ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN SCIENCE FOR OPTIMAL CANCER CARE

REVIEW

The landscape of conventional and artificial intelligence-based clinical prediction models in non-small-cell lung cancer: from development to real-world validation

H. R. Howard[1†], M. Hasanova[2,3†], A. Tiwari[4,5†], A. Ghose[3,5,6,7,8*], R. Winayak[9], T. Nahar[10], V. Chauhan[5,11], S. Arun[12], K. Palmer[13], A. Houston[14,15], N. S. Kumar[16,17], S. Lalwani[5,6], A. Januszewski[5,6], J. Conibear[5,6], A. Jain[18,19], A. Cortellini[20,21,22], A. Passaro[23], A. Addeo[24], V. Noronha[25‡], G. L. Banna[26,27‡] & S. Boussios[7,28,29,30,31,32,33‡]

[1]Cancer Institute, University College London, London; [2]Division of Biosciences, University College London, London; [3]OncoFlow™, London; [4]Cancer Services, Princess Alexandra Hospital NHS Trust, Harlow; [5]Barts Cancer Institute, Queen Mary University of London, London; [6]Barts Cancer Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London; [7]Department of Research and Innovation, Medway NHS Foundation Trust, Gillingham, UK; [8]Health Systems and Treatment Optimisation Network, European Cancer Organisation, Brussels, Belgium; [9]Kent Oncology Centre, Maidstone and Tunbridge Wells NHS Trust, Maidstone; [10]School of Medicine Dentistry and Biomedical Sciences, Queen's University, Belfast; [11]Department of General Internal Medicine, Royal Infirmary of Edinburgh, NHS Lothian, Edinburgh; [12]Centre for Cancer Screening, Prevention and Early Diagnosis, Wolfson Institute of Population Health, Queen Mary University of London, London; [13]Department of Clinical Oncology, University College London Hospitals NHS Foundation Trust, London; [14]Barts Life Sciences, Barts Health NHS Trust, London; [15]Digital Environmental Research Institute, Queen Mary University of London, London; [16]College of Life Sciences, University of Leicester, Leicester; [17]National Medical Research Association, London, UK; [18]Apollo Cancer Centre, Indraprastha Apollo Hospital, New Delhi, India; [19]Deanesly Centre Cancer Services, New Cross Hospital, Royal Wolverhampton NHS Trust, Wolverhampton, UK; [20]Operative Research Unit of Medical Oncology, Fondazione Policlinico Universitario Campus Bio-Medico, Rome; [21]Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Rome, Italy; [22]Department of Surgery and Cancer, Hammersmith Hospital Campus, Imperial College London, London, UK; [23]Division of Thoracic Oncology, European Institute of Oncology IRCCS, Milan, Italy; [24]Oncology Service, Geneva University Hospital, Geneva, Switzerland; [25]Department of Medical Oncology, Tata Memorial Hospital, Homi Bhabha National Institute, Mumbai, India; [26]Department of Medical Oncology, Portsmouth Hospitals University NHS Trust, Portsmouth; [27]Faculty of Science and Health, School of Pharmacy and Biomedical Sciences, University of Portsmouth, Portsmouth, UK; [28]Department of Medical Oncology, Ioannina University Hospital, Ioannina; [29]Faculty of Medicine, School of Health Sciences, University of Ioannina, Ioannina; [30]AELIA Organization, Thessaloniki, Greece; [31]Kent and Medway Medical School, University of Kent, Canterbury; [32]Faculty of Medicine, Health and Social Care, Canterbury Christ Church University, Canterbury; [33]Faculty of Life Sciences and Medicine, School of Cancer & Pharmaceutical Sciences, King's College London, London, UK

Available online xxx

Globally, lung cancer remains the most common cause of cancer mortality, with non-small-cell lung cancer (NSCLC) being the most common subtype of lung cancer diagnosed. This review paper provides a comprehensive landscape of clinical prediction models (CPMs) in NSCLC, including in early-stage and metastatic disease, and the recent acceleration of artificial intelligence integration. Prediction models are developed using multimodal patient data to allow oncologists to make evidence-based decisions regarding patient treatment options. Despite these models in early-stage and metastatic NSCLC showing promise, their clinical application provides challenges, involving an unmet need for external validation, alongside a lack of prospective modelling. However, the continued advancements in this field, comprising production and accessibility of large-scale pathology databases and external validation of developed models, allow for continued research and progress. These models have potential to assist in personalised treatment selection, supporting oncologists in perceiving future risk factors or issues associated with a specific targeted therapy for an individual patient, ultimately optimising treatment to precise, personalised options for individuals diagnosed with NSCLC.

Key words: clinical prediction, predictive model, lung cancer, cancer prognosis, artificial intelligence, NSCLC

*Correspondence to: Dr Aruni Ghose, Barts Cancer Centre, St Bartholomew's Hospital, Barts Health NHS Trust, West Smithfield, London EC1A 7BE, UK; Barts Cancer Institute, Queen Mary University of London, London, UK; OncoFlow™, London, UK; Department of Research and Innovation, Medical Oncology, Medway NHS Foundation Trust, Gillingham, UK; Health Systems and Treatment Optimisation Network, European Cancer Organisation, Brussels, Belgium; Oncology Council, Royal Society of Medicine, UK. Tel: +447570839307

E-mail: aruni.ghose1@gmail.com (A. Ghose).

†These authors contributed equally.

‡Joint senior authors.

INTRODUCTION

Recording 1.8 million deaths in 2022, lung cancer (LC) remains the most common cause of cancer-related mortality globally.[1] Despite improvements in detection and survival, outcomes remain poor.[1]

Non-small-cell lung cancer (NSCLC) is the most common variety of LC.[2] It can be differentiated into two major histological subtypes: lung adenocarcinoma (LUAD) and lung squamous-cell carcinoma. Surgery is the standard of care in the early setting with or without neoadjuvant chemotherapy, chemoradiotherapy, or chemoimmunotherapy for resectable

disease whereas the mainstay for unresectable stages includes chemoradiotherapy with/without adjuvant immunotherapy.[3] The current first-line treatment landscape in the advanced or metastatic NSCLC (mNSCLC) setting consists of immune checkpoint inhibitors (ICIs) as monotherapy or with platinum-based chemotherapy as combination therapy depending on the patient's tumour programmed death-ligand 1 (PD-L1) expression.[4,5]

Over the past decade there has been an increase in the number of published clinical prediction models (CPMs) in the oncological literature. Despite their development using real-world patient data, translation to the clinic to aid and simplify decision making is limited.[6]

In the early NSCLC setting, the tumour—node—metastasis (TNM) staging system has been used for predicting patient outcomes and optimising treatment algorithms after surgical resection.[7] Recently, a few immune nutrition indices formulated to do the same from blood biomarkers including neutrophil/lymphocyte or platelet/lymphocyte or monocyte/lymphocyte ratios.[8] Following the advent of immunotherapy in the metastatic setting, prominent predictive tools included neutrophil/lymphocyte ratio, lactate dehydrogenase (LDH), and PD-L1 expression levels.[9,10] This transitioned to prediction models like the lung immuno-oncology prognostic score (LIPS-3) and NHS lung score in the context of first-line immuno and chemoimmunotherapy, respectively.[11-13]

Our comprehensive review article looks at the landscape of both early and advanced NSCLC CPMs from inception to present day, along with the usage of artificial intelligence (AI) modalities to further automate them. We carried out a comprehensive search in PubMed using keywords such as 'non-small cell lung cancer', 'artificial intelligence', 'clinical prediction models', 'early-stage', and 'metastatic'. The inclusion criteria covered papers published in recent years, primarily after 2020, and included exclusively original research articles.

PREDICTION MODELS IN EARLY-STAGE NSCLC

There are few known factors that can predict the choice of treatment. Current European Society for Medical Oncology guidelines for patients with early NSCLC stages I-IIIA focus on curative methods through a complete surgical resection.[14] Adjuvant chemotherapy is offered to patients with stage IIB and III NSCLC.[14] For patients who are medically inoperable or decline surgery, radical stereotactic ablative radiotherapy should be offered. The presence of certain tumour mutations can direct targeted therapies in the context of tyrosine kinase inhibitors for epidermal growth factor receptor (EGFR) mutations, as an example. LC prediction remains undeveloped, and the risk stratification is based on TNM staging. As emphasised in the literature, a single predictive model would allow for closer post-operative surveillance.[15] For example, we would be able to identify patients with aggressive stage I NSCLC who would benefit from adjuvant chemotherapy, which is not offered routinely in current practice. Conversely, those with low-risk disease could be prevented from receiving toxicity-associated therapies.

Key studies investigating clinical predictive factors in NSCLC are summarised in Table 1. A multicentre prospective study based in Central and Eastern Europe included 2052 patients with stage I-IIIA NSCLC from nine centres between 2007 and 2016 and followed up with them annually through 2020.[16] The hazard ratio (HR) of all-cause mortality increased with age, male sex, weight loss >5%, current smoking, and regular alcohol consumption. Individuals diagnosed at stage IIIA had a five and a half times more likely chance of death [HR 5.54 (4.10-7.48)]. Tumour histology, education level, and chronic obstructive pulmonary disease were not associated with a significant difference in survival rate. However, studies have shown that the survival rate for resected squamous-cell carcinoma differs from that of adenocarcinoma. Some studies have indicated that squamous-cell carcinoma histology was an independent favourable predictive factor.[17,18] However, the predictive value of each histology type is still being debated due to the continuously adapting criteria used to differentiate the types.

