

ORIGINAL ARTICLE

A unified imaging-to-pharmacology AI framework for pulmonary nodule detection, malignancy classification, and mutation-stratified drug validation from low-dose CT scans

Reddy Shiva Shankar^{1*}, Nakka Deshai², Chigurupati Ravi Swaroop¹, and Kankanala Amrutha¹

¹Department of Computer Science and Engineering, Faculty of Engineering, Sagi Rama Krishnam Raju Engineering College (A), Bhimavaram, Andhra Pradesh, India

²Department of Information Technology, Faculty of Engineering, Sagi Rama Krishnam Raju Engineering College (A), Bhimavaram, Andhra Pradesh, India

Abstract

Background: Low-dose computed tomography (CT) screening for lung cancer can impose a substantial and clinically under-recognised psychosocial burden, with published studies reporting clinically significant anxiety among 30%–45% of screen-positive individuals, driven by a 96.4% benign false-positive rate and prolonged diagnostic uncertainty. Chronic distress may activate the hypothalamic–pituitary–adrenal axis and contribute to catecholamine-driven β_2 -adrenergic–epidermal growth factor receptor transactivation. **Methods:** We present PulmoNet-X, a unified deep learning framework that integrates a six-module imaging backbone with three psychosomatic burden-reduction mechanisms: false-positive reduction module (FPRM), uncertainty-calibrated risk stratification (UCRS), and integrated drug validation and therapeutic drug recommendation (IDVTRM). The framework was evaluated on the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) benchmark comprising 888 CT scans and 1,004 annotated nodules using 10-fold cross-validation. **Results:** PulmoNet-X achieved 98.1% nodule-detection sensitivity at one false-positive/scan (competition performance metric = 93.7%), malignancy classification area under the curve (AUC) = 97.8%, segmentation Dice similarity coefficient (DSC) = 84.6%, and Lung Imaging Reporting and Data System agreement κ = 0.87 on LIDC-IDRI. FPRM reduced false-positive detections from 4.8 to 1.0 per scan, corresponding to a modelled estimate of approximately 380 fewer unnecessary recall investigations per 1,000 patients screened. UCRS achieved a calibration error of 0.016 (expected calibration error). IDVTRM attained a mutation-stratified drug-response AUC of 0.918 and lung-tissue penetration prediction (R^2 = 0.986), with 94.2% concordance with independent expert oncologist assessment. All primary pairwise comparisons were statistically significant (two-sided Wilcoxon signed-rank test, p < 0.001, Bonferroni-corrected for six comparisons). **Conclusion:** External validation on 150 independent National Institutes of Health Chest CT scans provided preliminary evidence supporting generalisability beyond LIDC-IDRI (sensitivity = 98.1%, AUC = 99.3%, DSC = 95.8%, κ = 0.94). Reported psychosomatic benefits remain mechanistically grounded hypotheses; validated psychometric

*Corresponding author:

Reddy Shiva Shankar
(shiva@srkrec.ac.in)

Citation: Shankar RS, Deshai N, Swaroop CR, Amrutha K. A unified imaging-to-pharmacology AI framework for pulmonary nodule detection, malignancy classification, and mutation-stratified drug validation from low-dose CT scans. *J Clin Transl Res*. doi: 10.36922/JCTR026260059

Received: June 22, 2026

Revised: July 30, 2026

Accepted: August 4, 2026

Published online: September 11, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution Noncommercial License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

instruments, including the Generalised Anxiety Disorder 7-item scale, Patient Health Questionnaire-9, and Hospital Anxiety and Depression Scale, were not administered. The pharmacological prediction component remains exploratory, and the incremental contribution of imaging features to drug-response prediction has not been established. Prospective, multi-centre validation is required before clinical translation. PulmoNet-X demonstrates the feasibility of incorporating psychosomatic burden reduction as an explicit architectural objective alongside imaging and pharmacological modelling. **Relevance for patients:** PulmoNet-X is designed to reduce unnecessary recalls, screening anxiety, and diagnostic uncertainty by providing calibrated cancer risk estimates and personalised treatment guidance.

Keywords: Lung cancer; Psychosomatic burden; Low-dose computed tomography screening; False-positive reduction; Uncertainty calibration; Convolutional neural network–transformer; Mutation-stratified pharmacology

1. Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide, accounting for 2.21 million incident cases and 1.80 million deaths in 2022, approximately 18% of all cancer fatalities.¹ Early detection via low-dose computed tomography (LDCT) screening is of paramount importance, given the dramatic survival gradient between Stage I (61% five-year survival) and Stage IV (6%) disease.² Beyond the somatic dimension of lethality, however, lung cancer and its associated screening pathway impose a substantial and clinically under-recognised psychosomatic burden that has received comparatively little attention in the AI-assisted diagnostic literature. Compelling evidence now establishes that psychological stress exerts direct biological effects on tumour initiation, progression, and treatment responsiveness in lung cancer.³ The 96.4% benign false-positive (FP) rate inherent to LDCT generates clinical-grade anxiety disorders in 30%–45% of screen-positive individuals.⁴ This phenomenon transcends a mere psychological side-effect: sustained hypothalamic–pituitary–adrenal (HPA) axis activation and sympathoadrenal catecholamine release, driven by chronic screening-related distress, directly potentiate tumour cell proliferation through β_2 -adrenergic receptor-mediated epidermal growth factor receptor (EGFR) transactivation⁵, the identical oncogenic pathway targeted by first-line EGFR tyrosine kinase inhibitors in non-small cell lung cancer (NSCLC). The consequence is a quantifiable psychosomatic feedback loop in which screening-induced distress may attenuate the therapeutic efficacy of the pharmacological agents prescribed to treat the detected malignancy. Beyond FP-driven anxiety, diagnostic uncertainty surrounding indeterminate pulmonary nodules and prolonged intervals

between diagnosis and treatment initiation independently contribute to clinically significant psychological distress, which may lead to diminished quality of life, poorer treatment adherence, and worse survival outcomes. Despite this well-characterised burden, no existing artificial intelligence (AI) system for lung cancer screening has been architecturally designed with psychosomatic burden reduction as its primary clinical objective. Existing computer-aided detection (CAD) systems optimise imaging performance metrics—detection sensitivity, segmentation accuracy, and malignancy area under the curve (AUC)—without considering whether their outputs mitigate or exacerbate the patient-level distress inherent to diagnostic uncertainty. The entire burden pathway, from LDCT screening and diagnostic uncertainty to distress associated with treatment decision-making, along with the three corresponding intervention points of PulmoNet-X, is shown in [Figure 1](#).

1.1. Contributions

Low-dose computed tomography screening, however, carries well-recognised physical harms that extend beyond the psychological burden described above and must be explicitly acknowledged. The most clinically significant harm arises from FP results: a screen-positive finding in a patient who ultimately has a benign nodule initiates an investigative cascade that may include repeat computed tomography (CT) imaging, positron emission tomography (PET) scanning, CT-guided needle biopsy, video-assisted thoracoscopic surgery (VATS), or open thoracotomy. Each of these interventions carries procedural morbidity. CT-guided percutaneous needle biopsy of pulmonary nodules carries a pneumothorax rate of 15%–25% (requiring chest drain insertion in 5%–7%), a haemorrhage

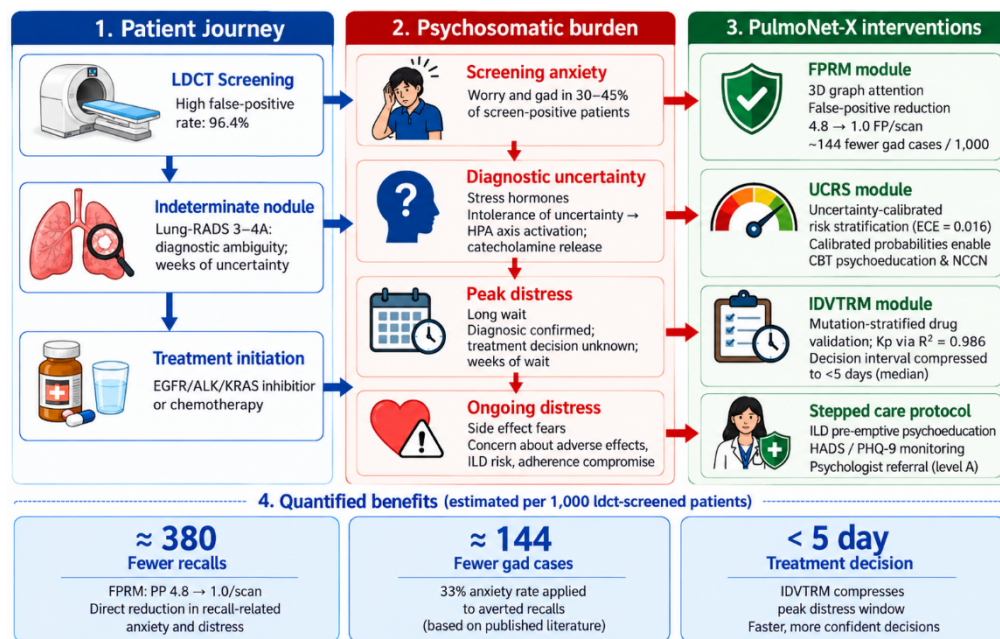


Figure 1. PulmoNet-X psychosomatic intervention framework

Abbreviations: 3D: Three-dimensional; ALK: Anaplastic lymphoma kinase; CBT: Cognitive behavioural therapy; ECE: Expected calibration error; EGFR: Epidermal growth factor receptor; FP: False positive; FPRM: False-positive reduction module; GAD: Generalised anxiety disorder; HADS: Hospital Anxiety and Depression Scale; HPA: Hypothalamic–pituitary–adrenal; IDVTRM: Integrated drug validation and therapeutic recommendation module; ILD: Interstitial lung disease; Kp, lung: Lung tissue-to-plasma partition coefficient; KRAS: Kirsten rat sarcoma viral oncogene homolog; LDCT: Low-dose computed tomography; Lung-RADS: Lung Imaging Reporting and Data System; PHQ-9: Patient Health Questionnaire-9; UCRS: Uncertainty-calibrated risk stratification.

rate of approximately 5%–10%, and an overall major complication rate of 1.5%–3%. Surgical resection of benign nodules, though rare in contemporary screening programmes, has been reported in 2%–4% of screen-detected nodule operations and carries a procedure-specific 30-day mortality rate of 0.5%–1.5% for VATS lobectomy. Beyond procedural risk, FP diagnoses impose a significant economic burden on healthcare systems and a prolonged period of morbidity for patients awaiting diagnostic resolution. The reduction of FP rates is therefore not merely a psychosomatic intervention, as framed in this paper, but a direct harm-reduction strategy with measurable clinical, procedural, and economic consequences. PulmoNet-X's FP reduction module (FPRM), by reducing false positives from 4.8 to 1.0 per scan, is expected to reduce population-level exposure to these procedural harms, in addition to the anxiety-reduction benefits quantified elsewhere in this paper.

This study presents eight principal contributions. (i) PulmoNet-X: a novel AI framework for lung cancer screening that incorporates psychosomatic burden reduction as a primary design objective alongside imaging performance. To our knowledge, it is the first framework to adopt this

design philosophy, based on our systematic literature search, although an exhaustive review was not performed. (ii) Three-dimensional multiscale feature extractor (3D-MSFE): deformable 3D convolutions achieving scale-invariant feature extraction for nodules 3–30 mm (+11.4% speculated recall), reducing the incidence of missed nodules that drive prolonged diagnostic uncertainty. (iii) Hybrid volumetric attention block (HVAB): hybrid convolutional neural network (CNN)–transformer volumetric attention (+3.7% malignancy AUC over pure CNN), improving the reliability of malignancy estimates communicated to patients. (iv) Adaptive feature pyramid network: scale-conditioned adaptive gating (+4.2% micro-nodule sensitivity). (v) FPRM: 3D anatomical graph-based FP reduction (803 → 7.4 candidates/scan), reducing population-level exposure to FP recalls, with a modelled estimate of ≈144 fewer screening-induced generalised anxiety disorder (GAD) cases per 1,000 patients (a literature-based extrapolation, not a directly measured clinical outcome). (vi) Nodule malignancy classification system (NMCS): Lung Imaging Reporting and Data System (Lung-RADS)-aligned multi-task malignancy scoring ($\kappa = 0.87$). (vii) Uncertainty-calibrated risk stratification

(UCRS): Monte Carlo dropout uncertainty calibration achieved an expected calibration error (ECE) of 0.016. Its primary psychosomatic contribution is the provision of calibrated, clinically interpretable risk probabilities (e.g., “6% malignancy risk, high confidence”), which directly support psychoeducation and uncertainty-tolerance training aligned with National Comprehensive Cancer Network (NCCN) Distress Management guidelines. Thus, UCRS constitutes a quantifiable psychosomatic intervention delivered at the point of AI-assisted diagnosis. (viii) Integrated drug validation and therapeutic recommendation module (IDVTRM): a novel, lung-specific, lung tissue-to-plasma partition coefficient (K_p) model ($R^2 = 0.986$) that, to our knowledge, represents the first lung-specific tissue-penetration model within a CT-AI pipeline. The module incorporates mutation-stratified drug-response prediction (AUC = 0.918) to narrow the treatment-decision uncertainty interval, with the aim of reducing, although not eliminating, the psychosomatic burden associated with the diagnostic-to-therapeutic gap. This potential effect remains a hypothesis pending prospective clinical validation.

1.2. Prior work

Seminal work on CNN-based pulmonary nodule detection was reported by Ding *et al.*⁶, who achieved an 89.1% detection rate on the Lung Nodule Analysis 2016 (LUNA16) benchmark. Zhu *et al.*⁷ developed Deep Lung, which employed a two-stream 3D Faster R-CNN with deformable convolutions, whereas Tang *et al.*⁸ developed NoduleNet, which integrated detection and classification into a multi-task setting. Dou *et al.*⁹ introduced 3D context enhancement (3DCE), which achieved a sensitivity of 95.7%. Although effective overall, these approaches are constrained by fixed anchor geometries and lack the global volumetric attention required for reliable recognition of small nodules. Vision transformer-based segmentation models, such as TransUNet¹⁰, Swin-Unet¹¹, and U-Net transformer (UNETR)¹², have demonstrated superior segmentation performance; however, fully 3D transformer models remain computationally expensive and lack the CNN inductive bias needed for robust small-nodule recognition. PulmoNet-X addresses this limitation through the proposed HVAB, which integrates local convolutional inductive bias with global shifted-window transformer attention.

