smoking duration was better than 20-pack-year<sup>24</sup>. Risk of lung cancer is influenced by a range of factors beyond age and smoking, including ethnicity, genetics, and

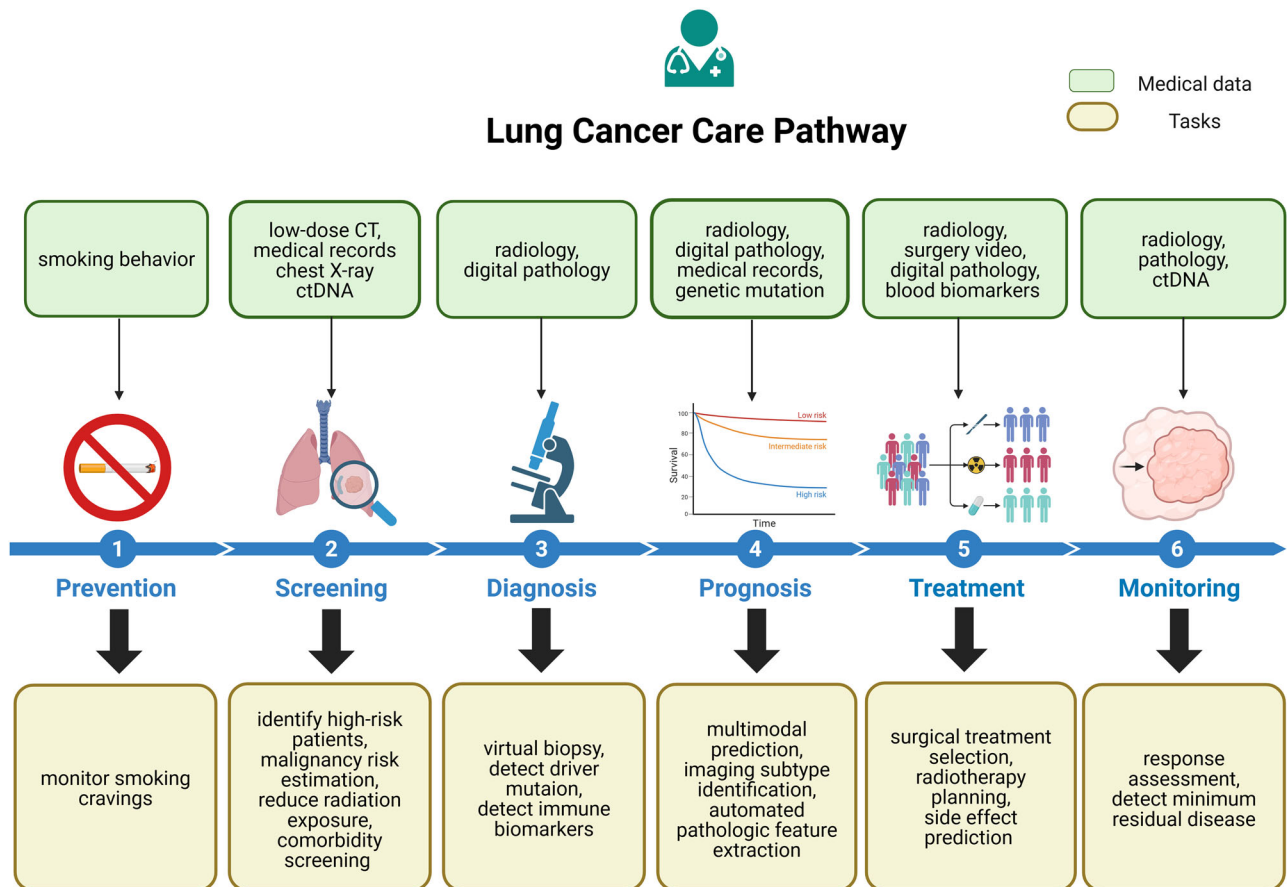


Fig. 1 | AI applications in lung cancer care pathway. Created in BioRender. Zhu, E. (2025) <https://BioRender.com/f83y797>.

**Table 1 | Summary of representative AI studies in lung cancer**

Reference	Year	Steps	Tasks	Number of cases	Data modality	Algorithm	Performance
Engelhardt et al. <sup>20</sup>	2021	Prevention	Predict environment-associated risk of smoking craving	48	Environment images	GNN	AUC >0.8
Senyurek et al. <sup>25</sup>	2020	Prevention	Smoking detection	40	Puff topography	GNN	F1-score 78%
Gould et al. <sup>28</sup>	2021	Screening	Predict lung cancer risk	196,102	Clinical and laboratory data	XGBoost	AUC 0.86
Lu et al. <sup>29</sup>	2020	Screening	Predict lung cancer risk	41,856	Medical records and chest X-ray	GNN	AUC 0.755
White et al. <sup>30</sup>	2017	Screening	Predict lung cancer risk	4,813,985	Web Search Logs	Statistical classifier	True-positive rate ranging from 3% to 57%; False-positive rates ranging from 0.00001 to 0.001
Arcilla et al. <sup>33</sup>	2019	Screening	Malignancy risk estimation of lung nodules	6,716	Low-dose CT	GNN	AUC 0.944
Mikhael et al. <sup>34</sup>	2023	Screening	Malignancy risk estimation of lung nodules	15,000	Low-dose CT	GNN	AUC 0.86–0.94
Kirienko et al. <sup>45</sup>	2018	Diagnosis	Differentiation between primary and metastatic lung cancer	534	PET	Linear discriminant analysis (radiomics)	AUC 0.92
Wang et al. <sup>47</sup>	2023	Diagnosis	Diagnosis of malignant pleural effusion	918	CT	GNN	AUC 0.842
Perez-Johnston et al. <sup>49</sup>	2022	Diagnosis	Predict pathologic subtypes	219	CT	K-means clustering (radiomics)	N/A
Coudray et al. <sup>55</sup>	2018	Diagnosis	Subtype classification of NSCLC, driver mutation prediction	1,634	Whole-slide images	GNN	AUC 0.970
Rossi et al. <sup>51</sup>	2021	Diagnosis	EGFR mutation prediction	109	CT	Support-vector machine (radiomics)	AUC 0.850
Yamamoto et al. <sup>52</sup>	2014	Diagnosis	ALK mutation prediction	172	CT	Modified Random Forest (radiomics)	Accuracy 78.8%
Wu et al. <sup>60</sup>	2022	Diagnosis	PD-L1 expression assessment	173	Whole-slide images	GNN	Strong consistency between AI and pathologists (F1 = 0.9429–0.9458).
Chen et al. <sup>53</sup>	2023	Diagnosis	PD-L1 expression assessment	194	Contrast-enhanced CT	Linear Regression with FDR control; Elastic Net Regularization (radiomics)	AUC 0.70–0.72
Sun et al. <sup>54</sup>	2018	Diagnosis	CD8 <sup>+</sup> T cell assessment	135	CT	Elastic net regression (radiomics)	AUC 0.67
Park et al. <sup>61</sup>	2022	Diagnosis	Tumor-infiltrating lymphocytes assessment	923	Whole-slide images	Limit SCOPEIO (self-developed analytical tool)	AUC 0.9715
Zhao et al. <sup>65</sup>	2018	Prognosis	Predict invasiveness	523	CT	GNN	AUC 0.880
Hosny et al. <sup>73</sup>	2018	Prognosis	Prognostic prediction	1,194	CT	GNN	AUC 0.70
Lu et al. <sup>74</sup>	2020	Prognosis	Prognostic prediction	1,057	Whole-slide images	GNN	AUC 0.63
Zhong et al. <sup>75</sup>	2022	Treatment	Predict lymph node metastasis	3,096	CT	GNN	AUC 0.82
Das et al. <sup>76</sup>	2023	Treatment	Interpretation of pulmonary function tests	24	Pulmonary function test	Self-developed AI software	Accuracy increased by 10.4%
Kadomatsu et al. <sup>79</sup>	2022	Treatment	Intraoperative air leak sites detection	110	Endoscopic images	GNN	sensitivity and specificity of 81.3% and 68.9%
Choi et al. <sup>81</sup>	2024	Treatment	Predict cardiac toxicity after radiotherapy	209	PET/CT	Tree-based Pipeline Optimization Tool (radiomics)	93% predictive accuracy
Dercle et al. <sup>84</sup>	2023	Treatment	Predict nivolumab response	758	CT	Random Forest (radiomics)	Sensitivity and specificity for predicting 3-month OS were 86% and 77.8%,
Sead et al. <sup>86</sup>	2023	Treatment	Predicting response to immune checkpoint inhibitors	976	CT	GNN	C-index 0.75
Arbour et al. <sup>102</sup>	2021	Monitoring	RECIST evaluation	453	Text reports	Fully Connected Deep Learning Model	AUC 0.9
Tran et al. <sup>103</sup>	2024	Monitoring	Identify patients with increased risk of relapse	85	CT and ctDNA	Cox regression model	AUROC 0.682