Approximately 25%-30% of all NSCLCs are squamous-cell lung cancer (SQLC).[19] Research on identifying relevant clinicopathological variables in SQLC still requires significant efforts. One study investigated the clinicopathological predictive variables for resected SQLC exploring the impact of adjuvant and neoadjuvant treatment.[19] SQLC cases from 494 patients were retrospectively analysed to create a Cox model to identify risk classes using two different methods. Model A had two risk classes while model B had three. The predicted outcomes were overall survival (OS), cancer-specific survival (CSS), and disease-free survival (DFS). Significant predictors following multivariate analyses included age >68 years ($P < 0.0001$), T-descriptor according to TNM classification ($P < 0.0001$), lymph nodes (1-2 versus 3-4; $P < 0.0001$), lymph node status (positive versus negative; P 0.0001), and grading (1-2-3; $P < 0.0001$).[19]

Histopredictive models have been developed aiming to further characterise risk based on structural morphology. A multicentre, retrospective study thoroughly investigated the role of diversity of nuclear morphology in the epithelium region which was referred to as CellDiv features.[20] Features were broadly classified into nuclear shape and intensity. Patients predicted as high-risk based on the CellDiv model in the independent data validation set had significantly worse 5-year OS [LUAD = HR 1.62 (95% CI 1.15-2.3) $P = 0.0058$; lung squamous cell carcinoma = HR 2.24 (95% CI 1.04-4.80) $P = 0.039$]. The CellDiv features also showed an association with certain anti-apoptotic signalling mechanisms. Another study utilised histological assessment alongside patient, demographic, and clinical factors to devise a model.[21] Three risk factors were identified: T2—the size and extent of the primary tumour; presence of necrosis, and a pathological architectural score of 5-6. While histology-based models may play an important role in guiding treatment decisions and prognostication in LC, they have significant limitations related to tissue

Table 1. Landscape of clinical prediction models in early non-small-cell lung cancer

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination
Sheikh et al. 2023[16]	Prospective	2052	Stage I-IIIA NSCLC	Cox proportional hazard model	Overall survival, progression-free survival	Demographic and clinical features	Study location, age at diagnosis, sex, education level, COPD, chronic diseases, weight loss in the past 2 years, smoking, regular alcohol drinking, tumour histology, tumour stage	Older age, alcohol drinkers, high tumour stage, males, weight loss associated with worse outcomes	N/A
Pilotto et al. 2015[19]	Retrospective	573	Resected squamous-cell carcinoma	Cox proportional hazard model	Disease-free/ cancer-specific/ overall survival	Clinicopathological data	Age, T-descriptor according to TNM, nodes, and grading	Age $\leq$68 years, T-descriptor according to TNM, 1-2 negative nodes, grading 1-2	Disease-free survival Low risk = 64.6% Intermediate risk = 39.8% High risk = 21.8% Cancer-specific survival Low risk = 84.4% Intermediate risk = 55.4% High risk = 30.9% Overall survival Low risk = 77.3% Intermediate risk = 47.9% High risk = 27.2%
Lu et al. 2020[20]	Retrospective	1057	LUAD, LUSC	Cox proportional hazards model	5-year overall survival	CellDiv: Histopathological features quantifying morphological diversity	11 nuclear features Shape: area, eccentricity, solidity, circularity, major/ minor axis length of best-fit eclipse Appearance: mean intensity, intensity range, mean inside/ outside boundary intensity, boundary, saliency	LUAD: among features related to nuclear shape, minor axis length of best-fit eclipse had the strongest weighting; among features related to nuclear intensity, mean intensity, and mean inside boundary intensity had the strongest positive correlation, while mean inside boundary had the strongest negative correlation. LUSC: major axis length, solidity, and circularity had the strongest weighting; mean intensity and outside boundary intensity had the strongest weighting	PR-AUC LUSC-TCGA: CellDiv = 0.89 Clinical variables = 0.84 Both = 0.85 LUSC-UB: CellDiv = 0.63 Clinical = 0.52 Both = 0.64 LUAD-TCGA: CellDiv = 0.72 Clinical = 0.77 Both = 0.80 Overall (CellDiv) = 0.68 $\pm$ 0.01

Continued

https://doi.org/10.1016/j.esmoop.2025.105557

Table 1. Continued

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination
Liu et al. 2019[21]	Retrospective	198	Stage 1 lung adenocarcinoma	Cox regression model	5-year overall survival	Mixed clinicopathological features	Age, sex, smoking history, T stage (T1/T2), Architectural score (3-4, 5-6), lymphovascular invasion, neuron invasion, necrosis, type of surgery, adjuvant chemotherapy	Necrosis, infiltration of neutrophils, T stage, Architectural score	5-year overall survival Low risk (score $\leq$103): 88.0% High risk (score $\geq$103): 49.0%
Pilotto et al. 2018[22]	Retrospective	1375	Resected squamous-cell carcinoma	Cox univariate model	Disease-free survival/cancer-specific survival/ overall survival	Clinicopathological features	Age, T-descriptor of the TNM classification, lymph node status, grading	Age, T-descriptor of the TNM classification, lymph nodes (1-2 versus 3-4), lymph node status (positive/ negative), grading	Disease-free survival Low risk = 51.0% Intermediate risk = 33.5% High risk = 25.8% Cancer-specific survival Low risk = 82.7% Intermediate risk = 64.7% High risk = 53.3% Overall survival Low risk = 56.7% Intermediate risk = 37.9% High risk = 30.9%
Li et al. 2017[23]	Retrospective	2414	Non-squamous NSCLC with clinical annotations	Cox proportional hazards regression model	Overall survival	Clinical and pathological features	Age, Stage, IRGPI risk score (25 gene pairs)	Meta-training: stage, grade, immune risk In meta-testing: age, stage, immune risk DCC: Age, stage, immune risk GSE30219: age, stage, immune risk TCGA: stage, immune risk	Overall survival IRGPI in stage-1 of meta-testing: Low risk: ~73% High risk: ~39% IRGPI in stage II of meta-testing: Low risk: ~39% High risk: ~19% IRGPI in stage I of independent validation: Low risk: ~35% High-risk: ~8% IRGPI in stage II of independent validation: Low risk: ~55% High risk: ~14%

Continued

Table 1. Continued

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination
Wu et al. 2020[24]	Retrospective	1091	Early-stage LUAD patients with clinical information	Multivariate Cox regression analysis	Overall survival	Clinical and pathological features	Age, sex, smoking, stage, TNM scoring, survival status, ALK fusion, EGFR mutation, KRAS mutation, TP53 mutation, STK11 mutation, MYC copy, 21 prognostic genes	GSE30219: age, risk score GSE31210: age, stage, risk score GSE50081: stage, risk score Combined dataset: age, sex, stage, risk score	GSE30219 survival probability: High risk: ~0.10 Low risk: ~0.25 GSE31210 survival probability: High risk: ~0.55 Low risk: ~9.97 GSE50081 survival probability: High risk: ~0.24 Low risk: ~0.64 Combined dataset survival probability: High risk: ~0 Low risk: ~0.36

AUC, area under the curve; COPD, chronic obstructive pulmonary disease; DCC, Director's Challenge Consortium; IRGPI, immune related gene prognostic index; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; N/A, not applicable; NSCLC, non-small-cell lung cancer; PR, precision recall; TCGA, The Cancer Genome Atlas; TNM, tumour–node–metastasis; UB, University of Bern.

availability, interobserver bias, varying therapeutic options based on subtype, and tumour heterogeneity; however, integrating histological classification and molecular profiling can help overcome some of these limitations.

While predictive models for early-stage NSCLC show promise, factors such as modest improvement in survival prediction and the lack of routine measurement of clinical variables may hinder widespread adoption. Further prospective studies are needed to refine these models and assess their practical utility in guiding personalised treatment approaches.

PREDICTION MODELS IN METASTATIC NSCLC

In the mNSCLC setting, several prediction models have been developed to predict patients' response to therapy using genetic and clinical data, as highlighted in Table 2. The use of prediction models utilising gene expression signatures in mNSCLC has recently gained traction. Chen et al. developed a model that significantly predicted OS in NSCLC patients by studying NSCLC cells under hypoxic and non-hypoxic conditions and identifying 17 genes strongly related to metastatic markers including epigenetic modification, epithelial—mesenchymal transition, glucose metabolism, and hypoxia. They showed that these genes were able to risk stratify patients with NSCLC and significantly predicted each group's OS.[25]

Furthermore, gene expression models have also been developed by studying transcriptomics from online databases, including The Cancer Genome Atlas (TCGA) program and Gene Expression Omnibus (GEO). Ding et al. evaluated differentially expressed genes (DEGs) between metastatic and non-metastatic NSCLC patients by identifying these DEGs in the TCGA and then assessing their function using platforms including Gene Ontology and Kyoto Encyclopaedia of Genes and Genomes. In total, they identified 2058 DEGs, whose functions included development and differentiation of epidermal cells and keratinocytes, features closely related to metastasis. Moreover, they developed a six-gene panel from these DEGs to classify NSCLC patients as low- or high-risk for metastasis, with significant OS differences in either group.[26]

Similarly, Wang et al. developed a genetic-based model in NSCLC patients to predict metastasis, focusing specifically on brain metastases, an area with poor prognoses and limited effective prediction models. They used the GEO platform to identify 472 cancer-related DEGs in patients with NSCLC and brain metastases. Of these, they narrowed the gene signature down to 11 to create a RiskScore model for these patients. Their model predicted a significant difference in OS between patients with high and low RiskScore. Interestingly, both gene expression models highlighted a significant immunophenotypic difference between high- and low-risk groups, with the high-risk groups exhibiting a greater concentration of immune cells including tumour-associated macrophages (TAMs), dendritic cells, and monocytes. Paradoxically, immune cell concentration at tumour tissue has been associated with tumour metastasis, possibly due to

ineffective penetration and tumour killing. In addition, immune cells themselves can be immunosuppressive: TAMs, for instance, contribute to an immunosuppressive tumour microenvironment by producing cytokines, chemokines causing checkpoint inhibition in T cells. Clinically, these findings may be applied to predict as for whom immunotherapy would be beneficial in NSCLC patients, and to model their response.[26,27]

In addition to gene expression models, clinical models have also been developed. One such example is the model developed by Ganti et al. which sought to predict OS in a population of geriatric patients with mNSCLC, a population for whom therapy options are oftentimes withheld, and a specific CPM had never been developed. They retrospectively analysed clinical data from 38 National Cancer Institute trials between 1991 and 2011, to carry out a multivariate analysis on 1467 patients of at least 70 years of age with mNSCLC. They identified four independent prognostic factors and assigned weighted points to each: male sex (+3), Eastern Cooperative Oncology Group (ECOG) performance status (PS) = 1 (+3), PS = 2 (+8), stage IV at presentation (+11), and weight loss (+4). Patients were then classified as good or poor prognosis according to their score: $\leq$8 and $\geq$9, respectively. Their model predicted a significant difference in OS between the two patient groups.[28]