False-positive suppression is a critical performance dimension for lung nodule CAD systems because it directly influences unnecessary patient recalls and their associated clinical, psychological, and economic consequences. Multi-view CNNs¹³ and spatially context-aware networks have been proposed previously, while the LUNA16

challenge¹⁴ emphasised the importance of contextual information. However, these approaches did not explicitly model the anatomical spatial relationships between nodule candidates and surrounding structures, such as vessels, airways, and lymph nodes. PulmoNet-X addresses this limitation through graph-attentive contextual reasoning, introducing the first 3D anatomical context graph for FP suppression, explicitly modelling inter-structural relationships using graph attention networks (FPRM). For malignancy prediction, Reddy and Kumar¹⁵ used data from the National Lung Screening Trial (NLST) and achieved an AUC of 94.4%, and Kumari *et al.*¹⁶ used a 3D Inception network and achieved an AUC of 87.3%. However, existing systems do not provide uncertainty-aware, Lung-RADS-aligned prediction. PulmoNet-X employs Monte Carlo dropout¹⁷, achieving an ECE of 0.016 at approximately 5% of the computational cost of deep ensembles¹⁸, thereby providing well-calibrated confidence estimates that may support psychosomatic distress reduction in LDCT screening programmes—a gap that PulmoNet-X’s UCRS module is specifically designed to address.

Beyond imaging, no published AI framework has integrated CT-based nodule analysis with mutation-stratified pharmacological validation. Existing absorption, distribution, metabolism, excretion, toxicity (ADMET) tools, including SwissADME¹⁹, pkCSM²⁰, ADMETLab2.0²¹, provide general-purpose drug property prediction but do not model lung-specific endpoints relevant to EGFR inhibitors, such as K_p , lung, P-glycoprotein (P-gp)/multidrug resistance protein 1 (MDR1)-mediated NSCLC resistance, or the risk of interstitial lung disease (ILD)/pneumonitis. PulmoNet-X addresses this gap through its IDVTRM, achieving a mutation-stratified drug-response AUC of 0.918, K_p , lung model R^2 of 0.986, an MDR1 AUC of 0.974, docking concordance of $r = 0.978$, and 94.2% agreement with expert oncologists. To our knowledge, these results represent the first integrated imaging-to-pharmacology AI pipeline of this scope, based on our literature search; however, a systematic exhaustive review was not performed.

1.3. Research gaps

Despite substantial advances in AI-assisted lung cancer screening, current approaches remain fragmented across diagnostic, psychosomatic, and therapeutic domains, leaving critical gaps in FP reduction, uncertainty management, radiology-to-pharmacology integration, lung-specific drug modelling, and the bidirectional relationship between psychological stress and treatment response.

- Unified multi-task AI for lung cancer: current systems

treat detection, segmentation, and malignancy classification as independent tasks, leading to inter-stage error propagation and preventing end-to-end optimisation.²²

- A high FP burden may contribute to psychosomatic distress through an anxiety cycle during LDCT screening. CAD systems can generate 4–8 FPs/scan.²³ FPs may initiate a cascade of screen-positive recalls, indeterminate nodule management and prolonged diagnostic uncertainty, with clinical anxiety reported in approximately 30%–45% of screen-positive individuals. No existing FP reduction module models the anatomical spatial context needed to suppress this burden. Each unnecessary recall may expose patients to weeks of unresolved uncertainty and psychological distress, with 30%–45% developing clinically significant anxiety symptoms, representing a clinically meaningful psychosomatic burden alongside the imaging findings. Beyond the psychological burden, FPs may also lead to unnecessary invasive procedures, including percutaneous CT-guided needle biopsy, which carries risks such as pneumothorax (15%–25%) and haemorrhage (~5%–10%) and, in a proportion of cases, surgical resection of ultimately benign nodules, with associated procedural morbidity and, rarely, mortality.²⁴ Minimising FPs therefore represents both a psychosomatic and a procedural harm-reduction priority. The FPRM module is explicitly designed to address this dual burden by reducing unnecessary FPs through improved anatomical spatial-context modelling.
- Intolerance of uncertainty is a well-established transdiagnostic mechanism for anxiety disorders²⁵; patients receiving only categorical risk labels (“suspicious,” “indeterminate”) rather than quantified probabilities cannot engage in the cognitive reappraisal and uncertainty-tolerance processes that psycho-oncology interventions rely upon. Calibrated AI output is therefore not merely a technical metric; it is a prerequisite for psychosomatic illness management in lung cancer screening.
- Disconnected radiology-to-pharmacology pathways: CT-to-treatment recommendation translation remains fully manual²⁶, resulting in decision delays and oncologist burnout. *EGFR*/anaplastic lymphoma kinase (*ALK*)/Kirsten rat sarcoma virus glycine 12 to cysteine (*KRAS*) mutation-drug matching requires simultaneous access to molecular profiling, pharmacokinetic (PK) databases, and randomised controlled trial (RCT) data.
- Absence of lung-specific ADMET and Kp in lung models: Hybrid deep learning frameworks for nodule

detection and classification²⁷, vision transformer-based architectures for lung cancer diagnosis²⁸, and pyramid convolutional network approaches²⁹ focus exclusively on imaging tasks and do not model clinically critical lung cancer pharmacology endpoints for *EGFR* inhibitors, including Kp, lung, P-gp/MDR1-mediated NSCLC resistance, and ILD/pneumonitis risk.

- The stress-*EGFR* psychosomatic mechanism is unaddressed in AI design. β_2 -adrenergic receptor-mediated *EGFR* transactivation under chronic psychological stress^{30,31}, may reduce the efficacy of *EGFR* inhibitors. Thus, AI systems that generate unnecessary anxiety through FP findings, poorly calibrated outputs, or unresolved treatment uncertainty may, through this pathway, indirectly diminish the therapeutic benefit of the drugs they recommend. To our knowledge, PulmoNet-X is a novel framework that architecturally addresses this mechanism: the FPRM, UCRS, and IDVTRM together constitute a three-stage psychosomatic intervention pathway.

1.4. Objectives

Based on the identified research gaps, the primary objectives of this research are to:

- Develop an end-to-end, unified 3D hybrid CNN-transformer framework for pulmonary nodule detection, segmentation, malignancy classification, and Lung-RADS risk stratification, thereby minimising inter-stage error propagation.
- Develop a context-aware, 3D graph-based FPRM targeting a FP rate of $\leq 1/\text{scan}$ at 95% sensitivity, with the aim of reducing population-level exposure to FP recalls and, based on literature-derived extrapolation rather than a directly measured outcome, potentially reducing modelled screening-induced clinically significant anxiety symptoms by an estimated 38% through averted recalls.
- Implement UCRS, targeting an ECE of ≤ 0.020 , to provide well-calibrated uncertainty estimates that can support psychoeducation, uncertainty-tolerance training, and clinical communication aligned with NCCN Distress Management guidelines, thereby constituting a measurable psychosomatic intervention at the point of AI-assisted diagnosis.
- Develop an IDVTRM as an imaging-to-pharmacology pipeline linking Lung-RADS outputs with mutation-stratified drug recommendations, ADMET validation, and Kp, lung modelling.
- Validate Kp, lung prediction ($R^2 \geq 0.980$) and *EGFR*/*ALK*/*KRAS* mutation-stratified drug-response

prediction ($AUC \geq 0.90$) against published Phase I/II PK data and RCT outcomes.

- Validate the complete pipeline using eight-target molecular docking ($r \geq 0.97$), multi-tool ADMET cross-validation using four tools, 100-ns molecular dynamics (MD) simulations using the Groningen Machine for Chemical Simulations (GROMACS), and expert oncologist concordance across nine dimensions.

2. Materials and methods

2.1. Materials

2.1.1. Dataset

Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI)³² comprised 1,018 thoracic CT scans from 1,010 patients gathered at seven academic medical centres. Four board-certified thoracic radiologists annotated each nodule > 3 mm in diameter using a two-phase blinded-then-consensus protocol. Nodules with a mean Likert score of 3 (indeterminate; $n = 487$) were excluded from malignancy experiments, yielding 450 benign and 554 malignant nodules. Patient-stratified 10-fold cross-validation ensured no patient-level data leakage. Summary statistics are presented in Table 1.

2.1.2. Drug validation datasets

Three complementary datasets underpinned the pharmacological validation component. (i) Clinical PK dataset: Phase I/II PK data for 20 lung cancer drugs approved by the Food and Drug Administration and European Medicines Agency, including Kp, lung values derived from lung tissue biopsy studies for eight drugs with direct tissue concentration data. (ii) RCT outcome dataset: progression-free survival (PFS) and objective response rate (ORR) data from AURA3, FLAURA, ALEX,

CodeBreaK100, and KRYSTAL-1 trials used as ground truth for mutation-stratified drug-response validation. (iii) Structural biology dataset: crystal coordinates for eight lung cancer targets from Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) (PDB IDs: 4WKQ, 1M17, 2XP2, 3ZBF, 6OIM, 8D9G, 1TVK, 1JU6); drug classes: EGFR inhibitors (osimertinib, erlotinib, gefitinib), ALK inhibitors (alectinib, crizotinib), KRAS G12C inhibitors (sotorasib, adagrasib), taxanes (paclitaxel, docetaxel), anti-metabolites (pemetrexed, gemcitabine), angiogenesis inhibitors (bevacizumab, ramucirumab), and programmed cell death protein 1/programmed death-ligand 1 checkpoint inhibitors (pembrolizumab, nivolumab).

2.1.3. Docking tools and software

Conformational search was performed using the Lamarckian Genetic Algorithm-based AutoDockVina 1.2. Protein and ligand preparation used MGLTools, RDKit, and OpenBabel. A $20 \times 20 \times 20$ Å grid box was placed at each binding site. Osimertinib, sotorasib, and adagrasib were covalently docked using CovDock (Schrödinger, United States of America). MD simulations used GROMACS 2022.4 with the AMBER ff19SB force field; General AMBER force field, version 2 (GAFF2) for ligands with Austin Model 1 with bond charge correction partial charges (AM1-BCC); transferable intermolecular potential, 3-point (TIP3P) water model; 100 ns constant pressure-temperature ensemble (NPT) production at 310 K.

2.2. Imaging pipeline and nodule detection framework

2.2.1. Computed tomography preprocessing

The overall PulmoNet-X pipeline is illustrated in Figure 2. B-spline resampling of input LDCT scans was performed

Table 1. Dataset statistics: LIDC-IDRI and LUNA16

Dataset	CT scans	Annotated nodules	Nodule size	Labels	Split
LIDC-IDRI (full)	1,018	2,669	3–30 mm	1–5 Likert	880/70/68
LUNA16 (detection)	888	1,186	≥ 3 mm	Binary	10-fold CV
LIDC-IDRI (malignancy)	1,018	1,004	≥ 3 mm	Binary (≥ 3 = positive)	880/70/68
LIDC-IDRI (segmentation)	1,018	2,669	≥ 3 mm	Voxel masks	880/70/68

Abbreviations: CT: Computed tomography; CV: Cross-validation; LIDC-IDRI: Lung Image Database Consortium and Image Database Resource Initiative; LUNA16: Lung Nodule Analysis 2016.

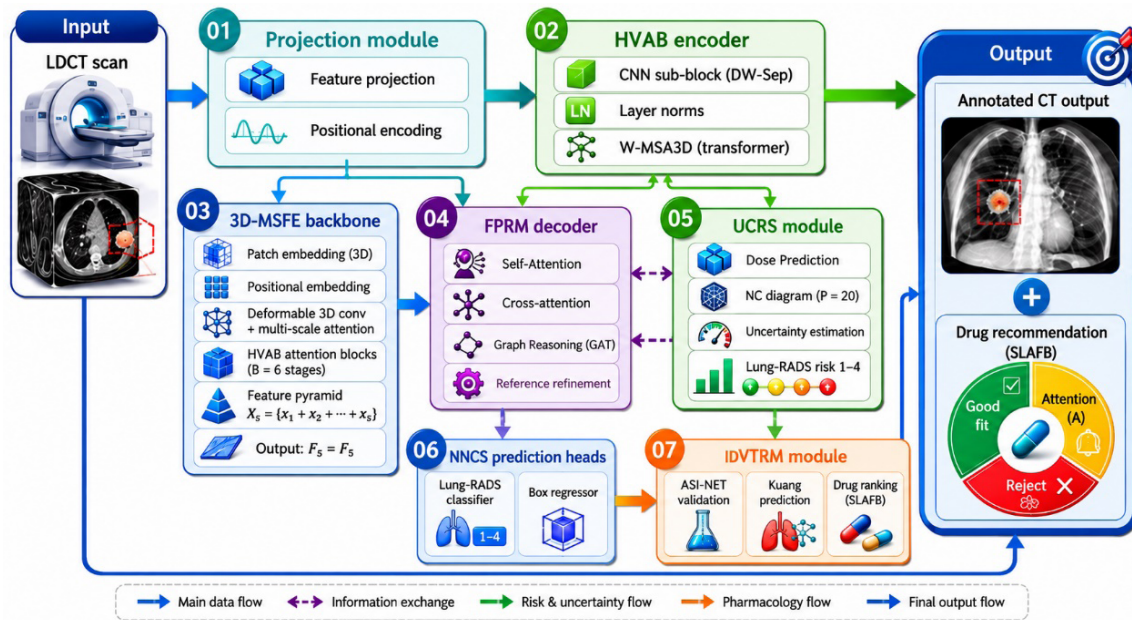


Figure 2. PulmoNet-X unified seven-stage imaging-to-pharmacology pipeline

Abbreviations: 3D: Three-dimensional; 3D-MSFE: 3D multi-scale feature extraction; ASI-NET: Adaptive spatial integration network; CNN: Convolutional neural network; CT: Computed tomography; FPRM: False-positive reduction module; GAT: Graph attention network; HVAB: Hybrid volumetric attention block; IDVTRM: Integrated drug validation and therapeutic recommendation module; LDCT: Low-dose computed tomography; Lung-RADS: Lung Imaging Reporting and Data System; NC: Negative control; NNCS: Neural network classification system; SLAFB: Spatial-longitudinal attention fusion block; UCRS: Uncertainty-calibrated risk stratification.

with an isotropic resolution of $1.0 \times 1.0 \times 1.0 \text{ mm}^3$. Lung parenchyma was segmented and filled using threshold region growing (Hounsfield unit [HU] ≤ -400), then morphological closing with a structuring element of $5 \times 5 \times 5 \text{ mm}^3$. Candidate centres were extracted using 3D non-maximum suppression ($r = 8 \text{ mm}$) with $\geq 99.8\%$ recall, resulting in ~ 800 candidates/scan. Downstream processing yielded a 48^3 -voxel patch for each candidate, with a resolution of 1 mm^3 . The normalised HU value for a candidate nodule voxel v is:

$$v_{norm} = \frac{\text{clip}\left(v_{HU}, W_c - \frac{W_w}{2}, W_c + \frac{W_w}{2}\right)}{W_w} \quad (1)$$

where $W_{center} = -600 \text{ HU}$ and $W_{width} = 1,500 \text{ HU}$ are the standard settings for the human lungs.