environmental exposures<sup>25–27</sup>. Traditional lineal regression models often struggle to process and interpret such complex and high-dimensional data effectively. To address this limitation, AI tools have been tested for identifying individuals at high risk of developing lung cancer. These tools leverage hidden patterns in routine clinical data<sup>28</sup>, chest X-rays<sup>29</sup>, extensive web search histories<sup>30</sup>, and survey responses<sup>31</sup>.

A significant challenge with current screening criteria is the high false-positive rate, as many detected nodules are either benign or of intermediate risk<sup>32</sup>. This results in unnecessary follow-ups and patient anxiety. While guidelines exist to help radiologists estimate malignancy risk, these assessments often remain subjective and heavily reliant on individual expertise. Deep learning-powered algorithms have been tested for automating cancer diagnosis, demonstrating promising performance in lung cancer screening. For instance, a deep learning algorithm developed by Google can analyze both current and prior CT scans of a patient. The model achieved a state-of-the-art performance (94.4% area under the curve) on 6716 National Lung Cancer Screening Trial (NLST) cases and outperformed six radiologists with absolute reductions of 11% in false positives and 5% in false negatives. If prior CT was available, the model performance was on-par with the same radiologists<sup>33</sup>. Another deep learning-based AI algorithm, Sybil, has demonstrated robust performance in predicting the future risk of developing lung cancer from a single LDCT. Sybil achieved area under the receiver-operator curves for lung cancer prediction at 1 year of 0.92 and 6 years of 0.75 on NLST<sup>34</sup>. Such AI can be used to personalize screening intervals to optimize resource utilization. Moreover, AI has enhanced interobserver agreement among radiologists for both risk stratification and management recommendations<sup>35</sup>. Patients undergo substantial radiation exposure during follow-up. Deep learning can revolutionize image reconstruction, by enabling lower doses of contrast agents, reducing ionizing radiation, and shortening image acquisition times. This advancement facilitates the application of ultra-low-dose CT for large-scale lung cancer screening<sup>36</sup>.

A meta-analysis showed that AI-based LDCT screening tools achieve high sensitivity (94.6%) but only moderate specificity (93.6%), translating to false-positive rates of ~6.4% and false-negative rates of ~5.4%<sup>37</sup>. Moreover, AI performance can also be skewed by biases in training datasets—such as variations in image quality, scan conditions, and vendor platforms—leading to inconsistent detection rates across institutions<sup>38</sup>. To mitigate these limitations, models should be developed and validated on large, multi-center, and demographically diverse cohorts, with systematic bias-audit frameworks and prospective external testing prior to clinical deployment.

Screening with LDCT presents opportunities for AI to simultaneously detect other smoking-related diseases, such as chronic obstructive lung disease<sup>39</sup>, and cardiovascular disease<sup>40</sup>. Beyond CT scans, deep learning has shown superior performance compared to radiologists in detecting lung nodules on X-rays<sup>41,42</sup>. In addition to radiographical scan, AI can also leverage blood biomarkers, including ctDNA<sup>43</sup> and plasma protein markers<sup>44</sup> to aid in the early-stage detection of lung cancer.

## Diagnosis

Lung cancer is a heterogeneous disease with diverse clinicopathological characteristics. AI could improve diagnosis in three domains by analyzing radiomics, digital pathology, and genomic sequencing data.

**Radiomics.** Advances in radiomics and deep learning technologies allow clinicians to derive comprehensive pathological insights from routine radiology scans prior to final pathological confirmation. This approach has been demonstrated to help differentiate lung cancer and benign lesions<sup>45</sup>, primary and metastatic lung lesions<sup>46</sup>, malignant and benign pleural effusions<sup>47</sup>, as well as adenocarcinoma and squamous cell carcinoma<sup>48</sup>, and even subtypes of adenocarcinoma<sup>49</sup>. It has demonstrated the ability to predict driver mutations, such as EGFR 19Del (exon 19) and L858R (exon 21)<sup>50</sup>, T790M (exon 20)<sup>51</sup>, and ALK rearrangement mutation<sup>52</sup>. Radiomics could also predict expression of PD-L1<sup>53</sup> and CD8<sup>+</sup> T cell<sup>54</sup> to guide immunotherapy.

**Digital pathology.** The field of computational pathology, despite being initially hindered by the high costs associated with digitalization, has gained significant momentum due to advances in AI. AI enabled the automatic diagnosis of lung cancer across various specimen types, including H&E slides<sup>55,56</sup>, cryosection tissue slides<sup>57</sup>, cytopathology samples<sup>58</sup>, and lymph node biopsy<sup>59</sup>. It also demonstrated the ability to predict driver mutations<sup>55</sup>, PD-L1 expression<sup>60</sup> and tumor-infiltrating lymphocytes<sup>61</sup> from H&E slides.

**Genomic sequencing.** Advances in predictive biomarker discovery have paved the way for targeted therapies and immunotherapies in lung cancer treatment. AI enhances somatic mutation identification in next-generation sequencing, outperforming standard genetic analysis approaches<sup>62,63</sup>. By decoding genomic and transcriptomic data, AI can accurately determine the cell-of-origin for cancers of unknown primary, aiding in diagnosis and treatment planning<sup>64</sup>. In the context of immune biomarkers, AI is able to predict other biomarkers like tumor mutation burden<sup>65</sup>, neoantigens<sup>66</sup>, and T-cell receptor-antigen binding specificity<sup>67</sup>.

## Prognosis

Lung cancer is primarily staged using the tumor, nodal, and metastasis classification system. In some cases, patients may require additional invasive procedures, such as endobronchial ultrasound biopsy, to assess nodal involvement. AI has emerged as a valuable tool for integrating multi-modal data, including medical records, radiology, pathology, and molecular data, to enhance staging accuracy and risk stratification. For instance, some pilot AI studies have been tested on routine radiology scans to predict the invasiveness of adenocarcinoma<sup>68,69</sup>, distant metastasis<sup>70</sup>, as well as to identify novel imaging subtypes<sup>71,72</sup>. In addition, AI automated the extraction of intricate features in medical imaging to offer new insights in prognostic stratification. For radiology, AI predictions were significantly associated with overall survival with AUC 0.70–0.71, outperforming clinical feature predictions with AUC 0.58–0.66<sup>73</sup>. For pathology, AI was able to predict overall survival with AUC 0.64–0.85, outperforming clinical feature predictions with AUC 0.52–0.84<sup>74</sup>.