Liu et al. developed a similar model to predict OS in mNSCLC patients, focusing on patients receiving chemoradiotherapy. They carried out a multivariate analysis on 243 mNSCLC patients who had previously received chemoradiotherapy and similarly identified male sex and Karnofsky PS <80 (equivalent to ECOG PS >1) as some of the similar poor prognostic risk factors. In total, their model contained six independent poor prognostic factors: male sex, Karnofsky PS <80, four or more chemotherapy cycles, haemoglobin level <120 g/l, neutrophil count $>5.8 \times 10^9$/l, and platelet count $>220 \times 10^9$/l. According to these risk factors, they divided the patient cohort into three groups stratified by risk: low, medium, and high. Between each, they showed a significantly different OS during the first measured 3 years.[29] Both these CPMs highlight male sex and poor PS as poor prognostic risk factors, findings that have been well demonstrated in previous studies.[17,18]

In addition, models have been developed to incorporate haematological cell counts as prognostic risk factors to predict OS in mNSCLC patients. This is especially seen in the setting of immunotherapy. The LIPS-3 is one such example where Banna et al. generated a model to predict OS in 784 mNSCLC patients given first-line pembrolizumab, with a molecular baseline PD-L1 expression level $\geq$50%. The model evaluated a mix of patients (age, sex, smoking history, body mass index) and clinical laboratory risk factors, including neutrophil-to-lymphocyte ratio (NLR), LDH level, and PD-L1 expression level. They highlighted three risk factors as significant prognostic markers for OS: NLR <4 (1-year OS of 76.6% versus 44.8% with NLR >4 in the validation cohort; HR 2.29), ECOG PS 2 (HR 2.04), and pretreatment steroid therapy (HR 1.67), creating a LIPS-3 tool.[12]

Similarly, The Lung Immune Therapy Evaluation (LITE) Risk generated by Navani et al. utilised NLR as a prognostic marker to predict OS in 495 mNSCLC patients treated with ICI monotherapy. The risk factors included in their model were ECOG PS status, raised LDH level, and NLR $\geq$3. The training cohort looked at patients receiving first-line pembrolizumab, while the test cohort at second line with either nivolumab or atezolizumab. Patients were divided into three groups based on their risk score: favourable (score $\leq$1), intermediate (score 1-2), and poor (score $\geq$3). Using the LITE Risk score, they predicted a median OS of 28.3 months, 9.1 months, and 2.1 months for the favourable, intermediate-, and poor-risk groups, respectively. Similarly, they showed that the differences in OS were significantly different between the groups, with an HR of 2.08 between intermediate and favourable and an HR of 5.21 between poor and favourable in the test cohort.[30] Another haematological marker used as part of cancer prognostic tools is the derived NLR (dNLR), which has shown prior utility in prognostic tools for advanced melanoma patients on pembrolizumab and ipilimumab.[31,32] Slightly different from the NLR, the dNLR offers a more detailed view of leukocyte populations, by calculating the ratio between neutrophils and other white blood cell populations, including granulocytes and monocytes. The lung immune prognostic index (LIPI) created by Mezquita et al. included baseline dNLR as well as LDH as prognostic tools to predict immunoresistance in 466 mNSCLC patients treated with programmed cell death protein 1 (PD-1)/PD-L1 inhibitors. The LIPI score generated three risk groups (poor, intermediate, and good) with significantly different median OS: 3 months, 10 months, and 34 months, respectively. They showed that mNSCLC patients with poor prognosis to ICI treatment had poor baseline LIPI scores: dNLR >3 and an LDH value greater than the upper limit of normal independently significantly predicted OS for patients treated with immunotherapy compared with a chemotherapy-treated control group (HR 2.22 and 2.51, respectively).[33] Compared with the LIPS-3 study which was conducted on patients on first-line pembrolizumab, the LITE Risk and LIPI used both first- and subsequent-line ICI monotherapies and significantly predicted for each, broadening their clinical utility.

While the LITE Risk, LIPS-3, and LIPI tools can be applied to mNSCLC patients on immunotherapy alone, they may not accurately predict OS in those patients receiving chemoimmunotherapy. The recent KEYNOTE-407 study showed that the addition of pembrolizumab to platinum-based chemotherapy in metastatic squamous NSCLC patients resulted in an increased OS of 17.1 months versus 11.6 months in the placebo group.[35] Despite the benefit of chemoimmunotherapy in patients with mNSCLC, few CPMs have been developed. However, the recent Spinnaker retrospective study sought to address this gap by analysing 308 mNSCLC patients treated with chemo-immunotherapy. As part of their study, they confirmed a

similar progression-free survival (PFS) of 8.0 months compared with other trials investigating chemo-immunotherapy in the mNSCLC setting. Moreover, they identified a high metastatic burden (three or more sites), squamous histology, and a systemic immune-inflammatory index (SII, an immune activity marker considering NLR) $\geq$1444 as poor prognostic markers for survival. They were able to differentiate three groups based on these markers to predict for OS and PFS.[13]

Chen et al.[34] presented the Lung Cancer Immunotherapy-Radiomics Prediction Vector (LCI-RPV), designed to predict treatment response in patients undergoing PD-1 or PD-L1 checkpoint inhibitor immunotherapy. A discovery cohort of 85 patients with histologically confirmed NSCLC using contrast-enhanced computed tomography (CT) and RNA sequencing from public databases, and with CD274 gene expression, was used to train the model. Following segmentation of the tumour and surrounding lung tissue, 1998 radiomic features were extracted and narrowed to 1647 reproducible features based on intraclass correlation. Linear regression and elastic net regularisation were used to produce a 15-feature composite signature integrating texture, wavelet, and fractal metrics from the tumour, peritumoural annulus, and background lung parenchyma. LCI-RPV was validated on two independent UK-based cohorts: the Imperial College Healthcare NHS Trust (ICHNT, $n = 66$) and the Lung Cancer Western European Study (LCWES, $n = 43$) of NSCLC patients treated with immunotherapy demonstrating moderate discrimination for predicting PD-L1 expression [area under the curve (AUC) = 0.70], high expression thresholds (AUC = 0.72 for >50% PD-L1), 3-month treatment response (AUC = 0.68), and pneumonitis occurrence (AUC = 0.64). The model stratified patients into high- and low-risk groups for 3-year OS (HR = 2.26 and 2.45 in ICHNT and LCWES, respectively, $P < 0.05$). Gene set enrichment analyses revealed LCI-RPV correlated positively with pathways related to inflammation, hypoxia, and apoptosis, while single-cell RNA sequencing linked key genes in the model to expression in tumour-infiltrating myeloid and T cells. By providing a noninvasive, imaging-based biomarker (CD274 expression), this model offers a tool for guiding immunotherapy decisions especially in radiotherapy-naive patients or when tissue biopsy is not feasible, paving a way for more personalised and risk-stratified care in advanced NSCLC.

In summary, the landscape of prediction models in mNSCLC is rapidly advancing, with newer, more detailed ones, based on specific patient cohorts and therapies. Coupled with emerging data on metastasis-associated gene expression profiles, clinicians can increasingly risk stratify patients and more precisely determine treatment selection.

AI INTEGRATION INTO NSCLC PREDICTION MODELS

More recently, AI has played an increasingly integral role in the development of prediction models and LC data analysis within the clinical setting. The complexity of NSCLC requires accuracy and efficiency in the decision-making process, integrating a multitude of information across patient positron emission tomography (PET) and CT scans, radiomics data, histopathology images, and biomolecular data, to arrive at a reliable and accurate model.

AI technology in LC prognosis can be broadly classified into machine learning (ML), deep learning (DL), and natural language processing (NLP).[35] ML allows a streamlined analysis of a patient's medical data and assists in the decision-making process. DL is a subset of ML that makes use of deep networks that achieve powerful learning capabilities and is ultimately able to make intelligent decisions on its own. NLP enables the computer to extract useful information from textual or verbal data, which expands the scope of AI to analysis of medical reports as well as big medical data. To compare AI models to standard approaches, metrics such as AUC scores are used to measure the overall performance of a model. It is given as a ratio between 0 and 1, and the closer the score is to 1, the more efficient the model. Alternatively, the concordance-index (C-index) is used, which is equal to the AUC for models with binary survival outcome variables.