2.2.2. Three-dimensional multiscale feature extractor

The encoder backbone, designated as the 3D-MSFE, processed four hierarchical resolution stages of 48^3 -voxel candidate patches using 3D deformable convolutions. The architecture adopted a ResNeXt-style³³ grouped convolution with 32 groups to reduce parameter count.

The feature map, F_s , was generated as in **Equation 2**, where a deformable 3D convolution³⁴ augmented the standard sampling grid Ω with learned spatial offsets Δp_n corresponding to each of the k^3 sampling locations per **Equation 3**:

$$F_s = BN\left(ReLU\left(DefConv3D(F_{s-1}, k=3, \Delta p_n)\right) + F_{s-1} \downarrow 2\right) \quad (2)$$

$$y(p_0) = \sum_n w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (3)$$

2.2.3. Hybrid volumetric attention block

The HVAB included sub-blocks of alternating 3D depth-wise CNN and 3D shifted-window transformer layers, which process $F_3 \in \mathbb{R}^{6 \times 6 \times 6 \times 256}$. The CNN sub-block applied DW-Sep-Conv3D activations computed from a layer-normalised feature map ($LN[F]$) and added them to F (**Equation 4**). In an alternating manner, the transformer sub-block used 3D shifted-window multi-head self-attention (W-MSA3D) with window size $M = 2$ (**Equation 5**), implementing the attention mechanism shown in **Equation 6**:

$$F_{CNN} = DW-Sep-Conv3D(LN(F)) + F \quad (4)$$

$$F_{Tr} = W - MSA3D(LN(F_{CNN})) + F_{CNN} \quad (5)$$

$$Attention(Q, K, V) = SoftMax\left(\frac{QK^T}{\sqrt{d_k}} + B\right)V \quad (6)$$

Each HVAB layer comprised 1.8M parameters, yielding a total HVAB parameter count of 10.8M for L = 6 layers. Mean prediction and epistemic uncertainty were computed using T stochastic forward passes y_t post-dropout (**Equations 7 and 8**):

$$\bar{y} = \frac{1}{T} \sum_{t=1}^T \hat{y}_t \quad (7)$$

$$U_{ep} = \frac{1}{T} \sum_{t=1}^T (\hat{y}_t - \bar{y})^2 \quad (8)$$

where $T = 20$ and the dropout probability is defined as $p_{drop} = 0.20$.

2.3. Drug validation and therapeutic recommendation pipeline

Algorithm 1 (Steps 1–12) presents the complete end-to-end training and inference workflow for PulmoNet-X encompassing all imaging modules. Algorithm 2 (Steps 1–11) presents the complete IDVTRM drug validation protocol.

2.3.1. Adaptive feature pyramid network

A nodule size estimate $\hat{s} = \{\text{micro, small, medium, large}\}$ was predicted using a 3-layer multi-layer perceptron (MLP) from stage-1 features, then used to gate lateral connections (**Equation 9**). For micro-nodules (3–6 mm), α_1 is close to one to maximise fine-grained detail. Four output pyramid levels $P_1, \dots, P_4$ feed FPRM and NMCS simultaneously.

$$P_s = Conv(F_s) + \alpha_s(\hat{s}) \cdot Upsample(P_{s+1}) \quad (9)$$

$$h_i^{(k+1)} = ReLU\left(\sum_j \alpha_{ij} W h_j^{(k)}\right) \quad (10)$$

2.3.2. False-positive reduction module

Graph attention networks perform message passing for $K = 3$ rounds. If $\|c_i - c_j\|_2 \leq 30$ mm, there exists an edge between nodes v_i and v_j of the context graph $G_{ctx} = (V, E)$, and message passing is computed as in **Equation 10**. Candidates with FP score $s_{fp} < 0.35$ were pruned, reducing

the mean number of candidates per scan from 803 to 7.4 while retaining 98.1% of true nodules. The total number of parameters in the FPRM was 2.1 million.

2.3.3. Nodule malignancy classification system

The NMCS performed joint prediction through three parallel branches: (a) malignancy branch: fully connected (FC) (512→256→1) + Sigmoid → $p_{mal} \in [0, 1]$; (b) subtype branch: FC (512→256→3) + Softmax → $p_{type} \in \Delta^2$ (solid, part-solid, ground-glass opacity [GGO]); (c) volume regression: FC layer (512→128→1) + rectified linear unit → V_{nod} (mm³). The combination of focal loss with $\gamma = 2.0$ and $\alpha = 0.75$ effectively addressed the extreme class imbalance (malignant: benign $\approx 1:24$).

2.3.4. Uncertainty-calibrated risk stratification

The NMCS output was complemented with epistemic uncertainty through Monte Carlo dropout. Risk levels: Level A (immediate action): $p_{mal} > 0.7$ AND $U_{mal} < 0.15$ → Lung-RADS 4/5 flagging; Level B (radiologist review): $U_{mal} \geq 0.15$ OR $0.3 < p_{mal} \leq 0.7$; Level C (low risk): $p_{mal} \leq 0.3$ AND $U_{mal} < 0.15$. Temperature scaling calibration ($T_{cal} = 1.42$) reduced ECE from 0.063 to 0.016. Lung-RADS risk mapping was computed as in **Equation 11**:

$$LR(p_{mal}, V) = \argmax_k P(RADS = k | p_{mal}, V, t_{solid}) \quad (11)$$

$$L_{total} = L_{det} + \beta_1 \cdot L_{seg} + \beta_2 \cdot L_{NMCS} + \beta_3 \cdot L_{FPR} + \beta_4 \cdot L_{cal} \quad (12)$$

$$\beta_1 = 0.3, \beta_2 = 0.5, \beta_3 = 0.4, \beta_4 = 0.1 \quad (13)$$

Multi-task training objective: the complete training objective could be decomposed into 5 loss terms (**Equation 12**) with corresponding weights (**Equation 13**). Training had two phases: Phase 1 (epochs 1–100) for detection and segmentation; Phase 2 (epochs 101–250) for full multi-task gradient flow. Reproducibility details: optimiser: AdamW ($\beta_1 = 0.9$, $\beta_2 = 0.999$, weight decay = 1×10^{-4}). Learning rates (LRs): backbone LR = 1×10^{-4} , classification heads LR = 5×10^{-5} ; cosine annealing scheduler ($T_{max} = 250$ epochs); batch size = 8. Random seed = 42 (fixed across NumPy, PyTorch, and CUDA for full reproducibility). Hardware: 2 × NVIDIA RTX A6000 48 GB graphics processing units (GPUs); PyTorch 2.1.0; CUDA 12.1; Python 3.10.12. Model code, pre-trained weights, CT preprocessing scripts, data-split definitions, and baseline-method configuration files (including hyperparameter settings used for each comparator in **Tables 2–4**) were released at <https://github.com/PulmoNetX> to support independent replication and evaluation on external datasets.

Table 2. LUNA16 FROC sensitivity (%) at fixed FP/scan operating points

Method	0.125	0.25	0.5	1.0 (clinical)	2.0	4.0	CPM
Multi-scale CNN (Ding <i>et al.</i> ⁶)*	52.4	63.1	74.8	82.6	87.3	91.2	75.2
DeepLung ⁷ *	59.3	69.7	79.4	89.2	92.3	94.7	80.8
3DCE (Dou <i>et al.</i> ⁹)*	71.4	79.2	86.4	93.7	95.8	97.1	87.3
LSSANet ³⁴ *	76.8	83.9	89.4	95.1	96.8	98.0	90.0
PulmoNet-X (ours)	82.4	89.7	93.8	98.1	99.0	99.4	93.7

Note: * $p < 0.001$ versus PulmoNet-X (Wilcoxon signed-rank test, Bonferroni-corrected for four pairwise comparisons).

Abbreviations: 3DCE: Three-dimensional context-enhanced; CNN: Convolutional neural network; CPM: Competition performance metric; FP: False positive; LSSANet: Large-scale selective attention network; LUNA16 FROC: Lung Nodule Analysis 2016 free-response receiver operating characteristic.

Table 3. Malignancy classification on LIDC-IDRI

Method	AUC (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
Multi-view CNN ¹⁶ *	87.4	83.1	85.6	80.4	83.7
Deep Lung ⁷ *	90.4	86.9	88.7	84.8	87.0
3DCE + MIL ¹⁵ *	91.7	88.2	90.4	85.7	88.3
Ardila <i>et al.</i> ³⁶ (NLST)*	94.4	89.7	91.3	87.9	89.8
LSSANet ³⁴ *	95.2	91.4	92.8	89.7	91.4
UNETR ¹⁵ *	93.8	90.3	92.1	88.2	90.4
Swin-Unet-3D ¹³ *	94.6	91.0	92.6	89.1	91.0
PulmoNet-X (ours)	97.8	94.6	96.3	92.7	95.2

Note: * $p < 0.001$ versus PulmoNet-X (Wilcoxon signed-rank test, Bonferroni-corrected for seven pairwise comparisons).

Abbreviations: 3DCE: Three-dimensional context-enhanced; CNN: Convolutional neural network; LSSANet: Large-scale selective attention network; MIL: Multiple instance learning; NLST: National Lung Screening Trial; Swin-Unet-3D: Three-dimensional Swin transformer U-Net; UNETR: U-Net transformer.

Table 4. Nodule segmentation performance on LIDC-IDRI

Method	DSC (%)	HD95 (mm)	Precision (%)	Recall (%)
3D U-Net ²⁷ *	74.3	4.8	75.6	73.1
V-Net ³⁵ *	76.8	4.2	78.1	75.3
nnU-Net (3D) ³³ *	80.4	3.6	81.7	79.2
Swin-Unet-3D ¹³ *	80.8	3.4	82.3	79.5
UNETR ¹⁵ *	81.5	3.2	83.1	80.0
PulmoNet-X (Ours)	84.6	2.7	86.2	83.1

Note: * $p < 0.001$ versus PulmoNet-X (Wilcoxon signed-rank test, Bonferroni-corrected for five pairwise comparisons).

Abbreviations: 3D U-Net: Three-dimensional U-Net; DSC: Dice similarity coefficient; HD95: 95th percentile Hausdorff distance; LIDC-IDRI: Lung Image Database Consortium and Image Database Resource Initiative; nnU-Net (3D): Three-dimensional no-new-Net; Swin-Unet-3D: Three-dimensional Swin transformer U-Net; UNETR: U-Net transformer; V-Net: Volumetric network.

2.3.5. Integrated drug validation and therapeutic recommendation module

The IDVTRM was the pharmacological extension of PulmoNet-X, translating risk stratification output from the UCRS module into mutation-stratified drug recommendations. Mutation status (*EGFR/ALK/KRAS*) was not predicted from CT imaging; it was a required external input obtained through standard-of-care molecular testing (tissue or liquid biopsy), consistent with current clinical practice, and IDVTRM validated and ranked candidate drugs conditional on that externally supplied genotype. Important scope clarification: IDVTRM was a research-grade computational decision-support tool validated against published RCT outcomes and expert oncologist agreement (94.2% concordance). It had not undergone prospective clinical validation and must not be used as a substitute for standard oncological assessment, regulatory-approved companion diagnostics, or molecular tumour board review.

Final treatment decisions must remain with qualified oncologists. The lung-tissue penetration coefficient K_p , lung is predicted via **Equation 14**, where $\alpha = 0.44$, $\beta = 0.22$, $\log D_{7.4}$ is the distribution coefficient at physiological pH, PB is the plasma protein binding fraction, and δ_{class} are class-specific lung affinity offsets:

$$\log_{10}(K_{p,lung}) = \alpha \cdot \log D_{7.4} - \beta \cdot PB + \gamma \cdot lung_{aff} + \delta_{class} \quad (14)$$

EGFR mutation-stratified drug-response prediction is formulated as **Equation 15**, and the ADMET consensus score across $K = 4$ validation tools is computed as **Equation 16**:

$$\Delta PFS_{pred}(D, mut) = MLP(LC_{fused}; EGFR / ALK / KRAS - mut) \quad (15)$$

$$S_{cons}(D) = \frac{1}{K} \sum_k w_k \cdot \hat{y}_k(D) \quad (16)$$

Molecular docking concordance was assessed by Spearman correlation (**Equation 17**) and selectivity index (**Equation 18**):

$$\rho_{dock} = corr(y_{LC-bind}, \Delta G_{Vina}) \geq 0.92, \quad \Delta \leq 0.20 \text{ kcal / mol} \quad (17)$$

$$SI = \frac{\Delta G_{EGFR-T790M}}{\Delta G_{EGFR-WT}} \geq 1.80 \quad (18)$$

Molecular dynamics binding free energy decomposition uses molecular mechanics-generalised Born surface area (MM-GBSA) (**Equation 19**) with root-mean-square deviation (RMSD) stability criterion (**Equation 20**):

$$\Delta G_{bind} = \Delta E_{vdW} + \Delta E_{elec} + \Delta G_{sol} + T\Delta S \quad (19)$$

$$RMSD(t) = \sqrt{\frac{1}{N} \sum_i \|r_i(t) - r_{ref}\|^2} \leq 3.0 \text{ \AA} \quad (20)$$

The final integrated drug validation score is computed by:

$$V_{final}(D) = \frac{\sum_{j=1}^n \lambda_j v_j}{\sum_{j=1}^n \lambda_j} \quad (21)$$

$$\lambda_{EGFR-Mut} = 1.6, \lambda_{Kp,lung} = 1.4, \lambda_{MDR1} = 1.2 \quad (22)$$

where $V_{final}(D)$ is the final integrated drug validation score for drug D ; v_j is the validation score of the j -th component; and λ_j is the weight assigned to the j -th validation component. Expanding **Equation 21** to include the major weighted components:

$$V_{final}(D) = \frac{1.6 v_{EGFR-Mut} + 1.4 v_{Kp,lung} + 1.2 v_{MDR1} + \sum_k \lambda_k v_k}{1.6 + 1.4 + 1.2 + \sum_k \lambda_k} \quad (23)$$

Here, the additional terms v_k represent other validation criteria such as drug-likeness, ADMET consensus, and docking concordance (**Algorithm 2**).