## Treatment

**Surgery.** Sublobar resection is not inferior to lobectomy in patients with peripheral, node-negative NSCLC measuring 2 cm or smaller. However, lymph node-negative status can only be definitively confirmed after surgery. To address this limitation, a deep learning-based AI model was developed to predict lymph node metastasis preoperatively. The AI demonstrated strong performance and had the potential to assist surgeons in accurately identifying patients who are suitable candidates for sublobar resection<sup>75</sup>. Commonly, sublobar resection is reserved for patients with compromised pulmonary function. A collaboration between pulmonologists and AI could interpret pulmonary function tests (PFTs), which is crucial for assessing a patient's surgical candidacy<sup>76</sup>. When patients were selected for segmentectomy, the variability and complexity of intrathoracic anatomy present a significant challenge. Virtual reality systems have been developed to reconstruct thoracic anatomy, aiding in preoperative surgery planning and potentially reducing the duration of complex surgeries<sup>77,78</sup>. During the final stages of surgery, AI can assist by detecting air-leak sites through the analysis of surgical videos, even in deflated lungs. This capability enhances the surgeon's ability to address potential complications before closing thoracic cavities<sup>79</sup>.

**Radiotherapy.** Radiotherapy is a critical therapeutic approach, especially for locally advanced lung cancer, where it still holds curative potential. Accurate delineation of the gross tumor volume and consistent contouring of Organs at Risk are essential yet challenging. AI-based algorithms were tested for auto contouring and radiotherapy planning, which is especially useful for low- and middle-income

countries<sup>80</sup>. Moreover, radiomic models have been used to predict lung cancer recurrence<sup>69,70</sup>, cardiotoxicity<sup>81</sup>, and lung toxicity<sup>82</sup>, after radiotherapy.

**Systemic therapy.** NSCLC was poorly immunogenic. However, advances in immunotherapy have identified two key immune checkpoints relevant to NSCLC: CTLA-4 and the PD-1/PD-L1 axis. Clinical studies have consistently demonstrated that anti-PD-1 and anti-PD-L1 antibodies significantly improve patient survival compared to chemotherapy, marking a breakthrough in treatment strategies for this disease<sup>83</sup>. PD-L1 expression is the primary biomarker for predicting treatment response to immune checkpoint inhibitors. However, responses have also been observed in patients without detectable PD-L1 expression in their tumors. This phenomenon is likely attributed to the heterogeneity of PD-L1 expression both within a single tumor (intratumoral) and among different tumors (intertumoral). Such variability introduces inherent bias when relying on biopsy samples, which may not accurately represent the overall tumor micro-environment. Beyond traditional immune biomarkers, radiomic biomarkers have provided early indicators of survival in patients<sup>84,85</sup>. These markers can also predict adverse reactions of immunotherapy, such as hyperprogression<sup>86</sup> cachexia<sup>87</sup>, and immunotherapy-induced pneumonitis<sup>53</sup>. The deep learning model effectively captured additional imaging patterns beyond known hand-crafted features, enhancing predictive accuracy<sup>88</sup>. On the other hand, blood biomarkers, such as ctDNA<sup>89</sup> and cytokines<sup>90</sup> have also been valuable for AI in predicting responses to immunotherapy. AI's capacity to integrate multimodal data—including radiomics, pathomics, and genomics—into comprehensive big data analyses holds great promise for identifying immunotherapy responders, ultimately advancing personalized treatment strategies for lung cancer<sup>91</sup>.

EGFR mutations are the most commonly targetable driver mutations in lung adenocarcinoma. Third-generation EGFR-TKIs have significantly extended patients' survival. However, treatment resistance remains challenging. Combination strategies involving chemotherapy<sup>92</sup> or VEGF inhibitor<sup>93</sup> have been shown to improve the durability of response to EGFR-TKIs. Despite these benefits, patients receiving combination therapies experience a higher incidence of severe adverse events. Therefore, oncologists must carefully identify and select high-risk patients who are most likely to benefit from these approaches. Two studies found that AI could predict progression risk to identify high-risk patients<sup>94,95</sup>.

**Clinical decision support systems.** AI holds great potential in clinical decision support by integrating radiology, pathology, genomics, and clinical data<sup>96</sup>. Clinical Decision Support Systems, when effectively integrated with AI, can provide physicians with personalized treatment information<sup>97</sup>. Some studies explored the application of AI tools like Watson for Oncology (WFO) in decision-making of lung cancer patients. Preliminary results indicated AI's potential in adhering to clinical guidelines and assisting in decision-making. However, a relatively high proportion of cases are still not supported by WFO, and it needs to learn the regional characteristics of patients<sup>98,99</sup>.

### Monitoring

Currently, the evaluation of treatment response and disease progression in lung cancer primarily relies on lesion size, as outlined in the Response Evaluation Criteria in Solid Tumors (RECIST). Early differentiation between responders and non-responders is critical for timely adjustments to treatment regimens. However, the validity of RECIST has been questioned in the context of targeted therapies and immunotherapies due to phenomena such as pseudoprogression<sup>100</sup>. Noninvasive radiomic biomarkers can predict pseudoprogression and hyperprogression in patients with lung cancer with AUC of 0.88 (pseudoprogression vs. hyperprogression) and 0.87 (hyperprogression vs. progression)<sup>101</sup>. Additionally, response assessment is a time-

intensive process requiring significant expertise and is subject to high intra- and inter-reader variability. Deep learning has shown promise in automating this process. Applications include automated RECIST evaluations for patients receiving immunotherapy<sup>102</sup>.

Minimal residual disease (MRD) is strongly associated with disease progression in lung cancer. Monitoring circulating tumor DNA (ctDNA) in plasma has emerged as a valuable method for detecting MRD and predicting patient survival<sup>103</sup>. Longitudinal ctDNA detection offers insights into treatment response and can guide therapeutic strategies for patients with metastatic non-small-cell lung cancer (NSCLC)<sup>89</sup>. Additionally, machine learning approaches have shown promise in analyzing ctDNA kinetics, enabling the optimization of personalized therapies for NSCLC<sup>104</sup>.

### Large language models in lung cancer

LLMs can respond to free-text queries without requiring specific task training. This enables AI to learn and comprehend medical domain knowledge extremely rapidly and accurately. Medical chatbots, for instance, have demonstrated the capability to generate responses to patients queries that are comparable to those of clinicians, both in quality and empathy<sup>105</sup>. For lung cancer, LLMs may be used as decision aids<sup>106,107</sup>. Although promising, inaccuracy is the most concerning problem. LLMs can generate fabricating facts because they learn statistical word associations rather than achieving true understanding. Also, the training data is often from the internet, which is not verified. They function best as assistive tools under human supervision rather than in autonomous roles<sup>108</sup>. In the context of clinical trials, AI facilitates the matching process by aligning patient medical records against the enrollment criteria. Multiple studies have reported that AI can effectively extract patients' data and matches it to relevant clinical trials<sup>109</sup>.

### Approved AI devices in lung cancer

Before AI algorithms can be implemented in clinical settings, official approval is required. The pace of AI development challenges the appropriate regulatory frameworks and requires more staffs to efficiently process submissions. This process involves more stringent clinical trials and validation testing than what is usually presented in academic publications. FDA categorizes these AI medical devices according to the level of potential risk posed to individual patients. Many AI devices in oncology fall under Class II (moderate risk) for which randomized controlled clinical trials are not typically required. To integrate AI product into widespread use, well-controlled clinical studies are necessary to show that the product's benefits outweigh its risks. Also, most AI products performed well in predefined tasks like detection but lack generalizability across different patient populations requires validation. Consequently, only a small proportion of AI algorithms are eligible to be deployed in clinical settings. Among the approved AI applications in lung cancer, they mainly focus on lung nodule detection, diagnosis, and radiotherapy planning, with all these algorithms being imaging-based. (Table 2) Multi-party collaboration is needed to optimize and adjust regulatory frameworks and processes, improve AI development, validation, and documentation standards, address challenges of advanced and evolving AI, and strengthen full lifecycle management and post-market surveillance<sup>110,111</sup>.

### Challenges and opportunities

AI research in lung cancer offers promising prospects for automated and precise management but translating these advances into clinical practice faces several hurdles. Key challenges are discussed in the following.