AI for NSCLC prognostics has had significant research and application use cases (as shown in Table 3), looking at imaging data, radiomic features, genetic mutations, molecular interactions, and drug response, to accurately predict survival rates and treatment outcomes for NSCLC patients. Primitive models for stratifying survival in stage I adenocarcinoma identified three risk factors, namely T2, necrosis, and pathological architecture score, and combined them to derive a prognostic score stratification model, achieving an AUC = 0.717.[21] She et al. looked at DL survival neural network models based on a combinatorial TNM staging system across 17 322 patients (majority of cancers being stage I and adenocarcinoma), to predict survival and test treatment outcomes. The model was compared with the TNM system for lung cancer-specific survival, showing more promising results (AUC = 0.739) compared with the TNM stage of the test data (AUC = 0.706).[36] Another group took this a step further, comparing TNM staging to survival prediction, as well as validating individual adjuvant treatment recommendation plans based on three variations of an ML model.[37] Across a total of 4617 resected stage III NSCLC patients in the study, the DL model (C-index = 0.834) performed more stably and accurately than the random forest (C-index = 0.678) and Cox proportional hazards model (C-index = 0.640), and better than the TNM staging system (C-index = 0.650).[37]

A more recent study in preoperative and post-operative NSCLC prognostics compared two Cox models, incorporating retrospective data across 392 patients and 66 clinicopathological features, to identify patients who might benefit from perioperative anticancer therapy, achieving C-indices of 0.70 (pre-surgical) and 0.75 (post-surgical).[38] ML methods based on multivariable CPMs have also been tried for predicting recurrence and death after curative-intent radiotherapy. Based on a retrospective, multicentre study in 657 patients across five hospitals, risk stratification models for predicting recurrence, recurrence-free survival

(RFS), and OS were developed and benchmarked against TNM staging and PS. The models achieved an AUC for recurrence (0.687), RFS (0.682), and OS (0.759) superior to TNM stage [recurrence (0.670), RFS (0.650), and OS (0.649)] and PS [recurrence (0.506), RFS (0.464), and OS (0.459)].[39]

Image-based methods have also been tested for AI-powered prognostication of NSCLC. Huang et al. developed a radiomics biomarker for CT image analysis to estimate DFS in 282 patients with stage IA-IIB NSCLC, which demonstrated a more accurate classification of DFS [C-index = 0.72, 95% CI 0.71 - 0.73] compared with conventional prognostic methods (C-index = 0.691, 95% CI 0.68 - 0.70).[40] Other models have enabled the analysis of pre- and post-surgery CT scans, or a combination of radiomics, radiotherapy, and surgery data based on seven independent datasets, to achieve better performance and higher accuracy of prediction tools in radiomics and surgery, respectively (AUC = 0.70, 95% CI; AUC = 0.71, 95% CI) when benchmarked against random forest models based on clinical parameters (AUC = 0.55, 95% CI; AUC = 0.58, 95% CI).[41] A more recent study by Zhang et al. analysed 20 333 CT images of 6371 patients to assist traditional TNM staging, prognostication, and therapy personalisation based on the EGFR genotype in resected stage I-III patients, achieving a C-index of 0.817 in the external validation set.[42] One group has even proposed combining tissue metabolome with CT scan analysis to derive an NSCLC classification and prediction model.[43]

Oftentimes, molecular-level analysis, genomics, and predictive biomarkers have been used for AI-prognostic model development. Genetic alterations with prognostic significance in NSCLC primarily include EGFR gene mutations or ALK gene rearrangements.[35] Jiang et al. explored survival prognosis in LUAD using TCGA across 524 patients, showing an independent association between smoking history, tumour stage, and surgical margin resection status, with an AUC score of 0.712 and a prognostic accuracy of 75%.[44] A model developed by Coudray et al. focused on extracting mutation-related morphological features from TCGA and other public databases, to predict response to targeted treatment based on the presence of mutated genes; results suggest an accuracy much higher than previous models (AUC = 0.97 versus AUC = 0.75-0.83), and even comparable to results from pathologists.[45] Li et al. used a similar approach, incorporating TCGA and GEO data from a large cohort to generate a prognostic model based on 16 genes for classifying patients into high- and low-risk groups, eventually capturing a stable prediction model, with the C-index reaching 0.704 in the external validation set.[46] Yang et al. tested a different combination of data analysis, looking at TCGA and immune-related genes obtained from the IMMPORT database, to construct an immune gene prediction model. Nineteen immune genes were eventually screened out and used to construct the model, achieving an accuracy of AUC = 0.673 based on patient classification into high- and low-risk groups, with an observed correlation between risk score and degree of neutrophil infiltration in the tumour microenvironment.[47]

Building on these models, Song et al. looked at the development of ML models based on treatment response in patients with unique genetic alterations present. This research validated a DL model for stage IV EGFR-positive LC across 342 select patients, to predict the efficacy of EGFR-tyrosine kinase inhibitor (TKI) chemotherapy in these patients. The median PFS in the external validation set for the low-risk group was 10.9 months, compared with 6.9 months for the high-progression group. The AUC scores for the two external validation sets across high and low risk were 0.73-0.75 and 0.70-0.72, respectively.[48]

One of the more comprehensive studies on AI prognosis, primarily in stage IA-IIIA surgically resected NSCLC, conducted at Kyushu University, combined data from 17 clinicopathological factors and 30 preoperative and 22 post-operative blood tests, to predict DFS, OS, and CSS. The AI model (XGBoost) was then compared to p-Staging; AI analysis of perioperative data rendered better prediction models than p-Staging for all NSCLC stages, and across DFS (71.0%, AUC = 0.890), OS (82.8%, AUC = 0.926), and CSS (88.7%, AUC = 0.960), and with significant improvements on previous models by Hosny et al. and Huang et al. The careful curation and analysis allowed a greatly improved predictive accuracy in this study.

A retrospective analysis of 1387 patients was conducted by Janik et al. using ML to risk stratify early NSCLC patients.[15] The final model which contained heterogeneous data had the highest accuracy of 76% and AUC of 81%. Variables explored by the model were the type of LC, TNM stage, chemotherapy type, tumour stage, presence of a wide range of symptoms, comorbidities, smoking history, family history, age, race, sex, histology grade, biomarkers, and treatment-specific features related to chemotherapy, radiotherapy, and/or surgery. While most cohort studies have looked at OS as the outcome, Janik et al. looked at factors associated with post-operative disease relapse to minimise the impact of non-cancer deaths.[15] The model, however, was unable to predict when the relapse would occur, but rather if it would happen. This does not allow for an understanding of how the disease may progress, which can help guide adjuvant treatment, leading to the model's limited interpretability.

Models using PET, CT scans, and biomolecular data have outperformed traditional methods like TNM staging, demonstrating higher AUC and C-index scores. She et al.'s[36] DL model had a C-index of 0.834 compared with the TNM system's 0.650. AI models incorporating genetic alterations and molecular data, such as EGFR mutations, demonstrated higher predictive accuracy, with models reaching AUC scores up to 0.97 (0/47). Image-based methods, such as radiomic biomarkers for CT analysis, also demonstrated DFS predictions (C-index = 0.72).[43] Comprehensive AI models, for example, XGBoost, achieved predictive accuracies (AUC up to 0.960) through analysis of multifaceted perioperative data. These studies demonstrate the advantage of integrating diverse data types to generate robust and accurate models.

Rakaee et al.[52] created a DL model (Deep-IO) to predict response to ICI, demonstrating superior predictive accuracy

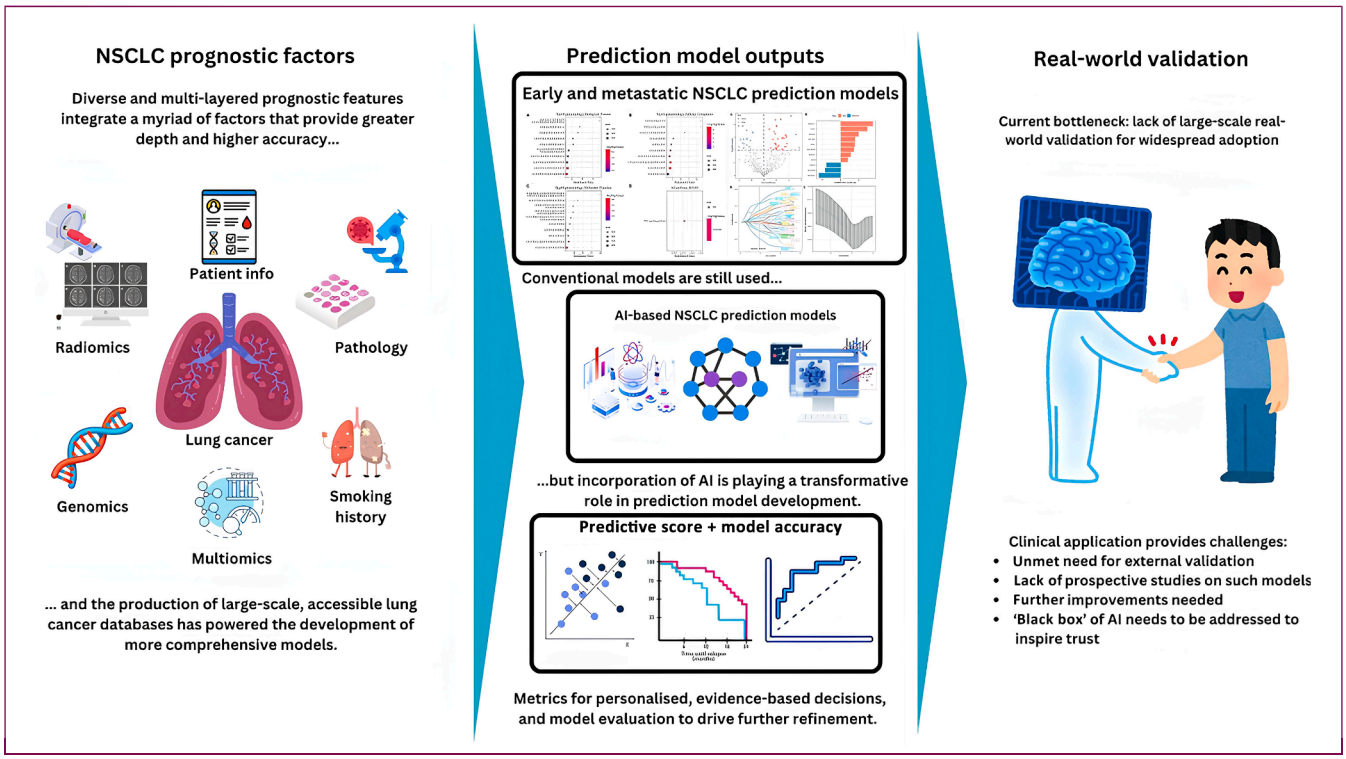

Figure 1. Summary of the development of conventional and AI-based clinical prediction models in the NSCLC landscape and their integration into real-world clinical practice.
AI, artificial intelligence; NSCLC, non-small-cell lung cancer.

compared with traditional biomarkers [PD-L1, tumour mutational burden (TMB), and TILs]. Deep-IO achieved an AUROC of 0.75 in internal testing and 0.66 in external validation. In the validation cohort, combining Deep-IO with PD-L1 improved predictive power and increased response stratification accuracy (AUROC = 0.70). The Deep-IO model was an independent predictor of PFS and OS, showing significant associations with clinical models. Combining Deep-IO and PD-L1 could be useful in stratifying patients more effectively, improving response prediction, and optimising immunotherapy treatment strategies.