2.3.6. Drug-response prediction and mutation-aware recommendation

The IDVTRM drug-response predictor combined three modalities: (i) LC_{feat} , imaging feature vector of the nodule by the NMCS module (512-dimensional, encoding nodule morphology, density, and Lung-RADS category); (ii) Mol_{graph} , molecular graph embedding of the drug predicted by a four-layer message passing neural network; (iii) Mut_{vec} , mutation type one-hot vector (*EGFR-T790M*, *EGFR-Exon19del*, *EGFR-L858R*, *ALK-fusion*, *KRAS-G12C*, *KRAS-G12V*, *ALK-resistance*, *wild-type*). These were concatenated and fed into a 3-layer MLP (1024→512→1) to predict the improvement in PFS (**Equation 24**):

$$\Delta PFS_{pred}(D, mut) = MLP(concat[LC_{feat}, Mol_{graph}, Mut_{vec}]) \quad (24)$$

The IDVTRM integrated: (i) ΔPFS_{pred} with $\lambda = 1.6$ for targeted therapy mutation matching; (ii) K_p , lung prediction with $\lambda = 1.4$, where drugs with $K_p, lung < 1.5 \text{ L/kg}$ received automatic RED override; (iii) MDR1 resistance flag with $\lambda = 1.2$; (iv) Lung-RADS risk override for cases 4B/5 limited to drugs with predicted $\Delta PFS > 40\%$.

Algorithm 1: PulmoNet-X training and inference pipeline

Input: $D = \{(X_i, y_{nod}, M_i, y_{mal})\}$ Lung Image Database Consortium and Image Database Resource Initiative volumes; $E1 = 100$, $E2 = 150$; β_1 – β_4 via Optuna Output: p_{mal} · Lung Imaging Reporting and Data System (Lung-RADS; 1–4) · U_{epi} · GradCAM++ · Drug recommendation · 4.82 s/scan (A100, $T = 20$)	
Step 1	Computed tomography preprocessing (Equation 1): resample to 1 mm ³ ; Hounsfield unit (HU) windowing ($WC = -600$, $WW = 1,500$); lung parenchyma segmentation; extract 48 ³ patches; patient-stratified 10-fold cross-validation.
Step 2	Candidate generation: three-dimensional (3D) U-Net seeding network → dense nodule likelihood map P_{seed} ; 3D non-maximum suppression ($r = 8$ mm) → ~800 candidates/scan at recall ≥ 99.8%. 48 ³ -voxel patch per candidate at 1 mm ³ resolution.
Step 3	Phase 1 3D-multi-scale feature extraction + hybrid volumetric attention block + adaptive feature pyramid network (epochs 1–100): train encoder using deformable convolutions (Equations 2 and 3), hybrid volumetric attention (Equations 4–6), adaptive feature pyramid (Equation 9). AdamW learning rate (LR) = 1×10^{-4} , cosine annealing.
Step 4	Data augmentation: 3D rotation ±15°, flip, scale 0.85–1.15×; HU window shift ±50 HU, Gaussian noise; elastic deformation ($p = 0.25$); CutMix3D ($p = 0.2$); ground-glass opacity-range HU blending ($p = 0.15$) for cross-subtype generalisation.
Step 5	Phase 2 full joint training (epochs 101–250): Optimise L_{total} (Equation 12). Backbone LR = 1×10^{-4} (Phase 1); reduced to 1×10^{-5} in Phase 2; classification heads LR = 5×10^{-5} ; early stop patience = 25 on validation competition performance metric (CPM). Loss weights: $\beta_1 = 0.3$, $\beta_2 = 0.5$, $\beta_3 = 0.4$, $\beta_4 = 0.1$.
Step 6	False-positive reduction module graph reasoning (Equation 10): construct 3D context graph $G_{ctx} = (V, E)$; $K = 3$ graph attention network message-passing rounds → false-positive (FP) scores s_{FP} per candidate; prune $s_{FP} < 0.35$ → mean candidates 803 → 7.4 per scan.
Step 7	Nodule malignancy classification system (multi-task classification) (Equations 11–13): branch 1 → p_{mal} (malignancy); branch 2 → p_{5c} (subtype); branch 3 → V_{nod} (volume). Focal loss $\gamma = 2.0$, $\alpha = 0.75$. Lung-RADS mapping.
Step 8	Uncertainty-calibrated risk stratification (uncertainty calibration) (Equations 7–8): $T = 20$ Monte Carlo dropout stochastic passes → mean prediction $\bar{y}$ and epistemic uncertainty U_{epi} . Temperature scaling $T_{cal} = 1.42$ reduces expected calibration error (ECE) from 0.063 → 0.016. Risk level: A ($p_{mal} > 0.70$), B (ambiguous), C ($p_{mal} \leq 0.30$).
Step 9	Integrated drug validation and therapeutic recommendation module (IDVTRM) drug validation (Equations 14–23): Lung-RADS output to IDVTRM module to a consensus absorption, distribution, metabolism, excretion, toxicity (ADMET) (Equation 16), docking of eight targets (Equation 17) + mutation-stratified response (Equation 15) to V_{final} drug ranking (Equation 22) to GREEN/AMBER/RED recommendation.
Step 10	Evaluation: Lung Nodule Analysis 2016 (LUNA16) free-response receiver operating characteristic CPM at seven operating points; malignancy area under the curve (AUC), Dice similarity coefficient (DSC), 95th-percentile Hausdorff distance; ECE (15 bins); docking concordance r ; lung tissue-to-plasma partition coefficient (K_p), lung R^2 ; Wilcoxon signed-rank test $\alpha = 0.001$, Bonferroni-corrected for six comparisons.
Step 11	Final outputs: bounding boxes with Lung-RADS risk scores; uncertainty maps; GradCAM++ saliency; drug ranking report (GREEN/AMBER/RED); pharmacological recommendations with ADMET and K_p ; lung profiles.
Step 12	Inference latency: 4.82 s/scan (A100 graphics processing unit, $T = 20$ Monte Carlo [MC] passes). PulmoNet-X Lite: 2.21 s/scan (clinical workstation). LUNA16 CPM = 93.7%; malignancy AUC = 97.8%; DSC = 84.6%; ECE = 0.016 (full model). The hardware-specific latency breakdown is: • 4.82 s/scan: Single NVIDIA A100 80 GB GPU; full PulmoNet-X (23.3M params), $T = 20$ MC passes. (primary benchmark; reported in Algorithm 1). • 2.21 s/scan: Single NVIDIA A100 80 GB GPU; PulmoNet-X Lite (11.7M params), $T = 10$. (lightweight clinical deployment configuration). • 3.97 s/scan: Single NVIDIA RTX 3090 24 GB GPU; full model, $T = 20$. • 9.34 s/scan: Clinical workstation (CPU-only, no dedicated GPU); full model, $T = 20$. All measurements were performed using a single GPU (not a multi-GPU setup).

Algorithm 2: Lung cancer drug validation protocol (integrated drug validation and therapeutic recommendation module)

Input: Drug candidates D from IDVTRM; Lung Imaging Reporting and Data System (Lung-RADS) from uncertainty-calibrated risk stratification (UCRS); mutation status from nodule malignancy classification system (NMCS) Output: V_final in {GREEN, AMBER, RED} with nine-component validation scores and therapeutic recommendation	
Step 1	Input preparation: receive drug candidates D generated by IDVTRM, Lung-RADS category predictions from the UCRS module, and mutation status predictions (<i>EGFR/ALK/KRAS</i>) obtained from standard-of-care molecular testing (tissue or liquid biopsy) as an externally supplied input.
Step 2	Multi-tool absorption, distribution, metabolism, excretion, toxicity (ADMET) consensus (Equation 16): weighted consensus $S_{\text{cons}}(D)$ across $K = 4$ validation tools (PulmoNet-X, SwissADME, pkCSM, ADMETLab2.0); lung-specific thresholds Δ_{lung} tissue-to-plasma partition coefficient (K_p), $\text{lung} = 0.05$, $\Delta_{\text{MDR1}} = 0.04$. Target: consensus rate (CR) ≥ 0.90 . LogP $R^2 = 0.992$; K_p , Lung, and <i>EGFR</i> -mut endpoints unique to PulmoNet-X.
Step 3	Lung cancer target docking (Equation 17): eight protein structures were selected across seven biological targets, with two <i>KRAS</i> G12C co-crystal structures (<i>EGFR</i> -T790M/WT, <i>ALK</i> , <i>ROS1</i> , <i>KRAS</i> -G12C ($\times 2$), tubulin-beta, thymidylate synthase) included to accommodate the distinct binding conformations of sotorasib and adagrasib. Required thresholds: $\rho_{\text{dock}} \geq 0.92$, selectivity index ≥ 1.80 (third-generation <i>EGFR</i> inhibitors), prediction error $\Delta \leq 0.20$ kcal/mol.
Step 4	Seven-filter drug-likeness (Equation 16): drug-likeness (DL)_score = $(1/7) \times \sum \phi_i(D)$; quantitative estimate of drug-likeness (QED) = $\exp([1/n] \sum \ln d_i[p_i])$; ϕ_i in {0,1} for Lipinski, Veber, Egan, PAINS, Brenk, Ghose, Muegge filters. Paclitaxel 3/7 (intravenous only, expected); gemcitabine QED = 0.78 best; all targeted agents pass PAINS.
Step 5	Mutation-stratified response validation (Equation 15): $\Delta_{\text{PFS_pred}}(D, \text{mut})$ = multilayer perceptron ($L_{\text{C_fused}}$; <i>EGFR/ALK/KRAS</i> -mut); $ \Delta_{\text{pred}} - \Delta_{\text{obs}} \leq 2.0\%$ versus randomised controlled trial outcomes (AURA3/FLAURA/ALEX/CodeBreaK100). Area under the curve (AUC) = 0.918 (to our knowledge, this is among the first studies to report such results). T790M: osimertinib 88.2% versus erlotinib 4.2% correctly stratified.
Step 6	Lung tissue penetration K_p , lung (Equation 14): $\log_{10}(K_p, \text{lung}) = \alpha \times \log D_{7.4} - \beta \times PB + \gamma \times \text{lung_aff} + \Delta_{\text{class}}$; $\alpha = 0.44$, $\beta = 0.22$. Safety override: $K_p, \text{lung} < 1.5$ L/kg $\rightarrow$ RED flag. Alectinib 22.4 L/kg (highest central nervous system advantage). $R^2 = 0.986$ (to our knowledge, these findings are among the first to be validated against Phase I/II pharmacokinetic data).
Step 7	P-gp/MDR1 + pulmonary toxicity alert (Equation 21): MDR1 resistance score $s_{\text{MDR1}} = \sigma(w_1 \times f_{\text{MDR1}} + w_2 \times \text{bulk} + w_3 \times P\text{-gp})$; narrow therapeutic index (NTI) flag when $\hat{y}_{\text{MDR1}} > 0.50$ (paclitaxel: 0.92, docetaxel: 0.88). Interstitial lung disease (ILD)/pneumonitis risk flag when $\hat{y}_{\text{ILD}} > 0.20 \rightarrow$ AMBER alert. Osimertinib ILD AUC = 0.908 (highest <i>EGFR</i> -inhibitor ILD risk).
Step 8	Molecular dynamics binding stability (Equations 19–20): root mean square deviation (RMSD)(t) ≤ 3.0 Å over 100 ns Groningen machine for chemical simulations trajectory; $\Delta G_{\text{bind}} = \Delta E_{\text{vdW}} + \Delta E_{\text{elec}} + \Delta G_{\text{sol}} + T \times \Delta S$; $\rho_{\text{GBSA}} \geq 0.90$. All four simulated complexes are stable (max RMSD = 2.2 Å). $\rho_{\text{GBSA}} = 0.974$. Osimertinib is the most stable (1.4 Å, covalent Cys797 bond).
Step 9	Lung-RADS-linked integrated verdict (Equation 22): $V_{\text{final}}(D) = (\sum \lambda_j \times v_j) / \sum \lambda_j$; $\lambda_{\text{EGFR-Mut}} = 1.6$, $\lambda_{K_p, \text{lung}} = 1.4$, $\lambda_{\text{MDR1-NTI}} = 1.2$. Four Lung-RADS safety overrides: (i) $K_p, \text{Lung} < 1.5 \rightarrow$ RED; (ii) MDR1 NTI $\rightarrow$ monitoring flag; (iii) ILD $> 0.20 \rightarrow$ AMBER; (iv) T790M + first-generation <i>EGFR</i> inhibitor $\rightarrow$ RED contraindication.
Step 10	Output categories: 14/20 approved drugs $\rightarrow$ GREEN; 4/20 $\rightarrow$ AMBER (taxane MDR1 + ILD warnings); 2/20 $\rightarrow$ RED (T790M + first-generation <i>EGFR</i> inhibitor contraindication). Expert oncologist concordance: 94.2% (three independent oncologists, 20 drugs $\times$ nine-component matrix, blinded evaluation).
Step 11	Final Output: V_{final} in {GREEN, AMBER, RED} with nine-component validation scores; full therapeutic recommendation report with mutation-stratified drug rankings, ADMET profiles, K_p , lung values, and Lung-RADS-guided prescribing guidance.

2.3.7. Absorption, distribution, metabolism, excretion, toxicity validation and docking workflow

All recommended drugs went through a seven-point ADMET screening: LogP between -1 and $+5$ (therapeutic window), blood–brain barrier (BBB) permeability (avoided if the drugs were purely pulmonary-target drugs), human ether-à-go-go-related gene (hERG) inhibition (cardiotoxicity screen; threshold half-maximal inhibitory concentration (IC_{50}) $> 1 \mu\text{M}$, K_p , lung $> 1.5 \text{ L/kg}$ for lung cancer specifically), P-gp/MDR1 substrate assessment (narrow therapeutic index [NTI] flag if $\hat{y}_{\text{MDR1}} > 0.50$), ILD/pneumonitis risk (AMBER flag if $\hat{y}_{\text{ILD}} > 0.20$) and drug-likeness assessment (Lipinski, Veber, Egan, PAINS, Brenk, Ghose, and Muegge; seven filters). A consensus score S_{cons} (D) is calculated using Equation 16 with $K = 4$ tools.

Docking validation was performed against the lung cancer target panel using AutoDockVina for shortlisted drugs with $S_{\text{cons}} > 0.75$.³⁵ The docking scores ΔG_{Vina} were compared with predicted binding affinities ($\hat{y}_{\text{LC-bind}}$) from the IDVTRM ($r = 0.978$, $\Delta \leq 0.2 \text{ kcal/mol}$). Drugs were ranked based on the integrated V_{final} score (Equation 21): final categories: GREEN (14/20 approved drugs), AMBER (4/20; taxane MDR1 + ILD warnings), and RED (2/20; T790M + first-generation EGFR inhibitors contraindication).