### Data sharing

Continuous data supply is crucial for the effective training, validation, and refinement of AI algorithms. To develop robust AI tools, large high-quality datasets from multiple institutions are needed to address the limitation, such

**Table 2 | FDA approved AI devices in lung cancer**

Name	Data modality	Task	FDA summary	Year
Riverain ClearRead CT	CT	detection of pulmonary nodules in asymptomatic population	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf16/k161201.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf16/k161201.pdf</a>	2016
Arterys Oncology DL	CT, MRI	segmentation of lung nodules and liver lesions, automated reporting	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf17/k173542.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf17/k173542.pdf</a>	2017
Philips Lung Nodule Assessment and Comparison Option (LNA)	CT	characterization of nodule type, location, and measurements	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf16/k162484.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf16/k162484.pdf</a>	2017
Siemens AI-Rad Companion (Pulmonary)	CT	segmentation of lesions of the lung, liver, and lymph nodes	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf18/k183271.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf18/k183271.pdf</a>	2019
Coreline AView LCS	CT	characterization of nodule type, location, measurements, and Lung-RADS category	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf20/k201710.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf20/k201710.pdf</a>	2020
MeVis Veolity	CT	detection of solid pulmonary nodules, alerts to overlooked regions	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf20/k201501.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf20/k201501.pdf</a>	2021
Coreline AVIEW	CT	quantitative analysis of lung nodule CT scanning, calculation the malignancy score	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf21/k214036.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf21/k214036.pdf</a>	2022
Coreline AVIEW Lung Nodule CAD	CT	Detection of lung nodules	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf22/k221592.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf22/k221592.pdf</a>	2023
GE Xeleris V Processing and Review System	PET/CT	Assess the fraction of total lung function for lung cancer resection	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf22/k221680.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf22/k221680.pdf</a>	2023
uMI Panorama	PET and CT	detection of lung nodules	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf22/k223325.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf22/k223325.pdf</a>	2023
Siemens syngo.CT Lung CAD	CT	detection of solid pulmonary nodules, alerts to overlooked regions	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf23/k231157.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf23/k231157.pdf</a>	2023
Mantecia MOZI TPS	CT	radiotherapy treatments planning	<a href="https://www.accessdata.fda.gov/cdrh_docs/pdf22/k223724.pdf">https://www.accessdata.fda.gov/cdrh_docs/pdf22/k223724.pdf</a>	2023

**Table 3 | Publicly available lung cancer datasets and their description and challenges**

Dataset	Details on Dataset	Cons	Pros	Restrictions in Data Sharing	Access Link
NSCLC-Radiomics	Radiomic analysis dataset for NSCLC from CT and PET scans, consisting of 422 cases.	<ul style="list-style-type: none"> <li>- Preprocessing variability</li> <li>- Missing clinical data</li> <li>- Limited diversity of cases</li> </ul>	<ul style="list-style-type: none"> <li>- Enables radiomics studies</li> <li>- Well-annotated dataset for imaging research</li> </ul>	<ul style="list-style-type: none"> <li>- Requires institutional agreements</li> <li>- Redistribution without permission prohibited</li> </ul>	<a href="https://wiki.cancerimagingarchive.net/display/Public/NSCLC-Radiomics">https://wiki.cancerimagingarchive.net/display/Public/NSCLC-Radiomics</a>
NSCLC Radiogenomics	A dataset integrating radiomic features from CT imaging and genomic data for NSCLC, with 211 cases.	<ul style="list-style-type: none"> <li>- High dimensionality of genomic and radiomic features</li> <li>- Need for robust feature selection methods</li> <li>- Limited sample size for comprehensive radiogenomic analysis</li> <li>- Requires expertise in multi-modal data integration</li> </ul>	<ul style="list-style-type: none"> <li>- Facilitates multi-modal studies integrating radiomics and genomics</li> <li>- High-quality imaging data</li> </ul>	<ul style="list-style-type: none"> <li>- Requires agreement to specific licensing terms</li> <li>- Redistribution is restricted and subject to approval</li> </ul>	<a href="https://www.cancerimagingarchive.net/collection/nsclc-radiogenomics/">https://www.cancerimagingarchive.net/collection/nsclc-radiogenomics/</a>
LIDC-IDRI	Largest publicly available lung cancer imaging dataset, containing over 1,010 CT scans with nodule annotations.	<ul style="list-style-type: none"> <li>- Annotation inconsistencies</li> <li>- Inter-reader variability in nodule classification</li> <li>- Limited PET/CT inclusion</li> <li>- No clinical outcome data</li> </ul>	<ul style="list-style-type: none"> <li>- Large-scale dataset with diverse cases</li> <li>- Comprehensive nodule annotations</li> </ul>	<ul style="list-style-type: none"> <li>- Free for research purposes</li> <li>- Redistribution or commercial use prohibited</li> </ul>	<a href="https://www.cancerimagingarchive.net/collection/lidc-idri/">https://www.cancerimagingarchive.net/collection/lidc-idri/</a>
TCGA-LUAD	Cancer Genome Atlas Lung Adenocarcinoma dataset, including genomic and imaging data.	<ul style="list-style-type: none"> <li>- Complex preprocessing required</li> <li>- High dimensionality of genomic data</li> <li>- Limited imaging modalities included</li> <li>- Expertise needed for genomic-imaging integration</li> </ul>	<ul style="list-style-type: none"> <li>- Combines genomic and imaging data</li> <li>- Enables survival and outcome studies</li> </ul>	<ul style="list-style-type: none"> <li>- Requires agreement to TCGA's data-use policy</li> <li>- Certain portions require dbGaP application</li> </ul>	<a href="https://portal.gdc.cancer.gov/projects/TCGA-LUAD">https://portal.gdc.cancer.gov/projects/TCGA-LUAD</a>
NLST	National Lung Screening Trial dataset with CT imaging and demographic information for screening participants.	<ul style="list-style-type: none"> <li>- Limited metadata</li> <li>- Variability in screening protocols</li> <li>- No PET/CT imaging data</li> <li>- Restricted access to patient-level details</li> </ul>	<ul style="list-style-type: none"> <li>- Large-scale dataset for screening research</li> <li>- Longitudinal data available</li> </ul>	<ul style="list-style-type: none"> <li>- Requires formal data-use agreement</li> <li>- Access limited to qualified researchers</li> </ul>	<a href="https://cdas.cancer.gov/datasets/nlst/">https://cdas.cancer.gov/datasets/nlst/</a>
AutoPET	Dataset for segmentation in PET/CT imaging, featuring annotated lung cancer cases.	<ul style="list-style-type: none"> <li>- High annotation variability</li> <li>- Preprocessing needed to standardize image formats</li> <li>- Focused only on segmentation tasks</li> <li>- Limited patient numbers</li> </ul>	<ul style="list-style-type: none"> <li>- High-quality segmentation annotations</li> <li>- Facilitates PET/CT segmentation research</li> </ul>	<ul style="list-style-type: none"> <li>- Open access via registration</li> <li>- Cannot use for commercial applications without permission</li> </ul>	<a href="https://autopet.grand-challenge.org/Dataset/">https://autopet.grand-challenge.org/Dataset/</a>
RIDER Lung CT	Quantitative imaging biomarker dataset, including serial CT scans of 32 lung cancer patients.	<ul style="list-style-type: none"> <li>- Temporal scan alignment issues</li> <li>- Lack of outcome metadata</li> <li>- Limited patient diversity</li> <li>- No PET/CT data included</li> </ul>	<ul style="list-style-type: none"> <li>- Supports temporal analysis studies</li> <li>- Enables biomarker development research</li> </ul>	<ul style="list-style-type: none"> <li>- Freely available under RIDER project license</li> <li>- Redistribution not permitted</li> </ul>	<a href="https://www.cancerimagingarchive.net/collection/rider-lung-ct/">https://www.cancerimagingarchive.net/collection/rider-lung-ct/</a>
MIDRC	MIDRC is a collaborative initiative initially focused on COVID-19 chest imaging and, as of 2023, has begun expanding to include cancer imaging, including lung cancer.	<ul style="list-style-type: none"> <li>- Limited lung cancer-specific data.</li> </ul>	<ul style="list-style-type: none"> <li>- Centralized repository for imaging data.</li> <li>- Supports multi-institutional studies.</li> </ul>	<ul style="list-style-type: none"> <li>- Requires registration for access.</li> <li>- Redistribution without permission prohibited.</li> </ul>	<a href="https://www.midrc.org/">https://www.midrc.org/</a>