Proteomics may play a critical role for enhanced prediction of NSCLC, providing comprehensive insights into the protein expression profiles of tumours. Key research incorporates identification of protein biomarkers; analysis of protein signatures in the blood, plasma, or tissue; identifying patient response to therapies; and stratifying patients into risk groups based on protein expression patterns. These allow distinguishing NSCLC subtypes, predicting disease progression, and indicating response to therapies for a personalised approach with genomic and clinical data.

Christopoulos et al.[53] introduced PROphet, an ML-based plasma proteomics test to improve treatment selection. They conducted a multicentre observational study with 540 NSCLC patients treated with PD-1/PD-L1 inhibitors and 85 chemotherapy patients to profile the plasma proteome and evaluate clinical benefit, survival, and treatment response. PROphet demonstrated superior prediction capacity to immunotherapy than PD-L1 alone. PROphet-negative patients with high PD-L1 expression (>50%) benefitted significantly from combination immunotherapy and chemotherapy (HR = 0.23, $P = 0.0003$). PROphet-positive patients also showed similar survival whether treated with immunotherapy alone or in combination with chemotherapy. Low PD-L1 expression (1%-49%) or PD-L1-negative tumours (<1%) showed differentiated survival outcomes based on PROphet results. PROphet allows for refined treatment decisions for mNSCLC beyond PD-L1 status alone, suggesting integration of plasma proteomics with standard PD-L1 testing improves patient outcomes and helps avoid unnecessary chemotherapy-related toxicity.

Rakaee et al.[54] investigated the association between ML-based TILs on histological images and outcomes of immunotherapy in patients with advanced NSCLC using digital scoring. The study discovered that high TIL levels (>250 cells/mm^2) were independently associated with improved PFS and OS in both cohorts. High TILs also predicted ICI response better than TMB in PD-L1-negative patients (TIL AUC = 0.77, TMB AUC = 0.65). Combining TILs with PD-L1 expression or TMB improved both specificity and positive predicitve value for ICI response compared with PD-L1 alone. TIL assessment offers a cost-effective opportunity, easily integrated into pathology workflows, to enhance precision therapy for NSCLC patients.

While AI is a pioneering technology within the field of NSCLC prediction, there are barriers which may hinder the widespread adoption of such models. These include a lack of large-scale LC pathology databases, a need for both external validation and prospective, as opposed to

https://doi.org/10.1016/j.esmoop.2025.105557

Table 2. Landscape of clinical prediction models in metastatic non-small-cell lung cancer

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination	Guideline approval
Banna et al. (2021)[12]	Retrospective	784	mNSCLC with PD-L1 of at least 50% treated with first-line immunotherapy	Multivariate Cox r egression analysis	Overall survival	Clinicopathological data	Performance status, use of steroids, NLR	Performance status, use of steroids, NLR	1-year overall survival Favourable risk = 78.2% Intermediate risk = 53.8% Poor risk = 10.7%	No—needs further validation with RCTs
Banna et al. (2022)[13]	Retrospective	308	mNSCLC	Multivariate stepwise Cox regression analysis	Overall survival, p rogression-free survival, NLR, and SII	Clinicopathological data	Number of metastatic sites, squamous histology, SII	Number of metastatic sites, squamous histology, SII	Median overall s urvival Favourable risk = 20.3 months Intermediate risk = 12.4 months Poor risk = 8.4 months C-index = 0.623 Progression-free survival: Favourable risk = 11.3 months Intermediate risk = 7.9 months Poor risk = 5.7 months C-index = 0.613	No—needs further validation with RCTs
Chen et al. (2019)[25]	Retrospective	N/A	NSCLC	N/A	Hypoxic effect on cells and their transcription profile, epigenetic effects	Genetic data	17 prognostic genes	17 prognostic genes	Three datasets showed significance in survival time between high- and low-risk groups: 0.0049, 0.04, and 0.025, respectively	N/A—pre-clinical
Ding et al. (2022)[26]	Retrospective	764 samples (31 non-metastatic, 733 metastatic)	NSCLC	Multivariate Cox regression and LASSO	Association between prognostic risk model and TNM stage, immune microenvironment, mechanisms	Genetic data	6 mRNAs	6 mRNAs	N/A	N/A—pre-clinical
Wang et al. (2023)[27]	Retrospective	599 samples	mNSCLC with brain metastases	Multivariate Cox regression analysis and LASSO analysis	Tumour mutational signature, immune signature, sensitivity to treatment	Genetic data	11 prognostic genes	11 prognostic genes	N/A	N/A—pre-clinical
Liu et al. (2021)[28]	Retrospective	243	m NSCLC treated with chemotherapy and radiation	Multivariate Cox regression analysis	Overall survival	Clinicopathological features	Sex, KPS score, number of chemotherapy cycles, Hb level, neutrophil count, platelet count	Sex, KPS score, number of chemotherapy cycles, Hb level, neutrophil count, platelet count	Median overall survival (months) Low risk = 17.0 Moderate risk = 14.0 High risk = 8.0 1-year OS rate: Low risk = 67.7% Moderate risk = 59.6% High risk = 26.2% 2-year OS rate: Low risk = 32.1% Moderate risk = 18.0% High risk = 7.9% 3-year OS rate: Low risk = 19.3% Moderate risk = 7.9% High risk = 0%	No—does not account for molecular therapy

Continued

Table 2. Continued

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination	Guideline approval
Ganti et al. (2019)[29]	Retrospective	1467	mNSCLC	Multivariable Cox proportional hazards model	Overall survival/ progression-free survival	Clinical data	Sex, performance status, tumour stage, weight loss	Sex, performance status, tumour stage, weight loss	Training cohort Good prognosis = 11.77 months Poor prognosis = 6.21 months Area under 1-year ROC = 0.61 Area under 2-year ROC = 0.62 C-index = 0.57 Testing cohort: Good prognosis = 13.15 months Poor prognosis = 8.52 months Area under 1-year ROC = 0.60 Area under 2-year ROC = 0.65 C-index = 0.57	No—needs further validation with RCTs
Navani et al. (2023)[30]	Observational cohort	495	mNSCLC treated with at least 1 dose of ICI monotherapy	LASSO Cox regression	Overall survival	Clinicopathological data	ECOG, LDH level, NLR	ECOG, LDH level, NLR	Median overall survival Favourable risk = 28.3 months Intermediate risk = 9.1 months Poor risk = 2.1 months Intermediate versus favourable HR 2.08 Poor versus favourable HR 5.21	No—needs further validation with RCTs
Mezquita et al. (2018)[33]	Retrospective	466	mNSCLC	Multivariate Cox proportional hazards regression model	Overall survival, progression-free survival, disease control rate	Pathological data	dNLR, LDH	dNLR, LDH	Median overall survival: Good risk = 34 months Intermediate risk = 10 months Poor risk = 3 months Progression-free survival Good risk = 6.3 months Intermediate risk = 3.7 months Poor risk = 2.0 months	No—needs further validation with RCTs
Chen et al (2023)[34]	Retrospective	194	mNSCLC	Linear regression	Overall survival	Pathological data	15 radiomic feature, fractal dimensions, texture features	PD-L1 status, treatment response	Moderate discrimination PD-L1 expression AUC = 0.66-0.72 Treatment response AUC = 0.68	No—needs further validation with RCTs

dNLR, derived neutrophil-to-lymphocyte ratio; ECOG, Eastern Cooperative Oncology Group; Hb, haemoglobin; ICI, immune checkpoint inhibitor; KPS, Karnofsky Performance Status; LASSO, least absolute shrinkage and selection operator; LDH, lactate dehydrogenase; mNSCLC, metastatic non-small-cell lung cancer; N/A, not applicable; NLR, neutrophil-to-lymphocyte ratio; PD-L1, programmed death-ligand 1; RCTs, randomised control trials; ROC, receiver operating characteristic; SII, systemic immune-inflammatory index; TNM, tumour—node—metastasis.

https://doi.org/10.1016/j.esmoop.2025.105557

Table 3. Landscape of artificial intelligence-based clinical prediction models in early and metastatic non-small-cell lung cancer

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination	Guideline approval
Liu et al. 2023[21]	Retrospective	198	Stage IA or IB invasive lung adenocarcinoma	Cox regression model	5-year overall survival (OS), disease-free survival (DFS)	Patient demographics, clinical and pathological parameters	Demographics: age and sex Pathological parameters: smoking history, type of surgery, adjuvant chemotherapy, T stage, lymphovascular invasion, necrosis, neuron invasion, architectural score	T stage, necrosis, and architectural score were key prognostic predictors	AUC: Prediction model = 0.717 T stage = 0.624 Architectural score = 0.611 Necrosis = 0.617	N/A—further studies and external validation are required
She et al. 2020[36]	Retrospective	17 322	Stages I-IV NSCLC	DeepSurv—a deep learning (DL) survival neural network Cox proportional hazards regression model	Prognostic model: Survival predictions and treatment recommendations Treatment recommendation: SEER database—population that followed treatment recommendations; CHINA database = population that did not follow recommendations	Patient characteristics, tumour stage, treatment strategies	127 features, including Cases (sex, age, and marriage status), tumour characteristics (location, size, histological grade, histological type, TNM stage), SEER code (CS extension, CS mets at dx, regional nodes examined, regional nodes positive, lung—pleural/elastic layer invasion by H&E or elastic stain, lung—separate tumour nodules—ipsilateral lung, lung—surgery to primary site), and treatment details (surgical type)	N/A	Prediction model C-index: DeepSurv model = 0.739 Cox regression model = 0.716 Treatment recommendation hazard ratio: SEER = 2.99 CHINA = 2.14	N/A—feasibility trial status; further studies and external validation needed
Jin et al. 2023[37]	Retrospective	4617	Surgically resected stage aI3 NSCLC	Model 1: DL neural network Model 2: Random forest Model 3: Cox proportional hazards model	5-year survival	Patient demographics, NSCLC-related characteristics, and treatment details	Demographic information (age and sex) and NSCLC-cancer-related characteristics (TNM stage, histology type, primary site, tumour size, regional node number examined, regional node positive number, and laterality of the tumour), and treatment details (surgery of primary site, radiation, and chemotherapy)	Regional positive nodes (−0.6634); regional examined nodes (−0.7648); tumour size (−0.5633); age (−0.4633)	C-index: DL neural network = 0.843 Random forest = 0.678 Cox proportional hazards = 0.678	N/A—improvements to model explainability needed
Palmer et al. 2023[38]	Retrospective	392	Patients undergoing radical surgery for early-stage NSCLC	Three regularised Cox models: Ridge, LASSO, Elastic Net	Recurrence-free survival (RFS)	66 clinicopathological features across 3 models	Elastic Net model: Systemic inflammatory response index (SIRI), eosinophil count, preoperative nodal stage, weight loss, performance status, and maximum standardised uptake value (SUV-max) Ridge model: Performance status, weight loss, SIRI, eosinophil count, lymphovascular invasion, visceral pleural invasion, and pathological stage	N/A	C-index: Elastic Net = 0.70 Ridge = 0.75 AUC ROC: Elastic Net = 0.72 Ridge = 0.79	N/A—prospective and external validation required