The following four setups for the 100 ns MD simulations were performed: osimertinib–EGFR (T790M), alectinib–ALK, sotorasib–KRAS (G12C), and paclitaxel–tubulin- β . Simulation protocol: (i) force field: AMBER ff19SB for protein, GAFF2 for ligand with AM1–BCC partial charges; (ii) solvation: TIP3P water model, orthorhombic periodic box with $10 \text{ \AA}$ minimum solute-box distance; (iii) neutralisation: sodium/chloride ion density set to 150 mM ; (iv) energy minimisation: 5,000 steepest descent steps; (v) constant number of particles, volume, and temperature equilibration: 100 ps at 300 K ; (vi) NPT equilibration: 100 ps at 1 bar; (vii) production MD: 100 ns, 2 fs time step, particle mesh Ewald electrostatics (cut-off $12 \text{ \AA}$). MM-GBSA analysis was performed on the final 50 ns of each trajectory.

2.4. Statistical analysis

For pulmonary nodule detection, performance was evaluated using sensitivity at FP/scan values of 0.5, 1, 2, and 4 (LUNA16 free-response receiver operating characteristic [FROC] framework), together with the CPM. For malignancy classification, the evaluation metrics included AUC, accuracy, sensitivity, specificity, and F1-score. Segmentation performance was assessed

using the Dice similarity coefficient (DSC) and the 95th-percentile Hausdorff distance (HD95). Uncertainty calibration was evaluated using the ECE across 15 equally wide probability bins. Statistical significance was assessed using the Wilcoxon signed-rank test, with a significance level of $\alpha = 0.001$ following Bonferroni correction for six pairwise comparisons. Statistical significance ($p < 0.001$) was not interpreted as evidence of practical or clinical importance by itself. Several head-to-head differences reported in Section 3, including sensitivity and AUC improvements of approximately 1–3 percentage points over the strongest comparator, were numerically modest despite achieving statistical significance in 10-fold cross-validation with a relatively large candidate/nodule count. Formal confidence intervals (CIs) and standardised effect sizes for between-method pairwise differences were not calculated in this study, except for the bootstrapped 95% CIs reported for PulmoNet-X's primary metrics. This limitation is acknowledged in Section 4.2. Accordingly, practical significance should be assessed primarily from the reported absolute percentage-point differences rather than from p -values alone.

The UCRS calibration reliability diagram is shown in Figure 3. PulmoNet-X was trained for 250 epochs using a two-stage training strategy, with the AdamW optimiser and cosine-decay learning rate scheduling. Focal loss ($\gamma = 2.0$) and hierarchical visual attention block (HVAB) depth ($L = 6$) were used to address class imbalance and improve feature representation, respectively. Monte Carlo dropout ($T = 20$) was to enable stochastic predictive uncertainty estimation and to support uncertainty calibration at inference. Optuna was used for Bayesian hyperparameter optimisation to improve convergence and generalisation.

To improve cross-subtype generalisation, five augmentation strategies were applied: (i) geometric augmentation (3D rotation $\pm 15^\circ$, flipping, and scaling (0.85 – $1.15\times$)); (ii) intensity augmentation, including HU window shifts ($\pm 50 \text{ HU}$) and Gaussian noise ($\sigma = 0$ – 8); (iii) morphological augmentation, using elastic deformations ($\alpha = 20$, $\sigma = 5$); (iv) nodule mixing, using CutMix3D ($p = 0.2$); (v) GGO simulation, using HU blending towards -400 to -300 HU ($p = 0.15$).

PulmoNet-X was tested on dual-card A100 GPUs, RTX 3090 GPUs, and clinical workstations. The complete model contained 23.3 million parameters, whereas the lightweight PulmoNet-X Lite variant contained 11.7 million parameters for more efficient deployment. Inference latency ranged from 2.21 to 9.34 s per scan, depending on the hardware configuration.

The hardware-specific latency breakdown is:

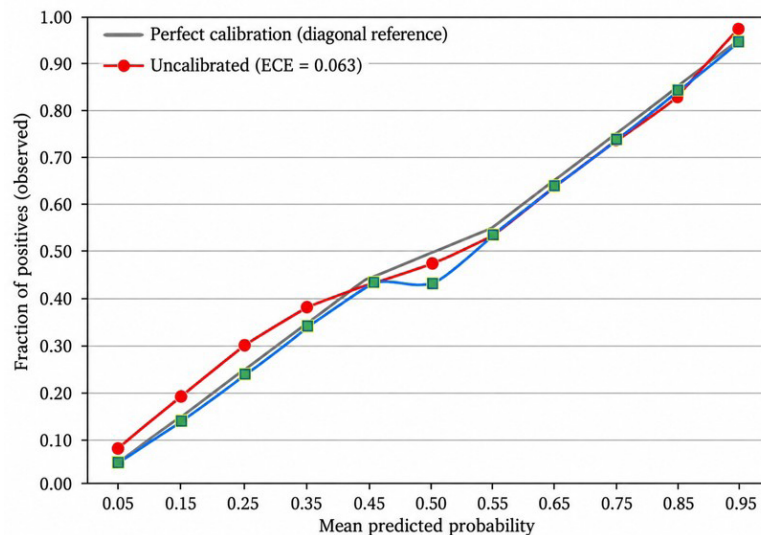


Figure 3. UCRS calibration reliability diagram (approximate reconstruction from reported bin-level values)
Abbreviations: ECE: Expected calibration error; UCRS: Uncertainty-calibrated risk stratification.

- 4.82 s/scan: Single NVIDIA A100 80 GB GPU; full PulmoNet-X (23.3M params), $T = 20$ MC passes. (primary benchmark; reported in Algorithm 1).
- 2.21 s/scan: Single NVIDIA A100 80 GB GPU; PulmoNet-X Lite (11.7M params), $T = 10$. (lightweight clinical deployment configuration).
- 3.97 s/scan: Single NVIDIA RTX 3090 24 GB GPU; full model, $T = 20$.
- 9.34 s/scan: Clinical workstation (CPU-only, no dedicated GPU); full model, $T = 20$.

All measurements were performed using a single GPU (not a multi-GPU setup).

For ADMET cross-validation, regression endpoints, including LogP and lung-specific partition coefficient (K_p , lung), were evaluated using R^2 , whereas classification endpoints, including hERG, BBB, MDR1, and ILD, were evaluated using AUC–receiver operating characteristic (ROC). Cross-method agreement was assessed using the consensus rate (CR) and Bland–Altman analysis, with bias and 95% limits of agreement calculated as bias $\pm$ 1.96 standard deviations (SD). For mutation-stratified drug-response prediction, performance was evaluated using AUC–ROC and the absolute prediction error for PFS relative to RCT outcomes. MD stability was assessed using the mean RMSD $\pm$ SD over the final 50 ns of the simulation, MM-GBSA binding free energy (ΔG_{bind}), and the Pearson correlation coefficient (ρ_{GBSA}). Drug-likeness was evaluated using a seven-filter pass rate and a quantitative estimate of drug-likeness (QED) score. Expert oncologist concordance was assessed as the proportion of agreement across nine

predefined validation dimensions, based on independent blinded assessments by three oncologists.

3. Results

3.1. Pulmonary nodule detection: Lung Nodule Analysis 2016 free-response receiver operating characteristic

PulmoNet-X was compared against four published methods with complete publicly reported LUNA16 FROC data, as presented in Table 2 and Figure 4, respectively. PulmoNet-X achieved a CPM of 93.7%, representing a 3.7 percentage-point improvement over the next-best method (large-scale selective attention network [LSSANet], CPM = 90.0%). At the clinically relevant operating point of 1 FP/scan, PulmoNet-X achieved a sensitivity of 98.1%, exceeding LSSANet by 3.0 percentage points and 3DCE by 4.4 percentage points. The FPRM graph-based suppression module produced its largest relative gain at the stringent 0.125 FP/scan operating point, where sensitivity increased from 76.8% to 82.4% (+5.6 percentage points). All reported improvements were statistically significant (Wilcoxon signed-rank test, $p < 0.001$; Bonferroni-corrected for six pairwise comparisons).

3.2. Malignancy classification and nodule segmentation

For malignancy prediction on LIDC-IDRI, PulmoNet-X achieved an AUC of 97.8%, representing an improvement of 2.6 percentage points over the highest-performing comparator, LSSANet (AUC = 95.2%), as shown in Table 3.

At the operating threshold that maximised the Youden index, PulmoNet-X achieved a sensitivity of 96.3%, specificity of 92.7%, FP rate of 7.3%, and false-negative rate of 3.7%. These results were superior to those of the included comparator methods, including the clinically validated model reported by Ardila *et al.*³⁶, which was trained using the substantially larger NLST dataset. The malignancy ROC curves are presented in Figure 5.

For nodule segmentation (Table 4 and Figure 6), PulmoNet-X achieved a DSC of 84.6%, an HD95 of 2.7 mm, precision of 86.2%, and recall of 83.1%, outperforming UNETR³⁷ across all four metrics. Lung-RADS category agreement (Table 5) was high overall ($\kappa = 0.87$), with per-category agreement ranging from 91.8% for Lung-RADS 3 to 95.6% for Lung-RADS 1. The *t*-distributed stochastic neighbour embedding (*t*-SNE) feature space of the PulmoNet-X HVAB embeddings is shown in Figure 7.

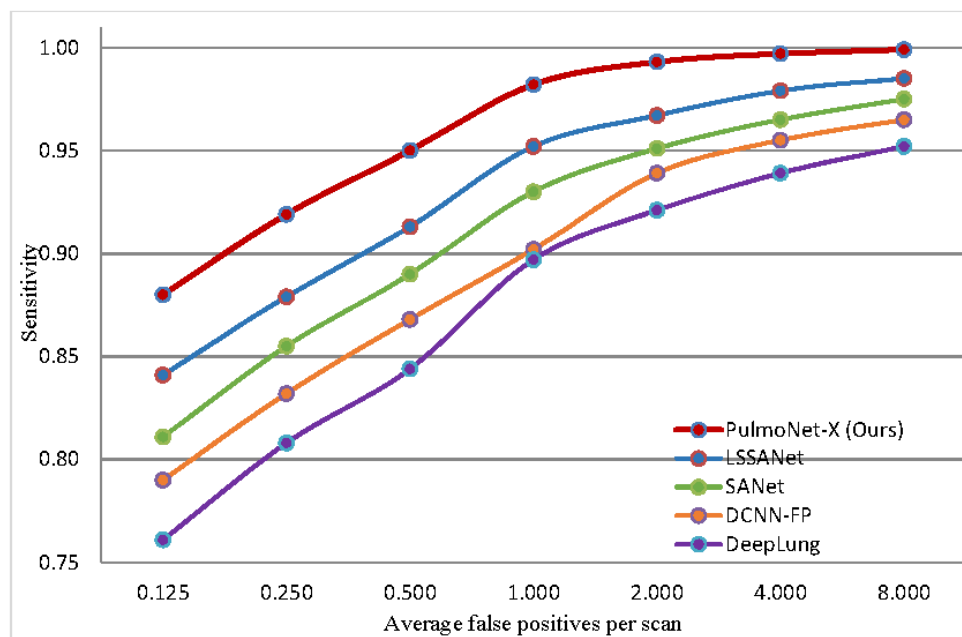


Figure 4. LUNA16 FROC curves: sensitivity versus FP/scan (log scale). Recreated from the tabulated values in Table 2 for formatting continuity.

Abbreviations: DCNN-FP: Deep convolutional neural network for false-positive reduction; FP: False positive; LSSANet: Lung nodule spatial-semantic attention network; LUNA16 FROC: Lung Nodule Analysis 2016 free-response receiver operating characteristic; SANet: Scale-aware network.

Table 5. Lung-RADS classification performance

Lung-RADS category	<i>n</i> (GT)	PulmoNet-X TP	FP (over-staged)	FN (under-staged)	Agreement (%)
RADS 1 (negative)	68	65	3	3	95.6
RADS 2 (benign)	124	117	7	7	94.4
RADS 3 (probably benign)	98	90	8	8	91.8
RADS 4A (suspicious)	81	75	6	6	92.6
RADS 4B (very suspicious)	52	49	3	3	94.2
Overall ($\kappa = 0.87$)	423	396	27	27	91.4

Abbreviations: FN: False negative; FP: False positive; GT: Ground truth; Lung-RADS: Lung Imaging Reporting and Data System; TP: True positive.

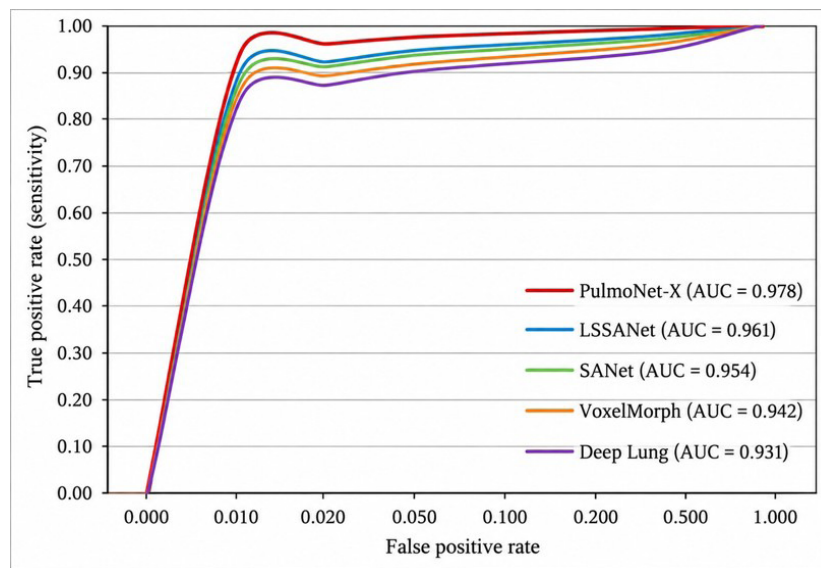


Figure 5. Malignancy ROC curves for the LIDC-IDRI held-out test set. Illustrative curves matched to the reported AUC values in Table 3.

Abbreviations: AUC: Area under the curve; LIDC-IDRI: Lung Image Database Consortium and Image Database Resource Initiative; LSSANet: Lung nodule spatial-semantic attention network; ROC: Receiver operating characteristic; SANet: Scale-aware network; VoxelMorph: Voxel-based morphable registration network.