as statistical power, diversity, and clinical practice variations. However, sharing data is challenging due to concerns about patient privacy and intellectual property protection<sup>112</sup>. To address this issue, three primary options are available. The first one is centralized learning, with institutions creating a shared legal agreement and security protocol to pull data together. While effective, this approach is costly. The second option is through creating deidentified public datasets. Some commonly used lung cancer databases have been summarized. (Table 3) This option is more affordable but may lack certain types of patient information, making it difficult to train AI for specific clinical applications. The third option is federated learning. Data remains private at each institution, but AI models are trained in a distributed manner<sup>113</sup>. It has been implemented in several cancer-related applications, including breast cancer<sup>114</sup>, brain cancer<sup>115</sup>, gastric cancer<sup>116</sup>, melanoma<sup>117</sup>, and lung cancer<sup>118</sup>.

### Bias and fairness

AI models inevitably inherit associated biases that favor a particular racial, ethnic, or gender groups, resulting in poor performance when applied to diverse populations<sup>119</sup>. For instance, only 50% of Black women and 63% of Black men diagnosed with lung cancer qualified for screening<sup>120</sup>. According to the report of the 75,774 patient from The Society of Thoracic Surgeons General Thoracic Surgery database, white patients and those with private insurance had a higher incidence of complex operations<sup>121</sup>. Efforts are underway to generate more diverse datasets and reduce biases in both breast cancer<sup>122</sup> and lung cancer<sup>123</sup>. Additionally, AI algorithms can be specifically designed to ensure fairness, enhancing their effectiveness across varied demographic and socioeconomic groups<sup>124</sup>.

### Interpretability

This remains a significant challenge for AI, particularly with deep learning approaches that operate as end-to-end systems, mapping inputs directly to outcomes without manually selected features. This black-box nature makes it difficult to understand which factors are driving decisions, potentially leading to misleading conclusions due to spurious confounders in the data<sup>125</sup>. Such opacity is often deemed unacceptable in healthcare decision-making, posing a significant barrier to clinical utilization<sup>126</sup>. Consequently, explainable AI has emerged as a highly active research area, aiming to make AI models more transparent and understandable<sup>127</sup>. Despite these efforts, the optimal form of explainability for clinical use remains unknown, and even FDA-approved AI devices currently offer limited interpretability.

### Reproducibility and translation

A robust AI model requires independent review and test by external groups, which is crucial for assessing potential biases in datasets and ensuring generalizability across diverse clinical settings. Despite these needs, most published AI studies still lack reproducibility. Imaging protocols such as CT scanner manufacturer, radiation dose, convolution kernel, iterative reconstruction, and section thickness significantly impact the diagnostic performance of deep learning algorithms. This variability reduced the reliability in clinical practice<sup>128</sup>. Motion artifacts from patient breathing and image noise further degrade data quality, complicating tasks like nodule detection and segmentation. Additionally, annotation variability among radiologists introduces subjectivity, affecting accuracy. To address these issues, standardized preprocessing pipelines are essential to address these challenges and ensure robust, generalizable AI models for lung cancer applications. Recently, the Image Biomarker Standardization Initiative (IBSI) has made significant strides in establishing standards<sup>129,130</sup>. A list of 16 criteria for the optimal development of a radiomic test serves as a guide for the implementation of future radiomic analyses<sup>131</sup>. Various guidelines have been proposed to provide essential frameworks to report necessary information about AI modeling, including MINIMAR (MINimum Information for Medical AI Reporting)<sup>132</sup>, SPIRIT-AI (Standard Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence)<sup>133</sup>,

CONSORT-AI (Consolidated Standards of Reporting of Trials–Artificial Intelligence)<sup>134</sup>, and ESMO-GROW (European Society for Medical Oncology Guidance for Reporting Oncology Real-World Evidence)<sup>135</sup>.

### Future directions

AI is evolving rapidly, with the ultimate objective being the development of a comprehensive model, known as generalist AI, capable of analyzing multi-modal data and addressing a wide range of tasks. Currently, most AI models in healthcare are uni-modal and uni-task, requiring separate models for different types of medical data—such as medical records, radiology, pathology, and genomic data—to solve even a single task. Novel deep learning architectures can integrate multimodal, thereby improving model performance<sup>2</sup>. Recently, Path-Chat, a chatbot enabling interactive discussions with pathologists, has been introduced, potentially providing expert-level insights related to specific cases<sup>136</sup>. Extending this concept, generalist deep learning models could integrate comprehensive patient information and interact with physicians similarly to how ChatGPT functions. Such models could allow physicians to define prediction tasks in natural language, with the model explaining its predictions. Generalist AI has the potential to significantly enhance diagnostic and prognostic methods in oncology, shifting from task-specific models to a holistic, integrated approach.

Beyond traditional medical data like radiologic images and genomic information which are costly and not time-sensitive, technological advances in smartphones and wearable sensors can collect extensive physiological and environmental data for each patient. AI holds substantial promise in managing these large datasets to identify individuals at high risk for cancers that are influenced by environmental and behavior factors, such as lung cancer. In the future, real-time AI-assisted lung cancer prevention could offer personalized early intervention and risk management strategies while accumulating valuable data for researchers to identify underlying risk factors. Integrating personal data can also facilitate remote monitoring, providing alerts to primary physicians and patients as necessary, during the diagnosis and treatment course of lung cancer.

### Limitations

The discussed AI applications themselves face significant limitations hindering immediate widespread clinical translation. These include challenges in data sharing and quality, inherent model biases, the “black-box” nature affecting interpretability, and a general lack of reproducibility and external validation in many studies. Most currently approved AI tools are imaging-based and target specific tasks, indicating that the full potential of multi-modal, generalist AI is yet to be realized in routine clinical practice.

### Conclusion

AI is significantly advancing lung cancer care across prevention, screening, diagnosis, prognosis, treatment, and monitoring by analyzing complex data to personalize patient management. (Fig. 1) Deep learning algorithms show immense potential in improving diagnostic accuracy, predicting treatment responses, and automating tasks. However, challenges such as data sharing, model bias, lack of interpretability, and reproducibility issues hinder widespread clinical adoption. Multi-party collaboration is needed to optimize and adjust regulatory frameworks and processes, improve AI development, validation, and documentation standards, address challenges of advanced and evolving AI, and strengthen full lifecycle management and post-market surveillance. The development of generalist AI which capable of integrating multimodal data, will provide holistic and interactive decision support.

### Data availability

No datasets were generated or analyzed during the current study.