Continued

Table 3. Continued

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination	Guideline approval
Hindocha et al. 2022[39]	Retrospective, multicentre	657 patients, across 5 hospitals	Patients receiving curative-intent radiotherapy for NSCLC	Logistic regression models based on TNM staging (for benchmarking) Machine learning (ML) models	Recurrence, RFS, OS 2 years after treatment	Patient demographics and clinical parameters	Age, PTV, primary size, SUV primary, SUV nodal, BMI, FEV1, TLCO, total dose, BED, sex, smoking status, T staging, N staging, nodal avidity, morphology (adenocarcinoma/no pathological diagnosis), morphology, treatment (chemo, conventional, SBRT)	RFS: PTV, size of the primary tumour, total number of fractions, whether treatment was with SBRT or BED Recurrence: PTV, size of primary tumour, number of fractions, SBRT treatment, BED OS: PTV, size of primary tumour, BED, T1 stage, treatment with SBRT, treatment with chemoradiotherapy, whether nodal avidity was present, smoking status	AUC ROC: RFS: ML model = 0.682 TNM model = 0.650 Recurrence: ML model = 0.687 TNM model = 0.670 OS: ML model = 0.659 TNM model = 0.649	N/A—more advanced, radiotherapy-specific, and image-based models are required
Huang et al. 2016[40]	Retrospective	282	Early-stage (IA-IIB) NSCLC	Cox regression models: radiomics signature versus clinicopathological risk factors	DFS	Radiomics, TNM staging, clinicopathological factors	Radiomics signature as an independent biomarker; TNM staging, clinicopathological factors	N/A	C-index: Radiomics signature = 0.72 Clinical-pathological nomogram = 0.691	N/A—building on previous model, but further validation required
Hosny et al. 2018[41]	Retrospective, multi-cohort	1194	Stage I-IIIb NSCLC patients treated with radiotherapy versus patients treated with surgery	3D convolutional neural network (CNN) Random forest (for benchmarking)	2-year OS	CT data, clinical parameters, preliminary genomic associations	CNN models: labelled CT scans, contrast-enhanced CT, MRI of the brain, bronchoscopy with TBNA, mediastinoscopy, radiotherapy treatment, chemotherapy treatment, PET—CT scans Random forest: clinical information (age, sex, staging, clinicopathological factors)	Large, uninterrupted areas of relatively higher CT density	AUC ROC: DL CNN: Radiotherapy = 0.70 Surgery = 0.71 Random forest: Radiotherapy = 0.55 Surgery = 0.58	N/A—rigorous evaluation in future prognostic studies necessary; improvements required include interoperability, model blind spots, and sensitivity to medical images
Zhang et al. 2023[42]	Retrospective	6371	Surgically resected, stage I-III NSCLC	Intelligent prognosis evaluation system (IPES) Cox regression (for benchmarking)	OS	Pre-therapy CT images, EGFR genotype	IPES model: Cisplatin-based chemotherapy, EGFR-TKI-targeted adjuvant therapy, EGFR mutation, sex, age, smoking status, complete resection, cancer/tumour family history, histology (adenocarcinoma/other), staging, death status Cox regression: Clinical metadata, TNM staging	N/A	C-index: IPES = 0.817; Cox models = 0.733	N/A—a more comprehensive integration of the clinicopathological factors is required
Martell et al. 2024[43]	Prospective	48	Surgically resected NSCLC patients who have not received chemo before surgery	Deep representation learning Pearson rank order correlation (for benchmarking)	Histology subtype, OS	CT scans, clinicopathological data, tissue metabolomics	Sex, TNM staging, histology, CT scan (contrast/non-contrast), treatment received, dosage, end result, tissue metabolome (amino acid, lipid, xenobiotics, nucleotide, peptide, energy, etc.)	Sedoheptulose-7-phosphate and uridine—metabolites with the highest correlation (0.72 and 0.64, respectively)	C-index: Radiomics = 0.62 CT = 0.57 TMR-CT = 0.74 Radiomics + clinical = 0.64 CT + clinical = 0.59 TMR-CT + clinical = 0.78	N/A—retrospective study

Continued

https://doi.org/10.1016/j.esmoop.2025.105557

Table 3. Continued

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination	Guideline approval
Jiang et al. 2019[44]	Retrospective	524	Lung adenocarcinoma (IA-IV)	Artificial neural network (ANN) Cox regression model	Progression-free survival (PFS) OS	Clinical data; genetic mutation data	Smoking history, tumour stage, surgical margin resection status, primary tumour site, TCGA database, pathological grade of the margin, types of gene mutation, patient's vital status	Smoking history, tumour stage, surgical margin level	AUC ROC: ANN model = 0.712 Cox regression model = 0.662	N/A—further improvements and more comprehensive analysis of TCGA needed for clinical use
Coudray et al. 2018[45]	Retrospective	1634	TCGA database	Deep convolutional neural network (CNN) Random sampling	Cancer prognostication based on genetic mutations	Whole slide H&E images, genetics/genomics	Morphology, immunohistochemical staining, necrotic regions, inflammation, blood clotting, fibrotic scars, genetic mutations	Genetic mutations most commonly associated with adenocarcinoma—STK11, EGFR, FAT1, SETBP1, KRAS, and TP53	AUC: CNN = 0.750 Random sampling = 0.687	N/A—classification needs to be extended to other subtypes and histological features
Li et al. 2019[46]	Retrospective	492 579	TCGA adenocarcinoma database GEO adenocarcinoma database	SigFeature (an ML algorithm)	1-year, 3-year, and 5-year OS	Genetic mutations	16-gene-based prognostic prediction model: C1QTNF6, ENPP5, PLEK2, RGS20, RHOV, SFTA3, GNG7, HMMR, IGF2BP1, PEBP1, UPK1B, SRGAP1, ZNF14, IER5L, SATB2, STYX	Over-expression of IGF2BP1, UPK1B, SRGAP1, SATB2, C1QTNF6, RHOV, IER5L, STYX, HMMR, PLEK2, and RGS20; and under expression of PEBP1, SFTA3, GNG7, ENPP5, and ZNF14	C-index: SigFeature: 1-year = 0.753 3-year = 0.726 5-year = 0.656;	N/A—retrospective
Yang et al. 2022[47]	Retrospective	1128 2498 318	TCGA database IMMPORT database Cistrome Cancer database	Cox proportional hazards regression model Decision tree model	5-year overall survival (OS)	Transcriptomic, immune, and clinicopathological data	193 differential immune genes, clinical parameters (sex, age, stage, tumour histology, survival)	Degree of neutrophil infiltration in the tumour microenvironment; immune genes with significant prognostic benefit: LTB4R, IL-33, and SHC3	AUC: TCGA data: Cox model = 0.673 Decision tree model = 0.601 GEO data: Cox = 0.712	N/A - large-scale, multicentre evidence necessary for further validation and verification
Song et al. 2020[48]	Prospective	342	Stage IV EGFR-positive NSCLC patients, receiving EGFR-TKI therapy	DL semantic signature Radiomics signature model	PFS	Clinical data, CT scans, EGFR subtype	Clinical data (sex, age, smoking status, performance status score, histopathological subtype), administered therapeutic regimen, EGFR variant subtype, enhanced/non-contrast CT	N/A	AUC: DL model = 0.73	N/A—screening of CT sections and an image-encoding process corresponding to DL semantic features is needed to better explain the model
Deng et al. 2022[49]	Retrospective	570 129	Stage IV NSCLC patients treated with EGFR-TKIs Stage IV patients treated with ICIs	EfficientNetV2-based survival benefit prognosis (ESBP) model Experienced oncologist	PFS	Preoperative CT scans, mutation presence, demographic data	Demographic information, including sex, age, smoking status, performance status score, histopathological subtype, EGFR mutation subtype (EGFR-TKI patients), tumour proportion score (ICI-treated patients), and the administered therapeutic regimen	N/A	C-index: ESBP = 0.690 10-year experienced oncologists = 0.651	N/A—although ESBP has been tested in the clinical setting based on how well it can assist clinical staff, further validation and optimisation on larger and more diverse datasets are needed