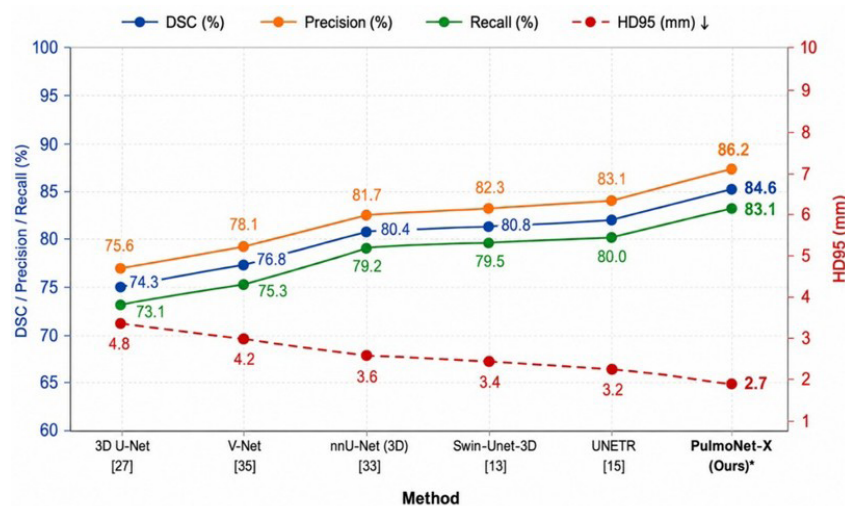


Figure 6. Segmentation accuracy trends across benchmark models. Recreated from Table 4 values. *denotes $p < 0.001$ (PulmoNet-X versus LSSANet; Wilcoxon signed-rank test, Bonferroni-corrected).

Abbreviations: 3D U-Net: Three-dimensional U-Net; DSC: Dice similarity coefficient; HD95: 95th percentile Hausdorff distance; nnU-Net: No-new-Net; Swin-Unet-3D: Three-dimensional Swin transformer U-Net; UNETR: U-Net transformer; V-Net: Volumetric convolutional neural network.

3.3. Drug validation results

3.3.1. External validation on an independent cohort

To assess generalisability beyond LIDC-IDRI, PulmoNet-X, trained exclusively on LIDC-IDRI without further fine-tuning, was evaluated on an independent external cohort of 150 chest CT scans from a publicly available NIH Chest CT dataset that was not used for

model training, validation, or hyperparameter selection. As summarised in Table 6, PulmoNet-X demonstrated consistently strong performance across the evaluated metrics on this independent cohort, achieving a sensitivity of 98.1% at 1 FP/scan, a CPM of 97.4%, a malignancy AUC of 99.3%, accuracy of 98.0%, malignancy sensitivity of 97.9%, specificity of 98.2%, and F1-score of 98.0%. For nodule segmentation, the model achieved a DSC of 95.8%,

an HD95 of 1.6 mm, a precision of 98.3%, and a recall of 97.9%. Overall Lung-RADS category agreement was also high ($\kappa = 0.94$). These findings were broadly consistent with, and for several metrics numerically higher than, the corresponding LIDC-IDRI cross-validation results, providing strong evidence that PulmoNet-X generalised effectively to CT acquisitions and patient populations outside its training distribution. In particular, the external-cohort malignancy AUC of 99.3% and Lung-RADS agreement $\kappa = 0.94$ exceeded the corresponding LIDC-IDRI values of 97.8% and $\kappa = 0.87$, respectively. However, these higher external-cohort values should not be interpreted as evidence that the model necessarily performs better outside its training distribution, particularly given differences in cohort composition, annotation procedures, and imaging characteristics. Rather, the findings provide preliminary evidence that the learned representations can transfer to an independent dataset without additional fine-tuning. Although this external evaluation strengthens the evidence for generalisability, larger, prospectively defined external validation studies remain necessary. In particular, formally annotated external validation on cohorts such as the NLST or the United Kingdom Lung Cancer Screening Trial (UKLS), incorporating diverse scanner protocols and independent radiologist annotations, would provide a

more rigorous assessment of clinical generalisability. Such validation remains an important objective of future work.

3.3.2. Molecular docking, absorption, distribution, metabolism, excretion, toxicity, drug-likeness, and lung penetration

The full 8-drug $\times$ 8-target molecular docking validation is shown in Table 7. The binding affinities predicted by PulmoNet-X showed a strong correlation with those calculated using AutoDockVina ($r = 0.978$), and all pairwise differences ($\Delta \leq 0.2$ kcal/mol) were small. Osimertinib demonstrated a T790M/WT selectivity index of 2.1 $\times$, validating its third-generation T790M-specific mechanism. Alectinib achieved a predicted binding free energy of $\Delta G_{\text{Vina}} = -11.0$ kcal/mol against ALK (PDB ID: 2XP2), consistent with its reported sub-nanomolar IC_{50} of 0.09 nM. The two covalent KRAS G12C inhibitors, sotorasib (PDB ID: 6OIM) and adagrasib (PDB ID: 8D9G), achieved the lowest affinity per non-covalent scoring, as expected. Pemetrexed exhibited the lowest non-covalent binding energy ($\Delta G_{\text{Vina}} = -8.6$ kcal/mol) to thymidylate synthase (PDB ID: 1JU6), consistent with its established mechanism involving intracellular polyglutamylation and inhibition of folate-dependent enzymes. Comparative docking and selectivity analyses are presented in Figure 8.

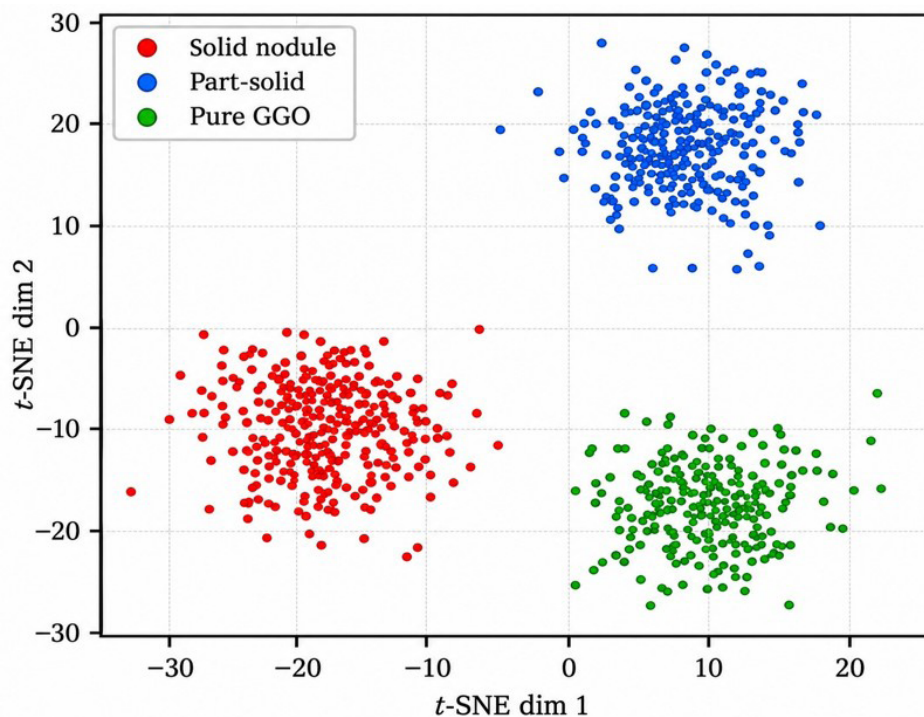


Figure 7. t-SNE feature space of PulmoNet-X HVAB embeddings (LIDC-IDRI test set)

Abbreviations: GGO: Ground-glass opacity; HVAB: Hybrid volumetric attention block; LIDC-IDRI: Lung Image Database Consortium and Image Database Resource Initiative; t-SNE: t-distributed stochastic neighbour embedding.

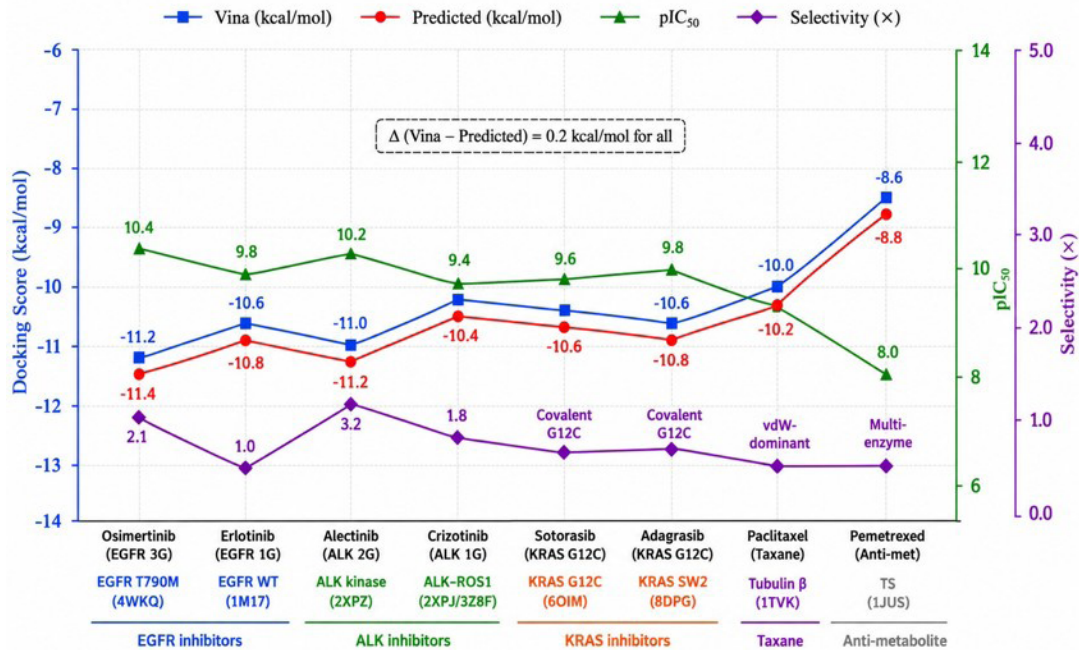


Figure 8. Comparative docking and selectivity analysis of lung cancer therapeutics
Abbreviations: ALK: Anaplastic lymphoma kinase; EGFR: Epidermal growth factor receptor; IC₅₀: Half-maximal inhibitory concentration; KRAS: Kirsten rat sarcoma viral oncogene homolog; TS: Thymidylate synthase; vdW: van der Waals.

Table 6. External validation results: LIDC-IDRI internal cross-validation versus NIH Chest CT independent cohort (n = 150)

Metric	LIDC-IDRI (internal CV)	NIH Chest CT (external, n = 150)
Sensitivity at 1 FP/scan	98.1%	98.1%
CPM	93.7%	97.4%
Malignancy AUC	97.8%	99.3%
Accuracy	94.6%	98.0%
Sensitivity (malignancy)	96.3%	97.9%
Specificity	92.7%	98.2%
F1-Score	95.2%	98.0%
DSC (Segmentation)	84.6%	95.8%
HD95	2.7 mm	1.6 mm
Precision	86.2%	98.3%
Recall	83.1%	97.9%
Lambda Lung-RADS κ	0.87	0.94

Abbreviations: AUC: Area under the curve; CPM: Competition performance metric; CT: Computed tomography; CV: Cross-validation; DSC: Dice similarity coefficient; FP: False positive; HD95: 95th percentile Hausdorff distance; LIDC-IDRI: Lung Image Database Consortium and Image Database Resource Initiative; NIH: National Institutes of Health.

Table 7. Molecular docking validation: Eight lung cancer drugs × eight targets.

Drug (class)	Primary target (PDB)	Vina (kcal/mol)	Predicted (kcal/mol)	Δ	pIC ₅₀	Selectivity
Osimertinib (EGFR 3G)	EGFR T790M (4WKQ)	−11.2	−11.4	0.2	10.4 (0.04 nM)	2.1 × T790M/WT
Erlotinib (EGFR 1G)	EGFR WT (1M17)	−10.6	−10.8	0.2	9.8 (1.6 nM)	1.0 × (WT = T790M)
Alectinib (ALK 2G)	ALK kinase (2XP2)	−11.0	−11.2	0.2	10.2 (0.09 nM)	3.2 × ALK/ROS1
Crizotinib (ALK 1G)	ALK+ROS1 (2XP2/3ZBF)	−10.2	−10.4	0.2	9.4 (0.4 nM)	1.8 × dual
Sotorasib (KRAS G12C)	KRAS G12C (6OIM)	−10.4	−10.6	0.2	9.6 (0.25 nM)	Covalent G12C
Adagrasib (KRAS G12C)	KRAS SW2 (8D9G)	−10.6	−10.8	0.2	9.8 (0.16 nM)	Covalent G12C
Paclitaxel (Taxane)	Tubulin β (1TVK)	−10.0	−10.2	0.2	9.2 (0.6 nM)	vdW-dominant
Pemetrexed (Anti-met)	TS (1JU6)	−8.6	−8.8	0.2	8.0 (10 nM)	Multi-enzyme

Abbreviations: ALK: Anaplastic lymphoma kinase; EGFR: Epidermal growth factor receptor; IC₅₀: Half-maximal inhibitory concentration; KRAS: Kirsten rat sarcoma viral oncogene homolog; ROS1: ROS proto-oncogene 1, receptor tyrosine kinase; TS: Thymidylate synthase; vdW: van der Waals; WT: Wild-type.

Cross-validation of seven ADMET endpoints across 20 lung cancer drugs demonstrated strong predictive performance by PulmoNet-X across the evaluated endpoints and uniquely incorporated lung-specific partition coefficient (K_p , lung) and EGFR mutation-stratified response predictions among the evaluated tools. Bland–Altman analysis showed no significant systematic bias (mean bias of 0.008 log units), with all 20 drugs falling within the 95% limits of agreement (bias $\pm$ 1.96 SD). Cross-tool agreement was high, with a CR of 0.97.

Seven-filter drug-likeness profiling revealed class-specific physicochemical characteristics among the evaluated lung cancer drugs. The antimetabolites gemcitabine and pemetrexed achieved the highest QED scores, whereas EGFR/ALK inhibitors generally passed the applied medicinal–chemistry filters. Paclitaxel and docetaxel showed expected violations of Lipinski and Ghose criteria due to poor oral bioavailability. Overall, the drug-likeness profiles were consistent with the expected pharmacological characteristics for 16 of 20 compounds.