Received: 1 October 2024; Accepted: 1 June 2025;

Published online: 01 July 2025

## References

1. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
2. Lipkova, J. et al. Artificial intelligence for multimodal data integration in oncology. *Cancer Cell* **40**, 1095–1110 (2022).
3. Ozdemir, B. & Pacal, I. An innovative deep learning framework for skin cancer detection employing ConvNeXtV2 and focal self-attention mechanisms. *Results Eng.* **25**, 103692 (2025).
4. Bayram, B., Kunduracioglu, I., Ince, S. & Pacal, I. A systematic review of deep learning in MRI-based cerebral vascular occlusion-based brain diseases. *Neuroscience* **568**, 76–94 (2025).
5. Pacal, I. Investigating deep learning approaches for cervical cancer diagnosis: a focus on modern image-based models. *Eur. J. Gynaecol. Oncol.* **46**, 17 (2025).
6. Pacal, I., Ozdemir, B., Zeynalov, J., Gasimov, H. & Pacal, N. A novel CNN-ViT-based deep learning model for early skin cancer diagnosis. *Biomed. Signal Process. Control* **104**, 107627 (2025).
7. Esteva, A. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118 (2017).
8. Bhayana, R. Chatbots and large language models in radiology: a practical primer for clinical and research applications. *Radiology* **310**, e232756 (2024).
9. Shmatko, A., Ghaffari Laleh, N., Gerstung, M. & Kather, J. N. Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. *Nat. Cancer* **3**, 1026–1038 (2022).
10. Moor, M. et al. Foundation models for generalist medical artificial intelligence. *Nature* **616**, 259–265 (2023).
11. Alexander, R. et al. Mandating limits on workload, duty, and speed in radiology. *Radiology* **304**, 274–282 (2022).
12. Jiang, H. et al. A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation. *Comput. Biol. Med.* **157**, 106726 (2023).
13. Lambin, P. et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **14**, 749–762 (2017).
14. Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V. & Madabhushi, A. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* **16**, 703–715 (2019).
15. Adams, L. C. et al. Leveraging GPT-4 for post hoc transformation of free-text radiology reports into structured reporting: a multilingual feasibility study. *Radiology* **307**, e230725 (2023).
16. Truhn, D. et al. Extracting structured information from unstructured histopathology reports using generative pre-trained transformer 4 (GPT-4). *J. Pathol.* **262**, 310–319 (2024).
17. Singhal, K. et al. Large language models encode clinical knowledge. *Nature* **620**, 172–180 (2023).
18. 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).
19. Balata, H. et al. Prevention and early detection for NSCLC: advances in thoracic oncology 2018. *J. Thorac. Oncol.* **14**, 1513–1527 (2019).
20. Engelhard, M. M., D'Arcy, J., Oliver, J. A., Kozink, R. & McClemon, F. J. Prediction of smoking risk from repeated sampling of environmental images: model validation. *J. Med. Internet Res.* **23**, e27875 (2021).
21. Engelhard, M. M. et al. Identifying smoking environments from images of daily life with deep learning. *JAMA Netw. open* **2**, e197939 (2019).
22. Senyurek, V. Y., Imtiaz, M. H., Belsare, P., Tiffany, S. & Sazonov, E. A CNN-LSTM neural network for recognition of puffing in smoking episodes using wearable sensors. *Biomed. Eng. Lett.* **10**, 195–203 (2020).
23. Jamal, A. et al. Current cigarette smoking among adults—United States, 2016. *MMWR Morb. Mortal. Wkly Rep.* **67**, 53–59 (2018).
24. Potter, A. L. et al. Pack-year smoking history: an inadequate and biased measure to determine lung cancer screening eligibility. *J. Clin. Oncol.* **42**, 2026–2037 (2024).
25. Field, J. K., Vulkan, D., Davies, M. P. A., Duffy, S. W. & Gabe, R. Liverpool Lung Project lung cancer risk stratification model: calibration and prospective validation. *Thorax* **76**, 161–168 (2021).
26. Dai, J. et al. Identification of risk loci and a polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese populations. *Lancet Respir. Med.* **7**, 881–891 (2019).
27. Chen, C. Y. et al. The role of PM2.5 exposure in lung cancer: mechanisms, genetic factors, and clinical implications. *EMBO Mol. Med* **17**, 31–40 (2024).
28. Gould, M. K., Huang, B. Z., Tammemagi, M. C., Kinar, Y. & Shiff, R. Machine learning for early lung cancer identification using routine clinical and laboratory data. *Am. J. Respir. Crit. Care Med.* **204**, 445–453 (2021).
29. Lu, M. T., Raghu, V. K., Mayrhofer, T., Aerts, H. & Hoffmann, U. Deep learning using chest radiographs to identify high-risk smokers for lung cancer screening computed tomography: development and validation of a prediction model. *Ann. Intern. Med.* **173**, 704–713 (2020).
30. White, R. W. & Horvitz, E. Evaluation of the feasibility of screening patients for early signs of lung carcinoma in web search logs. *JAMA Oncol.* **3**, 398–401 (2017).
31. Chen, S. & Wu, S. Identifying lung cancer risk factors in the elderly using deep neural networks: quantitative analysis of web-based survey data. *J. Med. Internet Res.* **22**, e17695 (2020).
32. Force, U. S. P. S. T et al. Screening for lung cancer: US Preventive Services Task Force recommendation statement. *JAMA* **325**, 962–970 (2021).
33. 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).
34. 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).
35. Kim, R. Y. et al. Artificial intelligence tool for assessment of indeterminate pulmonary nodules detected with CT. *Radiology* **304**, 683–691 (2022).
36. Jiang, B. et al. Deep learning reconstruction shows better lung nodule detection for ultra-low-dose chest CT. *Radiology* **303**, 202–212 (2022).
37. Thong, L. T., Chou, H. S., Chew, H. S. J. & Lau, Y. Diagnostic test accuracy of artificial intelligence-based imaging for lung cancer screening: a systematic review and meta-analysis. *Lung Cancer* **176**, 4–13 (2023).
38. Fukumoto, W. et al. External validation of the performance of commercially available deep-learning-based lung nodule detection on low-dose CT images for lung cancer screening in Japan. *Jpn J. Radio.* **43**, 634–640 (2025).
39. Tang, L. Y. W. et al. Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose CT. *Lancet Digit. Health* **2**, e259–e267 (2020).
40. Xu, K. et al. AI body composition in lung cancer screening: added value beyond lung cancer detection. *Radiology* **308**, e222937 (2023).
41. Lee, J. H., Hong, H., Nam, G., Hwang, E. J. & Park, C. M. Effect of human-ai interaction on detection of malignant lung nodules on chest radiographs. *Radiology* **307**, e222976 (2023).
42. Lee, J. H. et al. Performance of a deep learning algorithm compared with radiologic interpretation for lung cancer detection on chest radiographs in a health screening population. *Radiology* **297**, 687–696 (2020).