Continued

Table 3. Continued

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination	Guideline approval
Kinoshita et al. 2023[50]	Prospective	1049	Surgically resected stage I-IIIA NSCLC	eXtreme Gradient Boosting (XGBoost) ML algorithm p-Stage scores	DFS, OS, cancer-specific survival (CSS) at 5 years	17 clinicopathological factors, 30 preoperative, and 22 post-operative blood tests	Tumour malignancy (p-Stage, pathological N status, histological type, pleural invasion, lymphatic invasion, SUV-max, CEA, and CYFRA), pulmonary function, and smoking history. The preoperative blood test results, including total protein, albumin, total bilirubin, direct bilirubin, aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, γ-glutamyl transpeptidase, lactate dehydrogenase, urea nitrogen, creatinine, urine acid, sodium, potassium, chlorine, calcium, total cholesterol, triglyceride, glucose, CRP, white blood cell, neutrophil, lymphocyte, monocyte, haemoglobin, platelet, prothrombin time-international normalised ratio, and activated partial thromboplastin time, were averaged over the 1-month period before surgery	Preoperative prothrombin time-international normalised ratio and post-operative CRP were identified as variables of high importance for DFS, OS, and CSS; post-operative total protein, post-operative aspartate aminotransferase, post-operative creatinine, post-operative monocyte, preoperative lymphocyte, post-operative lymphocyte, and post-operative lactate dehydrogenase were identified as important variables in at least two among DFS, OS, and CSS	AUC: XGBoost: DFS = 0.890 OS = 0.926 CSS = 0.960 p-Stage: DFS = 0.756 OS = 0.758 CSS = 0.786	N/A—retrospective study on a single cohort; external validation is necessary
Torrente et al. 2022[51]	Retrospective	5275	NSCLC patients	AI-powered CLARIFY DSP		Clinical data, questionnaires data, data collected from wearable devices	N/A	N/A	N/A	N/A—however, CLARIFY project has been extensively validated and might soon be used as an independent decision-support tool
Janik et al. 2023[15]	Prospective, multi-site	1387	Early stage (I-II) NSCLC	ML model Previous work (Lee et al., 2020)	5-year OS	Clinical data, molecular data, radiological data	Tumour characteristics (stage, TNM, histology, grade, subtype); general (age, race, sex, previous cancer, family cancer history, Eastern Cooperative Oncology Group, synchronous tumours, and biomarkers: anaplastic lymphoma kinase immunohistochemistry, PD-L1, and epidermal growth factor receptor-negative); comorbidities; smoking; symptoms; radiotherapy (type, area, dose, fractioning, duration); chemotherapy (type, start time, and regimen); surgery (procedure, time, type, resection grade, and TN stages)	Having four family members with other cancers is the highest contributing factor, which increases the prediction by 0.19; missing information on the regimen of the chemotherapy or not having it done at all decreases the prediction by 0.07	AUC: ML model = 0.81 Previous work = 0.7677	N/A—further prospective and multi-site studies necessary, which would also incorporate radiomics and genomics

Continued

https://doi.org/10.1016/j.esmoop.2025.105557

Table 3. Continued

Author (year)	Study design	Sample size	Stage and type of lung cancer	Type of model	Outcome(s) investigated	Type of prognostic features included in the model(s)	Predictive features included in the final model(s)	Variables significantly associated with outcome	Model discrimination	Guideline approval
Rakaee et al. 2025[32]	Multicentre cohort study	Retrospective	958 patients (development cohort—614, validation cohort—344)	Supervised DL-based response stratification H&E with binary classification (responder versus non-responder)	Objective response rate PFS OS	PD-L1 expression TMB TILs Age Sex ECOG performance Histology Line of therapy	Deep-IO model probability score	Deep-IO score PD-L1 expression ECOG performance Status Histological subtype Sex	AUC: Internal set: 0.75 (95% CI 0.64-0.85) External validation: 0.66 (95% CI 0.60-0.72) Deep-IO + PD-LI 0.70 (95%, CI, 0.63-0.76)	Further validation needed

AUC, area under the curve; BED, biologically effective dose; BMI, body mass index; CI, confidence interval; CRP, C-reactive protein; CS, collaborative staging; CT, computed tomography; ECOG, Eastern Cooperative Oncology Group; EGFR, epidermal growth factor receptor; FEV1, forced expiratory volume in 1 second; GEO, Gene Expression Omnibus; H&E, haematoxylin—eosin; ICIs, immune checkpoint inhibitors; IL, interleukin; IO, immunotherapy; LASSO, least absolute shrinkage and selection operator; MRI, magnetic resonance imaging; N/A, not applicable; NSCLC, non-small-cell lung cancer; PD-L1, programmed death-ligand 1; PET, positron emission tomography; PTV, planning target volume; ROC, receiver operating characteristic; SBRT, stereotactic body radiation therapy; TBNA, transbronchial needle aspiration; TCGA, The Cancer Genome Atlas; TILs, tumour-infiltrating lymphocytes; TKI, tyrosine kinase inhibitor; TLCO, transfer factor for carbon monoxide; TMR, tissue-metabolic-radiomic; TNM, tumour—node—metastasis.

retrospective studies (which face selection bias and lack real-time data affecting accuracy), and the compatibility of health and biomedical data. Prospective validation would be essential to ensure safe and effective integration of AI prediction models in clinical workflows.

AI-based models have great scope for personalised, evidence-based treatment, enhancing the decision-making process, and helping oncologists anticipate problems or identify risk factors and eligibility for targeted therapy. A multidimensional approach and a multivariate analysis, incorporating whole CT scans and histopathology findings with the electronic health record, clinical reports, and underlying molecular drivers and mutations, will be key to creating mixed predictive analysis models with high degrees of accuracy to be reliably used in the clinical setting.

CONCLUSIONS

CPMs enable personalised risk assessment, prediction of OS, and enhanced treatment decision making. The advent of contemporary approaches including immunotherapies has sparked prediction model development, tailored to the unique challenges and opportunities of such treatments.

Numerous prediction models for early-stage NSCLC show a stride towards personalised medicine, offering the potential to stratify patients based on individualised risk profiles. While current guidelines predominantly advocate for surgical resection as the cornerstone of curative treatment, the landscape is evolving with the emergence of adjuvant therapies and targeted interventions. For mNSCLC, significant progress has been made in developing detailed CPMs and exploring metastasis-associated gene expression profiles. These advancements hold promise for more precise patient risk stratification and treatment selection. Gene expression models offer insights into the molecular underpinnings of metastasis and its impact on patient outcomes, giving rise to potential targets for therapeutic intervention, particularly in immunotherapy. Besides gene expression profiles, studies have shown the value of various histological and molecular parameters in prognostic modelling, as well as other key prognostic determinants, advocating for a holistic approach to prognostication.

ML techniques have played a pivotal role in identifying multifaceted prognostic markers, encompassing clinical, demographic, and histological parameters. Radiomics-based tools leveraging AI have shown promise in analysing imaging data, while molecular-level analysis incorporating genomic alterations has facilitated the development of personalised prediction models. AI-driven prediction models have the potential to revolutionise treatment selection and optimisation by identifying patients likely to benefit from specific therapies. Given the multifactorial nature of NSCLC prognosis, the development of multimodal AI prediction models, capable of integrating data from various sources, may improve the accuracy of predictions.

Although prediction models are being increasingly developed, their integration into real-world clinical practice has been limited and several challenges persist (Figure 1). Firstly, more prospective validation using large, representative datasets is required to ensure scientific robustness of models. Secondly, models for patients receiving combination therapies such as combined chemotherapy require continued research and validation efforts to translate them into clinical practice. Finally, for real-world use, it is crucial that prediction models yield clinically relevant outcomes and are aligned with current guidelines. Consideration should be given to the routine availability of variables. If variables are not routinely available, models must demonstrate significant survival stratification to justify the use of expensive or inaccessible testing for prognostication. This will also expand opportunities for external validation across centres in low- and middle-income countries. In summary, while several models show strong predictive potential, such as LIPS-3, LITE, and XGBoost-based models, we believe that none are currently ready for routine clinical adoption without further prospective validation and external real-world testing. Many models demonstrate promising discrimination metrics, but limited generalisability and the lack of integration with current clinical workflows remain key barriers.

Despite these challenges, AI holds immense promise as a tool for advancing personalised medicine in NSCLC. None the less, the role of the clinician remains invaluable in selecting relevant patient characteristics, as well as in preventing collinearity through an understanding of each variable's significance. By leveraging AI technologies, clinicians can navigate the complexities of NSCLC more effectively. We anticipate a rise in the development of AI-based prediction models in the years to come.

FUNDING

None declared.

DISCLOSURE

AC received grants for consultancies/advisory boards from MSD, BMS, IQVIA, AstraZeneca, REGENERON, Amgen, Daiichi Sankyo, Access Infinity, Ardelis Health, Alpha Sight, Guidepoint, Roche, and Alira Health; speaker fees from AstraZeneca, Pierre-Fabre, MSD, and Sanofi/REGENERON; payment for writing/editorial activity from BMS, MSD, and Roche; travel support from Sanofi/REGENERON, and MSD. All other authors have declared no conflicts of interest.