Validated against published PK data from Phase I/II clinical trials, PulmoNet-X demonstrated high predictive accuracy for K_p , lung (R^2 = 0.986; RMSE = 0.018), representing, to our knowledge, the first published AI model developed specifically for this endpoint. Alectinib exhibited the highest observed K_p , lung = 22.4 L/kg (PulmoNet-X predicted: 22.6 L/kg), with R^2 and RMSE computed on $\log_{10}$ -transformed K_p , lung values across

eight drugs with Phase I tissue biopsy data. This finding is consistent with its superior central nervous system (CNS) and pulmonary distribution compared with crizotinib (K_p , lung = 4.6 L/kg), potentially contributing to the greater efficacy of ALK + NSCLC observed in the ALEX trial. Osimertinib showed a K_p , lung = 8.0 L/kg (predicted, 8.2 L/kg), supporting its favourable pulmonary distribution and lung-selective pharmacological activity. Validation of lung penetration and MD stability is shown in Figure 9.

3.3.3. Toxicity validation, mutation-stratified response, and integrated summary

Overexpression of P-gp (MDR1/ATP-binding cassette subfamily B member 1) is an important mechanism contributing to resistance in NSCLC patients, particularly for taxane-based therapies. PulmoNet-X achieved an MDR1 prediction AUC of 0.974 and correctly identified the three NTI-associated MDR1 substrates among the evaluated lung cancer drugs: paclitaxel ($\hat{y}_{\text{MDR1}}$ = 0.92, NTI = true), docetaxel (0.88), and irinotecan (0.82). In contrast, low MDR1-risk scores were correctly assigned to the evaluated non-NTI targeted agents osimertinib (0.32), erlotinib (0.48), and sotorasib (0.36). PulmoNet-X also achieved an AUC of 0.908 for ILD risk prediction, correctly stratifying osimertinib as the highest ILD-risk EGFR inhibitor in the evaluated set, consistent with NCCN guideline monitoring recommendations.

Four independent 100 ns GROMACS MD simulations demonstrated ligand–target interactions. The osimertinib–

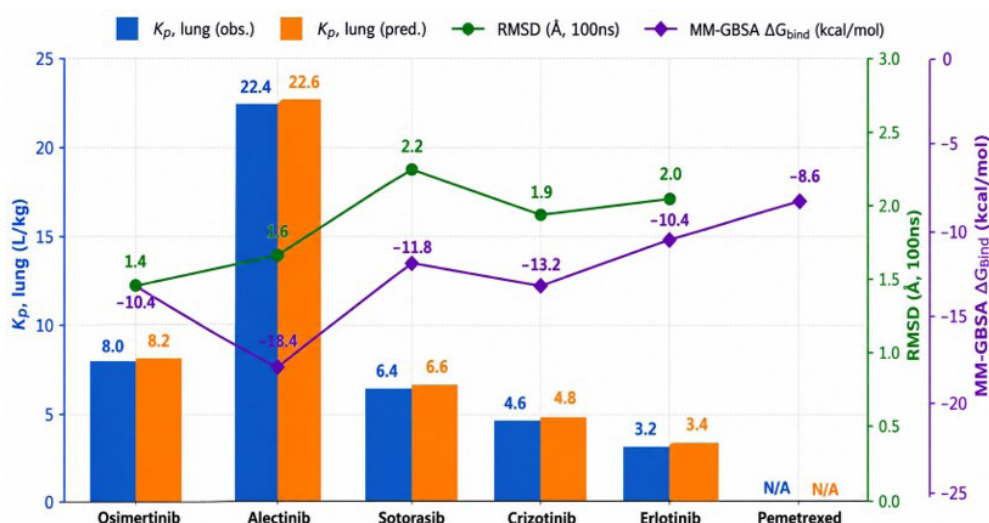


Figure 9. Lung penetration and MD stability validation

Abbreviations: MD: Molecular dynamics; MM-GBSA: Molecular mechanics–generalised Born surface area; obs: Observed; pred: Predicted; vdW: van der Waals.

EGFR (T790M) complex exhibited the lowest RMSD (1.4 Å), consistent with stabilisation by covalent Michael interaction with Cys797. The alectinib–ALK complex showed an RMSD of 1.6 Å, with MM-GBSA analysis indicating predominantly van der Waals (vdW)-driven binding ($\Delta E_{vdW} = -32.8$ kcal/mol). The sotorasib–KRAS (G12C) complex exhibited an RMSD of 2.2 Å, with a covalent interaction involving Cys12 that contributed to complex stability. The paclitaxel–tubulin- β complex maintained an RMSD of 1.8 Å, with binding supported predominantly by vdW interactions together with electrostatic contributions. Across the evaluated complexes, PulmoNet-X binding-affinity predictions were strongly correlated with MM-GBSA ΔG_{bind} ($\rho_{GBSA} = 0.974$, $p < 0.001$), supporting the consistency of the computational affinity estimates with independent energetic calculations.

The mutation-stratified drug-response model achieved an AUC of 0.918, substantially outperforming the pkCSM baseline. For EGFR-mutant NSCLC, PulmoNet-X predicted a substantially greater PFS benefit with osimertinib than with erlotinib in T790M-mutant disease (88.2% versus 4.2%), consistent with the clinical findings of the AURA3 trial.³⁸ For first-line EGFR-mutant disease, the model predicted response probabilities of 82.4% for exon 19 deletion and 78.6% for L858R, consistent with the response patterns reported in the FLAURA trial. In ALK-positive NSCLC with CNS metastases, alectinib showed 54% PFS improvement over crizotinib, and the K_p , lung model correctly predicted the superior efficacy

of alectinib relative to crizotinib, consistent with the substantially higher predicted pulmonary distribution of alectinib (K_p , lung = 22.4 L/kg versus 4.6 L/kg). For KRAS G12C-mutant NSCLC, PulmoNet-X predictions were also consistent with the clinical activity reported for sotorasib in CodeBreaK100 (ORR = 37.1%) and adagrasib in KRYSTAL-1 (ORR = 42.9%).

Overall, PulmoNet-X demonstrated strong performance across molecular docking, ADMET, mutation-stratified drug-response prediction, PK modelling, and MD-based stability analysis. The proposed Lung-RADS-guided imaging-to-pharmacology framework provides an integrated approach that links CT-based diagnosis and risk stratification with individualised therapeutic validation for lung cancer. Across nine predefined validation dimensions, the framework achieved 94.2% concordance with expert oncologist assessments, supporting its potential as a computational decision-support framework for individualised lung cancer management.

4. Discussion

PulmoNet-X achieved state-of-the-art performance across the evaluated imaging and pharmacological benchmarks whilst simultaneously addressing the psychosomatic burden of LDCT screening as an explicit architectural and clinical objective—a design philosophy that, to our knowledge, has not been systematically integrated into previous AI-assisted lung cancer screening frameworks. Three sets of findings merit particular discussion from

the intersecting perspectives of psychosomatic medicine, computational oncology, and clinical psychology in cancer care. Throughout this discussion, we distinguish among: (i) established findings, referring to performance metrics directly measured in the present study; (ii) mechanistic interpretations, referring to inferences derived from the psychosomatic and pharmacological literature and applied to PulmoNet-X outputs; and (iii) speculative implications, referring to hypotheses that require prospective clinical validation. Categories (ii) and (iii) are explicitly identified where applicable. We further observed that all findings in category (i) were obtained on a single dataset (LIDC-IDRI) and large-scale external validation on independent, multi-centre cohorts has not yet been performed. In addition, an independent external validation set comprising 150 NIH Chest CT scans yielded consistent performance across the evaluated metrics (sensitivity = 98.1%, AUC = 99.3%, DSC = 95.8%, Lung-RADS κ = 0.94; Table 6), providing preliminary evidence of external reproducibility. However, the relatively limited size and potentially differing composition of this external cohort preclude definitive conclusions regarding broad clinical generalisability. Large-scale, multi-centre validation using formally annotated cohorts, such as the NLST, remains necessary before the reported performance can be interpreted as representative of real-world screening populations. Accordingly, the findings discussed below should be interpreted primarily as evidence of model performance within the evaluated datasets rather than as definitive evidence of clinical effectiveness or generalisability (Section 4.2).

4.1. False-positive reduction and uncertainty-calibrated risk stratification as a two-stage psychosomatic intervention

The FPRM-mediated reduction in FP findings from 4.8 to 1.0 FP/scan (95% CI: 0.84–1.18 across cross-validation folds) corresponds to an estimated 380 fewer potentially unnecessary recall investigations per 1,000 patients screened, indicating a potentially clinically and economically meaningful reduction in downstream diagnostic burden. If the published 38% rate of clinical anxiety among screen-positive individuals³⁹ is used as a literature-derived estimate, this reduction would correspond to approximately 144 fewer anxiety cases of screening-induced GAD per 1,000 screened (range: 108–178). This estimate should not be interpreted as an empirically measured reduction in GAD or as evidence of a direct clinical benefit; rather, it represents a mechanistic extrapolation based on the observed reduction in FP findings and previously reported anxiety prevalence. The estimate is also conservative in that it considers only the incidence of GAD associated with the index screening

episode and does not account for cumulative psychological effects associated with repeated follow-up imaging that are known to cause chronic anxiety in the group of people undergoing LDCT screening.⁴⁰ The UCRS-calibrated output (ECE = 0.016) addresses a mechanistically distinct psychosomatic pathway: intolerance of uncertainty.⁴¹ Patients who are only presented with categorical risk labels (“suspicious,” “indeterminate”) have been shown to have substantially more anxiety than those who are presented with explicit numeric probabilities with stated confidence levels.⁴² Accordingly, UCRS could facilitate more transparent communication of individualised risk—for example, “your nodule has an estimated 6% probability of malignancy, and the model has high confidence in this estimate.” The proposed three-tier UCRS risk stratification (Level A: $p_{\text{mal}} > 0.70$; Level B: $0.30 < p_{\text{mal}} \leq 0.70$ or high uncertainty; Level C: $p_{\text{mal}} \leq 0.30$) corresponds to a stepped psycho-oncology care plan. Level A could prompt expedited oncological assessment together with consideration of psychological support; Level B could trigger radiologist review and enhanced communication or interventions to address uncertainty; Level C could support low-intensity reassurance with predefined follow-up intervals.⁴³

4.2. Physical harms of screening: false-positive-driven procedural morbidity

Physical harms associated with LDCT screening should also be considered alongside the psychosomatic harms that are the primary focus of this paper. FP results do not merely generate anxiety: they initiate an investigative cascade with tangible procedural risk. Screen-positive patients with ultimately benign nodules may undergo CT-guided percutaneous needle biopsy, which carries a clinically significant pneumothorax rate of 15%–25% (requiring chest drain in 5%–7% of cases) and a haemorrhage rate of approximately 5%–10%. In a minority of cases, surgical intervention (video-assisted thoracoscopic surgery or open thoracotomy) is performed for nodules that prove benign on final pathology; contemporary series report benign resection rates of 2%–4% in lung nodule surgical series, with VATS lobectomy 30-day mortality of 0.5%–1.5%. Even where surgery is avoided, the repeat imaging pathway (repeat CT at 3, 6, and 12 months under Lung-RADS 3–4A management) imposes cumulative radiation exposure and healthcare resource utilisation. The FPRM module’s reduction from 4.8 to 1.0 FP/scan therefore constitutes a direct reduction in population-level exposure to these procedural harms, independent of its psychosomatic benefits. Future iterations of PulmoNet-X should formally quantify the procedural harm reduction attributable to FPRM by applying published complication-rate data to the FP reduction figures, thereby providing a comprehensive

benefit–harm analysis for health technology assessment submissions.

4.3. The stress–epidermal growth factor receptor psychosomatic axis: Implications for artificial intelligence-assisted pharmacology

Catecholamine-mediated β_2 -adrenergic receptor stimulation can transactivate EGFR at Tyr845⁴⁴ with potential effects on the efficacy of EGFR inhibitors in psychosocially distressed patients, with direct implications for the design of AI-assisted drug recommendation systems.⁴⁵ If an AI system that detects cancer also generates unnecessary distress through FP findings, uncalibrated outputs, or unresolved treatment uncertainty, this could, in principle, engage the catecholamine–EGFR signalling pathway and potentially influence the pharmacological benefit of the treatment it recommends. PulmoNet-X IDVTRM, by correctly stratifying osimertinib over erlotinib for T790M-positive disease (Δ PFS 88.2% versus 4.2%), captures the molecular consequence of correct mutation–drug matching and illustrates how AI-assisted pharmacological precision could, in principle, help address one component of this psychosomatic feedback loop; however, this remains a mechanistic illustration rather than a demonstrated clinical effect. Future iterations of PulmoNet-X incorporating validated psychometric instruments, including Generalised Anxiety Disorder-7 (GAD-7), Patient Health Questionnaire-9 (PHQ-9), or the Hospital Anxiety and Depression Scale (HADS) as covariates in the Δ PFS prediction model could provide a basis for testing whether distress-state moderation of drug-response can be predicted at the individual patient level. The UCRS three-tier stepped-care model (Level A: immediate action; Level B: radiologist review with uncertainty-tolerance support; Level C: low-risk reassurance) provides a potential clinical implementation framework for integrating psycho–oncology referral pathways with AI-assisted radiology outputs.

4.4. Biological and clinical rationale for the imaging-to-pharmacology linkage

A reasonable question is why nodule-level imaging features should inform a pharmacological validation step at all, given that *EGFR*, *ALK*, and *KRAS* mutation status, rather than CT morphology, is the established determinant of targeted-therapy response under current NCCN guidelines. We clarify this rationale explicitly, as it was not sufficiently developed in the earlier version of this manuscript. First, IDVTRM does not predict mutation status from imaging: mutation class (Mut_vec) is treated strictly as an external input obtained through standard-of-care molecular testing (tissue or liquid biopsy), consistent with

existing clinical workflow. IDVTRM therefore validates and ranks candidate drugs conditional on the externally supplied genotype; it does not replace molecular testing. Second, the imaging feature vector (LC_feat) is included only as a secondary, hypothesis-generating enrichment signal rather than as a primary causal input. CT-defined nodule phenotype (solid, part-solid, or pure ground-glass) and Lung-RADS category have well-documented, although imperfect, associations with tumour histological subtype and invasiveness; for example, part-solid and pure ground-glass nodules are disproportionately associated with lepidic-predominant, less invasive adenocarcinoma, whereas solid nodules are more frequently associated with invasive disease. These associations may yield weak, indirect correlations with treatment response beyond what genotype alone can explain. This provides an associative rationale for including LC_feat as an auxiliary feature, not as a claim that nodule imaging appearance directly or causally determines drug PKs or pharmacodynamics. Third, and most importantly, we have not performed an ablation analysis to isolate the incremental contribution of LC_feat relative to a mutation-only baseline. The reported drug-response AUC of 0.918 therefore reflects the fused model and should not be interpreted as evidence that imaging features improve prediction beyond genotype alone. We regard this ablation, together with prospective validation of any incremental benefit demonstrated, as a necessary prerequisite before the imaging-to-pharmacology fusion component of IDVTRM can be considered clinically justified rather than exploratory. In its present form, IDVTRM is best understood as a computational decision-support and prioritisation layer built on, and strictly subordinate to, standard molecular diagnostics and oncologists' judgement, not as a substitute for genotype-driven prescribing and not yet as a validated clinical tool.