43. Chabon, J. J. et al. Integrating genomic features for non-invasive early lung cancer detection. *Nature* **580**, 245–251 (2020).
44. Vachani, A. et al. Development and validation of a risk assessment model for pulmonary nodules using plasma proteins and clinical factors. *Chest* **163**, 966–976 (2023).
45. Beig, N. et al. Perinodular and intranodular radiomic features on lung CT images distinguish adenocarcinomas from granulomas. *Radiology* **290**, 783–792 (2019).
46. Kirienko, M. et al. Ability of FDG PET and CT radiomics features to differentiate between primary and metastatic lung lesions. *Eur. J. Nucl. Med. Mol. Imaging* **45**, 1649–1660 (2018).
47. Wang, S. et al. Differentiation of malignant from benign pleural effusions based on artificial intelligence. *Thorax* **78**, 376–382 (2023).
48. Han, Y. et al. Histologic subtype classification of non-small cell lung cancer using PET/CT images. *Eur. J. Nucl. Med. Mol. Imaging* **48**, 350–360 (2021).
49. Perez-Johnston, R. et al. CT-based radiogenomic analysis of clinical stage I lung adenocarcinoma with histopathologic features and oncologic outcomes. *Radiology* **303**, 664–672 (2022).
50. Li, S., Ding, C., Zhang, H., Song, J. & Wu, L. Radiomics for the prediction of EGFR mutation subtypes in non-small cell lung cancer. *Med. Phys.* **46**, 4545–4552 (2019).
51. Rossi, G. et al. Radiomic detection of EGFR mutations in NSCLC. *Cancer Res.* **81**, 724–731 (2021).
52. Yamamoto, S. et al. ALK molecular phenotype in non-small cell lung cancer: CT radiogenomic characterization. *Radiology* **272**, 568–576 (2014).
53. Chen, M. et al. A novel radiogenomics biomarker for predicting treatment response and pneumotoxicity from programmed cell death protein or ligand-1 inhibition immunotherapy in NSCLC. *J. Thorac. Oncol.* **18**, 718–730 (2023).
54. Sun, R. et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol.* **19**, 1180–1191 (2018).
55. Coudray, N. et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* **24**, 1559–1567 (2018).
56. Chen, C. L. et al. An annotation-free whole-slide training approach to pathological classification of lung cancer types using deep learning. *Nat. Commun.* **12**, 1193 (2021).
57. Ozyoruk, K. B. et al. A deep-learning model for transforming the style of tissue images from cryosectioned to formalin-fixed and paraffin-embedded. *Nat. Biomed. Eng.* **6**, 1407–1419 (2022).
58. Xie, X. et al. Deep convolutional neural network-based classification of cancer cells on cytological pleural effusion images. *Mod. Pathol.* **35**, 609–614 (2022).
59. Pham, H. H. N. et al. Detection of lung cancer lymph node metastases from whole-slide histopathologic images using a two-step deep learning approach. *Am. J. Pathol.* **189**, 2428–2439 (2019).
60. Wu, J. et al. Artificial intelligence-assisted system for precision diagnosis of PD-L1 expression in non-small cell lung cancer. *Mod. Pathol.* **35**, 403–411 (2022).
61. Park, S. et al. Artificial intelligence-powered spatial analysis of tumor-infiltrating lymphocytes as complementary biomarker for immune checkpoint inhibition in non-small-cell lung cancer. *J. Clin. Oncol.* **40**, 1916–1928 (2022).
62. AlDubayan, S. H. et al. Detection of pathogenic variants with germline genetic testing using deep learning vs standard methods in patients with prostate cancer and melanoma. *JAMA* **324**, 1957–1969 (2020).
63. Sherman, M. A. et al. Genome-wide mapping of somatic mutation rates uncovers drivers of cancer. *Nat. Biotechnol.* **40**, 1634–1643 (2022).
64. Moon, I. et al. Machine learning for genetics-based classification and treatment response prediction in cancer of unknown primary. *Nat. Med.* **29**, 2057–2067 (2023).
65. Jain, M. S. & Massoud, T. F. Predicting tumour mutational burden from histopathological images using multiscale deep learning. *Nat. Mach. Intell.* **2**, 356–362 (2020).
66. Sarkizova, S. et al. A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat. Biotechnol.* **38**, 199–209 (2020).
67. Lu, T. et al. Deep learning-based prediction of the T cell receptor-antigen binding specificity. *Nat. Mach. Intell.* **3**, 864–875 (2021).
68. Zhao, W. et al. 3D Deep learning from CT scans predicts tumor invasiveness of subcentimeter pulmonary adenocarcinomas. *Cancer Res.* **78**, 6881–6889 (2018).
69. Wu, J. et al. Robust intratumor partitioning to identify high-risk subregions in lung cancer: a pilot study. *Int. J. Radiat. Oncol. Biol. Phys.* **95**, 1504–1512 (2016).
70. Wu, J. et al. Early-Stage non-small cell lung cancer: quantitative imaging characteristics of (18)F fluorodeoxyglucose PET/CT allow prediction of distant metastasis. *Radiology* **281**, 270–278 (2016).
71. Wu, J. et al. Radiological tumor classification across imaging modality and histology. *Nat. Mach. Intell.* **3**, 787–798 (2021).
72. Aerts, H. J. et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **5**, 4006 (2014).
73. Hosny, A. et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med.* **15**, e1002711 (2018).
74. Lu, C. et al. A prognostic model for overall survival of patients with early-stage non-small cell lung cancer: a multicentre, retrospective study. *Lancet Digit. Health* **2**, e594–e606 (2020).
75. Zhong, Y. et al. Deep learning for prediction of N2 metastasis and survival for clinical stage I non-small cell lung cancer. *Radiology* **302**, 200–211 (2022).
76. Das, N. et al. Collaboration between explainable artificial intelligence and pulmonologists improves the accuracy of pulmonary function test interpretation. *Eur. Respir. J.* **61**, 2201720 (2023).
77. Sadeghi, A. H. et al. Virtual reality and artificial intelligence for 3-dimensional planning of lung segmentectomies. *JTCVS Tech.* **7**, 309–321 (2021).
78. Li, X. et al. Accuracy and efficiency of an artificial intelligence-based pulmonary broncho-vascular three-dimensional reconstruction system supporting thoracic surgery: retrospective and prospective validation study. *EBioMedicine* **87**, 104422 (2023).
79. Kadomatsu, Y., Nakao, M., Ueno, H., Nakamura, S. & Chen-Yoshikawa, T. F. A novel system applying artificial intelligence in the identification of air leak sites. *JTCVS Tech.* **15**, 181–191 (2022).
80. Court, L. E. The Radiation Planning Assistant: addressing the global gap in radiotherapy services. *Lancet Oncol.* **25**, 277–278 (2024).
81. Choi, W. et al. Novel functional radiomics for prediction of cardiac positron emission tomography avidity in lung cancer radiotherapy. *JCO Clin. Cancer Inform.* **8**, e2300241 (2024).
82. Bourbonne, V. et al. Radiomics analysis of 3D dose distributions to predict toxicity of radiotherapy for lung cancer. *Radiother. Oncol.* **155**, 144–150 (2021).
83. Wakelee, H. et al. Perioperative pembrolizumab for early-stage non-small-cell lung cancer. *N. Engl. J. Med.* **389**, 491–503 (2023).
84. Dercle, L. et al. Baseline radiomic signature to estimate overall survival in patients with NSCLC. *J. Thorac. Oncol.* **18**, 587–598 (2023).
85. Zhao, J. et al. Assessing treatment outcomes of chemoimmunotherapy in extensive-stage small cell lung cancer: an integrated clinical and radiomics approach. *J. Immunother. Cancer* **11**, e007492 (2023).
86. Vaidya, P. et al. Novel, non-invasive imaging approach to identify patients with advanced non-small cell lung cancer at risk of