REFERENCES

1. Huang J, Deng Y, Tin MS, et al. Distribution, risk factors, and temporal trends for lung cancer incidence and mortality: a global analysis. *Chest*. 2022;161(4):1101-1111.
2. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol*. 2015;10(9):1243-1260.
3. Carter L, Apte V, Shukla A, et al. Stage 3 N2 lung cancer: a multidisciplinary therapeutic conundrum. *Curr Oncol Rep*. 2024;26(1):65-79.
4. Banna GL, Passiglia F, Colonese F, et al. Immune-checkpoint inhibitors in non-small cell lung cancer: a tool to improve patients' selection. *Crit Rev Oncol Hematol*. 2018;129:27-39.
5. Addeo A, Banna GL, Metro G, DI Maio M. Chemotherapy in combination with immune checkpoint inhibitors for the first-line treatment of patients with advanced non-small cell lung cancer: a systematic review and literature-based meta-analysis. *Front Oncol*. 2019;9:264.
6. Craddock M, Crockett C, McWilliam A, et al. Evaluation of prognostic and predictive models in the oncology clinic. *Clin Oncol (R Coll Radiol)*. 2022;34(2):102-113.
7. Goldstraw P, Chansky K, Crowley J, et al. The IASLC Lung Cancer Staging Project: proposals for revision of the TNM stage groupings in the forthcoming (Eighth) edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2016;11(1):39-51.
8. Shoji F, Kozuma Y, Toyokawa G, Yamazaki K, Takeo S. Complete blood cell count-derived inflammatory biomarkers in early-stage non-small-cell lung cancer. *Ann Thorac Cardiovasc Surg*. 2020;26(5):248-255.
9. Friedlaender A, Metro G, Signorelli D, et al. Impact of performance status on non-small-cell lung cancer patients with a PD-L1 tumour proportion score $\geq$50% treated with front-line pembrolizumab. *Acta Oncol*. 2020;59(9):1058-1063.
10. Banna GL, Signorelli D, Metro G, et al. Neutrophil-to-lymphocyte ratio in combination with PD-L1 or lactate dehydrogenase as biomarkers for high PD-L1 non-small cell lung cancer treated with first-line pembrolizumab. *Transl Lung Cancer Res*. 2020;9(4):1533-1542.
11. Addeo A, Friedlaender A, Banna GL, Weiss GJ. TMB or not TMB as a biomarker: that is the question. *Crit Rev Oncol Hematol*. 2021;163:103374.
12. Banna GL, Cortellini A, Cortinovis DL, et al. The lung immuno-oncology prognostic score (LIPS-3): a prognostic classification of patients receiving first-line pembrolizumab for PD-L1 $\geq$ 50% advanced non-small-cell lung cancer. *ESMO Open*. 2021;6(2):100078.
13. Banna GL, Cantale O, Muthuramalingam S, et al. Efficacy outcomes and prognostic factors from real-world patients with advanced non-small-cell lung cancer treated with first-line chemoimmunotherapy: the Spinnaker retrospective study. *Int Immunopharmacol*. 2022;110:108985.
14. eUpdate — Early and locally advanced non-small-cell lung cancer. Available at https://www.esmo.org/guidelines/guidelines-by-topic/esmo-clinical-practice-guidelines-lung-and-chest-tumours/early-stage-and-locally-advanced-non-metastatic-non-small-cell-lung-cancer-esmo-clinical-practice-guidelines/eupdate-early-and-locally-advanced-non-small-cell-lung-cancer-nsclc-treatment-recommendations2. Accessed April 29, 2024.
15. Janik A, Torrente M, Costabello L, et al. Machine learning-assisted recurrence prediction for patients with early-stage non-small-cell lung cancer. *JCO Clin Cancer Inform*. 2023;7:e2200062.
16. Sheikh M, Virani S, Robbins HA, et al. Survival and prognostic factors of early-stage non-small cell lung cancer in Central and Eastern Europe: a prospective cohort study. *Cancer Med*. 2023;12(9):10563-10574.
17. Kawaguchi T, Takada M, Kubo A, et al. Performance status and smoking status are independent favorable prognostic factors for survival in non-small cell lung cancer: a comprehensive analysis of 26,957 patients with NSCLC. *J Thorac Oncol*. 2010;5(5):620-630.
18. Sculier JP, Chansky K, Crowley JJ, Van Meerbeeck J, Goldstraw P. International Staging Committee and Participating Institutions. The impact of additional prognostic factors on survival and their relationship with the anatomical extent of disease expressed by the 6th Edition of the TNM Classification of Malignant Tumors and the proposals for the 7th Edition. *J Thorac Oncol*. 2008;3(5):457-466.
19. Pilotto S, Sperduti I, Novello S, et al. Risk stratification model for resected squamous-cell lung cancer patients according to clinical and pathological factors. *J Thorac Oncol*. 2015;10(9):1341-1348.
20. Lu C, Bera K, Wang X, 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*. 2020;2(11):e594-e606.

21. Liu DH, Ye ZH, Chen S, et al. Novel prognostic model for stratifying survival in stage I lung adenocarcinoma patients. *J Cancer Res Clin Oncol*. 2020;146(3):801-807.
22. Pilotto S, Sperduti I, Leuzzi G, et al. Prognostic model for resected squamous cell lung cancer: external multicenter validation and propensity score analysis exploring the impact of adjuvant and neoadjuvant treatment. *J Thorac Oncol*. 2018;13(4):568-575.
23. Li B, Cui Y, Diehn M, Li R. Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer. *JAMA Oncol*. 2017;3(11):1529-1537.
24. Wu P, Zheng Y, Wang Y, Wang Y, Liang N. Development and validation of a robust immune-related prognostic signature in early-stage lung adenocarcinoma. *J Transl Med*. 2020;18(1):380.
25. Chen YL, Zhang Y, Wang J, et al. A 17 gene panel for non-small-cell lung cancer prognosis identified through integrative epigenomic-transcriptomic analyses of hypoxia-induced epithelial—mesenchymal transition. *Mol Oncol*. 2019;13(7):1490-1502.
26. Ding H, Shi L, Chen Z, et al. Construction and evaluation of a prognostic risk model of tumor metastasis-related genes in patients with non-small cell lung cancer. *BMC Med Genomics*. 2022;15(1):187.
27. Wang R, Zhang X, He C, Guo W. An effective prognostic model for assessing prognosis of non-small cell lung cancer with brain metastases. *Front Genet*. 2023;14:1156322.
28. Ganti AK, Wang X, Stinchcombe TE, et al. Clinical prognostic model for older patients with advanced non-small cell lung cancer. *J Geriatr Oncol*. 2019;10(4):555-559.
29. Liu LF, Li QS, Hu YX, et al. Prognostic model to predict overall survival for metastatic non-small cell lung cancer patients treated with chemotherapy combined with concurrent radiation therapy to the primary tumor: analysis from two prospective studies. *Front Oncol*. 2021;11:625688.
30. Navani V, Meyers DE, Ruan Y, et al. Lung immune therapy evaluation (LITE) risk, a novel prognostic model for patients with advanced non-small cell lung cancer treated with immune checkpoint blockade. *Clin Lung Cancer*. 2023;24(3):e152-e159.
31. Ferrucci PF, Ascierto PA, Pigozzo J, et al. Baseline neutrophils and derived neutrophil-to-lymphocyte ratio: prognostic relevance in metastatic melanoma patients receiving ipilimumab. *Ann Oncol*. 2016;27(4):732-738.
32. Weide B, Martens A, Hassel JC, et al. Baseline biomarkers for outcome of melanoma patients treated with pembrolizumab. *Clin Cancer Res*. 2016;22(22):5487-5496.
33. Mezquita L, Auclin E, Ferrara R, et al. Association of the lung immune prognostic index with immune checkpoint inhibitor outcomes in patients with advanced non-small cell lung cancer. *JAMA Oncol*. 2018;4(3):351-357.
34. Chen M, Lu H, Copley SJ, 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*. 2023;18(6):718-730.
35. Paz-Ares L, Vicente D, Tafreshi A, et al. A randomized, placebo-controlled trial of pembrolizumab plus chemotherapy in patients with metastatic squamous NSCLC: protocol-specified final analysis of KEYNOTE-407. *J Thorac Oncol*. 2020;15(10):1657-1669.
36. She Y, Jin Z, Wu J, et al. Development and validation of a deep learning model for non-small cell lung cancer survival. *JAMA Netw Open*. 2020;3(6):e205842.
37. Jin L, Zhao Q, Fu S, Cao F, Hou B, Ma J. Development and validation of machine learning models to predict survival of patients with resected stage-III NSCLC. *Front Oncol*. 2023;13:1092478.
38. Palmer KR, Houston A, Macpherson H, et al. 87P Prognostic models of recurrence-free survival in non-small cell lung cancer. *J Thorac Oncol*. ;18(4):S92.
39. Hindocha S, Charlton TG, Linton-Reid K, et al. A comparison of machine learning methods for predicting recurrence and death after curative-intent radiotherapy for non-small cell lung cancer: development and validation of multivariable clinical prediction models. *EBioMedicine*. 2022;77:103911.
40. Huang Y, Liu Z, He L, et al. Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology*. 2016;281(3):947-957.
41. Hosny A, Parmar C, Coroller TP, et al. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med*. 2018;15(11):e1002711.
42. Zhang S, Liu X, Zhou L, et al. Intelligent prognosis evaluation system for stage I-III resected non-small-cell lung cancer patients on CT images: a multi-center study. *EClinicalMedicine*. 2023;65:102270.
43. Martell MB, Linton-Reid K, Hindocha S, et al. *Deep representation learning of tissue metabolome and computed tomography images annotates non-invasive classification and prognosis prediction of NSCLC. arXiv:2305.15590*. 2023.
44. Jiang N, Xu X. Exploring the survival prognosis of lung adenocarcinoma based on the cancer genome atlas database using artificial neural network. *Medicine*. 2019;98(20):e15642.
45. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559-1567.
46. Li Y, Ge D, Gu J, Xu F, Zhu Q, Lu C. A large cohort study identifying a novel prognosis prediction model for lung adenocarcinoma through machine learning strategies. *BMC Cancer*. 2019;19(1):886.
47. Yang D, Ma X, Song P. A prognostic model of non small cell lung cancer based on TCGA and ImmPort databases. *Sci Rep*. 2022;12(1):437.
48. Song J, Wang L, Ng NN, 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*. 2020;3(12):e2030442.
49. Deng K, Wang L, Liu Y, et al. A deep learning-based system for survival benefit prediction of tyrosine kinase inhibitors and immune checkpoint inhibitors in stage IV non-small cell lung cancer patients: a multicenter, prognostic study. *EClinicalMedicine*. 2022;51:101541.
50. Kinoshita F, Takenaka T, Yamashita T, et al. Development of artificial intelligence prognostic model for surgically resected non-small cell lung cancer. *Sci Rep*. 2023;13(1):15683.
51. Torrente M, Sousa PA, Hernández R, et al. An artificial intelligence-based tool for data analysis and prognosis in cancer patients: results from the clarify study. *Cancers (Basel)*. 2022;14(16):4041.
52. Rakaee M, Tafavvoghi M, Ricciuti B, et al. Deep learning model for predicting immunotherapy response in advanced non—small cell lung cancer. *JAMA Oncol*. 2025;11(2):109-118.
53. Christopoulos P, Harel M, McGregor K, et al. Plasma proteome-based test for first-line treatment selection in metastatic non-small cell lung cancer. *JCO Precis Oncol*. 2024;8:e2300555.
54. Rakaee M, Adib E, Ricciuti B, et al. Association of machine learning-based assessment of tumor-infiltrating lymphocytes on standard histologic images with outcomes of immunotherapy in patients with NSCLC. *JAMA Oncol*. 2023;9(1):51-60.