- hyperprogressive disease with immune checkpoint blockade. *J. Immunother. Cancer* **8**, e001343 (2020).
87. Mu, W. et al. Radiomics predicts risk of cachexia in advanced NSCLC patients treated with immune checkpoint inhibitors. *Br. J. Cancer* **125**, 229–239 (2021).
  88. Saad, M. B. et al. Predicting benefit from immune checkpoint inhibitors in patients with non-small-cell lung cancer by CT-based ensemble deep learning: a retrospective study. *Lancet Digit Health* **5**, e404–e420 (2023).
  89. Assaf, Z. J. F. et al. A longitudinal circulating tumor DNA-based model associated with survival in metastatic non-small-cell lung cancer. *Nat. Med.* **29**, 859–868 (2023).
  90. Wei, F. et al. Machine learning for prediction of immunotherapeutic outcome in non-small-cell lung cancer based on circulating cytokine signatures. *J. Immunother. Cancer* **11**, e006788 (2023).
  91. Wu, J., Mayer, A. T. & Li, R. Integrated imaging and molecular analysis to decipher tumor microenvironment in the era of immunotherapy. *Semin Cancer Biol.* **34**, 310–328 (2022).
  92. Planchard, D. et al. Osimertinib with or without chemotherapy in EGFR-mutated advanced NSCLC. *N. Engl. J. Med.* **389**, 1935–1948 (2023).
  93. Le, X. et al. A Multicenter Open-Label Randomized Phase II Study of Osimertinib With and Without Ramucirumab in Tyrosine Kinase Inhibitor-Naive EGFR-Mutant Metastatic Non-Small Cell Lung Cancer (RAMOSE trial). *J. Clin. Oncol.* **43**, 403–411 (2024).
  94. Song, J. et al. Development and validation of a machine learning model to explore tyrosine kinase inhibitor response in patients with stage IV EGFR variant-positive non-small cell lung Cancer. *JAMA Netw. open* **3**, e2030442 (2020).
  95. Wang, S. et al. Mining whole-lung information by artificial intelligence for predicting EGFR genotype and targeted therapy response in lung cancer: a multicohort study. *Lancet Digit. Health* **4**, e309–e319 (2022).
  96. Christie, J. R. et al. Artificial intelligence in lung cancer: bridging the gap between computational power and clinical decision-making. *Can. Assoc. Radio. J.* **72**, 86–97 (2021).
  97. Ankolekar, A. et al. Clinician perspectives on clinical decision support systems in lung cancer: Implications for shared decision-making. *Health Expect.* **25**, 1342–1351 (2022).
  98. Liu, C. et al. Using artificial intelligence (Watson for Oncology) for treatment recommendations amongst chinese patients with lung cancer: feasibility study. *J. Med. Internet Res.* **20**, e11087 (2018).
  99. Kim, M. S. et al. Artificial intelligence and lung cancer treatment decision: agreement with recommendation of multidisciplinary tumor board. *Transl. Lung Cancer Res.* **9**, 507–514 (2020).
  100. Gerwing, M. et al. The beginning of the end for conventional RECIST – novel therapies require novel imaging approaches. *Nat. Rev. Clin. Oncol.* **16**, 442–458 (2019).
  101. Li, Y. et al. Noninvasive radiomic biomarkers for predicting pseudoprogression and hyperprogression in patients with non-small cell lung cancer treated with immune checkpoint inhibition. *Oncoimmunology* **13**, 2312628 (2024).
  102. Arbour, K. C. et al. Deep learning to estimate RECIST in patients with NSCLC treated with PD-1 blockade. *Cancer Discov.* **11**, 59–67 (2021).
  103. Tran, H. T. et al. Circulating tumor DNA and radiological tumor volume identify patients at risk for relapse with resected, early-stage non-small-cell lung cancer. *Ann. Oncol.* **35**, 183–189 (2024).
  104. Ding, H., Xu, X. S., Yang, Y. & Yuan, M. Improving prediction of survival and progression in metastatic non-small cell lung cancer after immunotherapy through machine learning of circulating tumor DNA. *JCO Precis. Oncol.* **8**, e2300718 (2024).
  105. Ayers, J. W. et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern. Med.* **183**, 589–596 (2023).
  106. Ueda, D. et al. ChatGPT's diagnostic performance from patient history and imaging findings on the diagnosis please quizzes. *Radiology* **308**, e231040 (2023).
  107. Haver, H. L., Lin, C. T., Sirajuddin, A., Yi, P. H. & Jeudy, J. Use of ChatGPT, GPT-4, and Bard to improve readability of ChatGPT's answers to common questions on lung cancer and lung cancer screening. *Am. J. Roentgenol.* **221**, 701–704 (2023).
  108. Thirunavukarasu, A. J. et al. Large language models in medicine. *Nat. Med.* **29**, 1930–1940 (2023).
  109. Chow, R. et al. Use of artificial intelligence for cancer clinical trial enrollment: a systematic review and meta-analysis. *J. Natl. Cancer Inst.* **115**, 365–374 (2023).
  110. Warraich, H. J., Tazbaz, T. & Califf, R. M. FDA Perspective on the regulation of artificial intelligence in health care and biomedicine. *JAMA* **333**, 241–247 (2025).
  111. Mello, M. M. & Cohen, I. G. Regulation of health and health care artificial intelligence. *JAMA* **333**, 1769–1770 (2025).
  112. Lawlor, R. T. The impact of GDPR on data sharing for European cancer research. *Lancet Oncol.* **24**, 6–8 (2023).
  113. Warnat-Herresthal, S. et al. Swarm learning for decentralized and confidential clinical machine learning. *Nature* **594**, 265–270 (2021).
  114. Ogier du Terrail, J. et al. Federated learning for predicting histological response to neoadjuvant chemotherapy in triple-negative breast cancer. *Nat. Med.* **29**, 135–146 (2023).
  115. Pati, S. et al. Federated learning enables big data for rare cancer boundary detection. *Nat. Commun.* **13**, 7346 (2022).
  116. Feng, B. et al. Robustly federated learning model for identifying high-risk patients with postoperative gastric cancer recurrence. *Nat. Commun.* **15**, 742 (2024).
  117. Haggemuller, S. et al. Federated learning for decentralized artificial intelligence in melanoma diagnostics. *JAMA Dermatol.* **160**, 303–311 (2024).
  118. Jochems, A. et al. Developing and validating a survival prediction model for NSCLC patients through distributed learning across 3 countries. *Int. J. Radiat. Oncol. Biol. Phys.* **99**, 344–352 (2017).
  119. Hsu, W. et al. External validation of an ensemble model for automated mammography interpretation by artificial intelligence. *JAMA Netw. Open* **5**, e2242343 (2022).
  120. Potter, A. L. et al. Persistent race- and sex-based disparities in lung cancer screening eligibility. *J. Thorac. Cardiovasc. Surg.* **168**, 248–260 e242 (2024).
  121. Allen, M. S., Harmsen, W. S., Mandrekar, J. & Rocco, G. Bias against complex lung cancer surgery. *Ann. Thorac. Surg.* **112**, 1824–1831 (2021).
  122. McKinney, S. M. et al. International evaluation of an AI system for breast cancer screening. *Nature* **577**, 89–94 (2020).
  123. Dickson, J. L. et al. Uptake of invitations to a lung health check offering low-dose CT lung cancer screening among an ethnically and socioeconomically diverse population at risk of lung cancer in the UK (SUMMIT): a prospective, longitudinal cohort study. *Lancet Public Health* **8**, e130–e140 (2023).
  124. Chen, R. J. et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat. Biomed. Eng.* **7**, 719–742 (2023).
  125. Duffy, G. et al. Confounders mediate AI prediction of demographics in medical imaging. *NPJ Digit. Med.* **5**, 188 (2022).
  126. Wang, F., Kaushal, R. & Khullar, D. Should health care demand interpretable artificial intelligence or accept “black box” medicine?. *Ann. Intern. Med.* **172**, 59–60 (2020).
  127. Wulczyn, E. et al. Interpretable survival prediction for colorectal cancer using deep learning. *NPJ Digit. Med.* **4**, 71 (2021).
  128. Zhao, W. et al. Convolution kernel and iterative reconstruction affect the diagnostic performance of radiomics and deep learning in lung adenocarcinoma pathological subtypes. *Thorac. cancer* **10**, 1893–1903 (2019).
  129. Zwanenburg, A. et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* **295**, 328–338 (2020).

130. Whybra, P. et al. The image biomarker standardization initiative: standardized convolutional filters for reproducible radiomics and enhanced clinical insights. *Radiology* **310**, e231319 (2024).
131. Huang, E. P. et al. Criteria for the translation of radiomics into clinically useful tests. *Nat. Rev. Clin. Oncol.* **20**, 69–82 (2023).
132. Hernandez-Boussard, T., Bozkurt, S., Ioannidis, J. P. A. & Shah, N. H. MINIMAR (MINimum Information for Medical AI Reporting): developing reporting standards for artificial intelligence in health care. *J. Am. Med. Inform. Assoc.* **27**, 2011–2015 (2020).
133. Cruz Rivera, S. et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat. Med.* **26**, 1351–1363 (2020).
134. Liu, X. et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. *BMJ* **370**, m3164 (2020).
135. Castelo-Branco, L. et al. ESMO guidance for reporting oncology real-world evidence (GROW). *Ann. Oncol.* **34**, 1097–1112 (2023).
136. Lu, M. Y. et al. A multimodal generative AI copilot for human pathology. *Nature* **634**, 466–473 (2024).

### Author contributions

E.Z.: Conceptualization, Writing—original draft, Writing—review and editing. A.M.: Writing—original draft, Writing—review and editing, Visualization. J.Z.: Conceptualization, Writing—review and editing. Y.X.: Writing—review and editing. X.L.: Writing—review and editing. C.Z.: Conceptualization, Writing—review and editing. J.V.H.: Conceptualization, Writing—review and editing. J.W.: Conceptualization, Writing—review and editing, Supervision. X.L.: Conceptualization, Writing—review and editing, Supervision.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to Jia Wu or Xiuning Le.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025