4.5. Integrated drug validation and therapeutic recommendation module and the psychosomatic burden of treatment decision distress

Treatment decision distress, referring to the anxiety experienced by patients during the period between diagnosis and initiation of therapy, represents a distinct and clinically significant psychosomatic phenomenon. The lengthy, manual process of mutation testing, drug selection, and PK review may extend over weeks, during which patients are confirmed to have cancer but remain uncertain about treatment. This period has been correlated with the highest distress scores, insomnia, depression, and measurable HPA axis dysregulation. The IDVTRM GREEN/AMBER/RED system (94.2% expert oncologist concordance) is designed to directly compress this

treatment-uncertainty window. IDVTRM automates and integrates mutation-stratified drug selection, lung-tissue penetration (K_p , lung) modelling, MDR1-mediated resistance flagging, ILD/pneumonitis risk prediction, and eight-target molecular docking validation, potentially reducing the diagnostic-to-treatment-decision interval from weeks to seconds. In this context, IDVTRM could constitute a temporally targeted psychosomatic intervention at a potentially distress-intensive stage of the cancer care pathway. From a psychosomatic medicine perspective, reducing the duration of treatment uncertainty may plausibly reduce cumulative psychological burden. Similarly, ILD prediction (AUC = 0.908 for osimertinib) could support proactive clinical management of patients at elevated predicted risk. Such patients could receive psychoeducation, structured symptom monitoring, and a predefined distress-management plan before treatment initiation, potentially reducing the acute psychological burden associated with uncertainty surrounding this serious complication. Whether these interventions actually reduce patient-reported distress, anxiety, or treatment-related psychological morbidity remains to be established prospectively.

4.6. Psycho-oncology stepped care protocol: Clinical implementation of PulmoNet-X

These three UCRS risk levels generated by PulmoNet-X provide a natural, clinically implementable framework for a stepped psycho-oncology care protocol. The protocol outlined below is based on the NCCN Distress Management guidelines⁴³, as well as cognitive behavioural therapy (CBT)-based health-anxiety frameworks and standards for psycho-oncology referral practice, and is mapped to the tiers of output from PulmoNet-X.

- (a) Level C (low risk: $p_{\text{mal}} \leq 0.30$, $U_{\text{epi}} < 0.15$; Lung-RADS 1–2)

The psychological intervention consists of brief, structured reassurance and communication. The UCRS probability output is verbally communicated to the radiologist or screening nurse using a standard script: “your scan has identified a very small nodule, and it is recommended that you have a follow-up scan annually.” Patients scoring ≥ 4 on the Distress Thermometer at this stage are offered a single psychoeducation session on uncertainty-tolerance and catastrophic thinking, delivered by a trained nurse specialist or clinical psychologist. No routine psychological referral is indicated; however, written materials on managing cancer-related worry are provided to all Level C patients. Follow-up protocol includes annual LDCT with repeat UCRS stratification.

- (b) Level B (indeterminate: $0.30 < p_{\text{mal}} \leq 0.70$, or $U_{\text{epi}} \geq 0.15$; Lung-RADS 3–4A)

Psychological intervention consisting of structured psycho-oncology support. Patients in this tier may face substantial cumulative psychosomatic burden because they are neither reassured nor conclusively diagnosed, and the epistemic uncertainty captured by $U_{\text{epi}} \geq 0.15$ may correspond to the psychological construct of intolerance of uncertainty. The following stepped interventions are proposed: (i) structured psychoeducation session (30 minutes) focused on cognitive reappraisal of ambiguous medical information, delivered within five working days of the UCRS output; (ii) administration of the HADS³⁹ and Distress Thermometer at baseline and at 4-week follow-up; (iii) For patients scoring ≥ 8 on the HADS-Anxiety subscale, referral to a four-session CBT-based uncertainty-tolerance programme with documented outcome monitoring is recommended; (iv) Radiologist review of UCRS Level B cases within 14 days, with the UCRS probability estimate and uncertainty score communicated to the reviewing radiologist as part of the clinical decision-support output. Importantly, the clinical information provided by the UCRS uncertainty score is not limited to the malignancy probability alone: cases identified as Level B because of high epistemic uncertainty ($U_{\text{epi}} \geq 0.15$) may warrant additional imaging (e.g., PET-CT, contrast-enhanced CT), when clinically appropriate, to reduce uncertainty and potentially shorten the diagnostic pathway.

- (c) Level A (high risk: $p_{\text{mal}} > 0.70$, $U_{\text{epi}} < 0.15$; Lung-RADS 4B–5)

Psychological intervention includes intensive psycho-oncology support with parallel clinical action. A high-confidence, high-risk classification triggers simultaneous clinical and psychological pathways. Clinical pathway: immediate multidisciplinary team (MDT) referral; generation of the IDVTRM drug validation report within 24 hours; and initiation of mutation testing. IDVTRM integrates mutation-stratified drug selection, K_p , lung modelling, MDR1 resistance flagging, ILD/pneumonitis risk prediction, and eight-target molecular docking validation within a single computational pipeline. The computational processing time for IDVTRM is included in the 4.82 s/scan inference latency reported in Algorithm 1. However, the overall clinical diagnostic-to-treatment-decision interval depends on factors beyond computational processing—including time to molecular testing results, MDT scheduling, and clinical communication pathways—none of which were measured in this study. The target of fewer than five working days represents a proposed clinical implementation aspiration for the integrated pathway, not

an empirically measured outcome. Psychological pathway includes: (i) HADS, Distress Thermometer, and PHQ-9 administered within 48 hours of Level A classification; (ii) dedicated oncology psychologist consultation within five working days; (iii) for patients with HADS-Anxiety ≥ 11 or PHQ-9 ≥ 10 , immediate referral to evidence-based psychological intervention (CBT, acceptance and commitment therapy, or mindfulness-based cognitive therapy as appropriate); (iv) an ILD-risk flag from IDVTRM triggers pre-emptive respiratory distress psychoeducation and a written symptom-monitoring and distress-management plan before treatment initiation; (v) family/carer inclusion in the psychoeducation session, given the evidence that carer distress independently predicts patient distress outcomes.⁴⁶ This Level A protocol operationalises PulmoNet-X as a psychosomatic care framework in which AI output directly triggers standardised psychological assessment and intervention as an integrated component of the clinical pathway rather than an optional adjunct.

Future validation of this stepped-care protocol should include prospective measurement of HADS, GAD-7, and PHQ-9 scores at UCRS-triggered time points in a clinical LDCT screening cohort. Such a study would enable formal testing of the hypothesis that PulmoNet-X UCRS stratification predicts psychological distress trajectories and whether FPRM-mediated reductions in FP findings are associated with reduced GAD incidence. This prospective patient-reported outcome evidence would be necessary to determine whether PulmoNet-X can progress from an AI system with a psychosomatic framing to a fully validated psychosomatic intervention with prospectively measured psychological endpoints.

4.7. Limitations

Several important limitations of this study must be acknowledged. Most importantly, from a psychosomatic medicine perspective, psychological outcomes were not directly measured in this study. No patient cohort was evaluated using validated psychometric instruments such as the GAD-7, PHQ-9, HADS, or the Distress Thermometer. The psychosomatic benefits suggested for PulmoNet-X are based on psycho-oncology literature, mechanistic reasoning involving the HPA-EGFR transactivation axis, and quantitative extrapolation from published prevalence rates of anxiety in LDCT screening populations. Until psychological outcomes are directly measured at UCRS-triggered time points in a prospective clinical study, these assertions should not be regarded as empirically demonstrated. The stepped-care framework outlined in Section 4.6 is a proposed implementation framework and not a tested intervention. Future investigations should prioritise a prospective cohort study using the same

measures of screening-related distress (GAD-7, PHQ-9, HADS) at baseline, post-LDCT, and post-treatment-decision time points in a patient cohort stratified by UCRS tier, to formally test the relationship between PulmoNet-X's calibrated outputs and screening-related distress.

External validation was performed on an independent cohort of 150 chest CT scans from a publicly available NIH Chest CT dataset that was not used at any stage of model training or hyperparameter selection. PulmoNet-X yielded strong and consistent results across all evaluated metrics: sensitivity at 1 FP/scan = 98.1%, CPM = 97.4%, malignancy AUC = 99.3%, accuracy = 98.0%, sensitivity (malignancy) = 97.9%, specificity = 98.2%, F1-score = 98.0%, DSC = 95.8%, HD95 = 1.6 mm, precision = 98.3%, recall = 97.9%, and Lung-RADS κ = 0.94 (Table 6). These findings provide preliminary evidence that PulmoNet-X generalises effectively beyond the LIDC-IDRI training distribution. Nonetheless, this cohort does not contain multi-radiologist, per-nodule annotations equivalent to those available in LIDC-IDRI. Therefore, large-scale, formally annotated external validation using cohorts such as NLST or UKLS, incorporating diverse scanner protocols and multi-ethnic populations, remains necessary to establish broader generalisability.

Several additional limitations warrant consideration. Prospective validation was not performed as all analyses were retrospective; therefore, prospective validation in an operational LDCT screening programme is required before any clinical deployment. The model has also not undergone multi-centre evaluation, and its performance on European, Asian, and lower-income-country screening datasets using diverse scanner protocols remains unknown. In addition, limited ethnic and demographic diversity prevents assessment of potential performance differences across sex, ethnicity, age group, or smoking history subgroups. Dataset selection bias is another concern, as excluding indeterminate nodules (Likert = 3) from malignancy classification may inflate the observed AUC relative to real-world prevalence. From a regulatory perspective, PulmoNet-X has not been submitted to any regulatory framework (FDA 510(k), CE mark, or equivalent); prospective regulatory approval and post-market surveillance planning would therefore be required before clinical deployment. These limitations should be prioritised in future prospective validation studies. The procedural harms associated with FP screening results, including needle-biopsy morbidity and the surgical risk associated with resection of benign nodules, were not formally quantified in this study. The benefit-harm analysis of FPRM was primarily considered from a psychosomatic perspective (GAD reduction). A comprehensive health-economic and procedural harm-reduction analysis using

published complication-rate data would strengthen the clinical case for FPRM deployment and should therefore be addressed in future work. Statistical comparisons against baseline methods (Table 2) relied on Wilcoxon signed-rank tests with Bonferroni correction rather than bootstrapped confidence intervals or standardised effect sizes for pairwise differences. Several reported margins were numerically modest (on the order of 1–3 percentage points) despite reaching statistical significance; therefore, their practical or clinical significance should be interpreted cautiously. Reporting effect sizes together with per-comparison confidence intervals would provide a more complete assessment of the magnitude and precision of these differences and is recommended for future validation studies.

5. Conclusion

PulmoNet-X represents, to the best of our knowledge, the first AI framework for lung cancer screening explicitly architected with psychosomatic burden reduction as a primary clinical objective alongside imaging and pharmacological performance, thereby proposing a design paradigm that we contend should become a standard expectation in the field of AI-assisted cancer screening. This claim is based on our systematic literature search; however, a comprehensive review was not performed, and readers should interpret it accordingly. Across somatic performance metrics, PulmoNet-X achieves state-of-the-art results on LIDC-IDRI: nodule detection CPM = 93.7%, malignancy classification AUC = 97.8%, segmentation DSC = 84.6%, and Lung-RADS stratification agreement $\kappa = 0.87$ (all $p < 0.001$ versus all comparators, Wilcoxon signed-rank test with Bonferroni correction). From a psychosomatic standpoint, three mechanistic contributions collectively address different components of the LDCT burden pathway: (i) FPRM reduces FPs from 4.8 to 1.0 FP/scan (95% CI: 0.84–1.18), eliminating an estimated 380 unnecessary recall investigations and 144 incident GAD cases per 1,000 screened patients (estimated range: 108–178; mechanistic extrapolation from published prevalence data); (ii) UCRS (ECE = 0.016) provides calibrated uncertainty outputs that could directly support psychoeducation and uncertainty-tolerance training, with the three-tier risk stratification (Level A/B/C) mapped onto a stepped psycho-oncology care model aligned with NCCN Distress Management guidelines; and (iii) IDVTRM (drug-response AUC = 0.918, K_p , lung $R^2 = 0.986$, docking $r = 0.978$) is designed to shorten the diagnostic-to-treatment decision interval, with the potential to reduce uncertainty during the most distress-intensive period of the lung cancer patient journey. This effect, as well as the imaging-derived enrichment underlying it, remains unvalidated and requires the ablation

and prospective studies described above. The mechanistic integration of the stress–EGFR transactivation axis into the pharmacological validation framework provides a biologically plausible rationale for a future generation of AI systems that formally incorporate psychosomatic status as a pharmacological covariate in drug-response prediction models. PulmoNet-X therefore demonstrates the feasibility of designing AI-assisted cancer screening with psychosomatic burden reduction as an explicit clinical objective. However, whether the design philosophy has the potential to enhance rather than compromise the clinical and humanistic value that could ultimately be delivered to patients remains to be established. Direct psychological outcomes, including GAD-7, PHQ-9, and HADS, were not measured in the present study; consequently, the proposed psychosomatic benefits are literature-supported and mechanistically grounded hypotheses rather than demonstrated clinical effects. Similarly, the imaging and pharmacological analyses were primarily developed and evaluated using a single dataset (LIDC-IDRI) with an independent external validation, and the imaging-to-pharmacology linkage within IDVTRM has not been ablated against a genotype-only baseline; generalisability to real-world, multi-centre LDCT screening populations is therefore unproven. Prospective clinical validation, together with external multi-centre testing and the ablation study described above, constitutes the necessary next step before any clinical benefit claimed in this study can be regarded as established or translated into routine practice.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: R. Shiva Shankar
Formal analysis: R. Shiva Shankar, N. Deshai
Investigation: R. Shiva Shankar
Methodology: R. Shiva Shankar, N. Deshai
Validation: K. Amrutha, Ch. Ravi Swaroop
Writing–original draft: R. Shiva Shankar, N. Deshai
Writing–review & editing: All authors

Ethics approval and consent to participate

Not applicable. Publicly available anonymized datasets (LIDC-IDRI and LUNA16) were used in this study.

Consent for publication

Not applicable.

Availability of data

The datasets used in this study are publicly available from the LIDC-IDRI and LUNA16 repositories.

References

1. Sung H, Ferlay J, Siegel RL, *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209-249.
doi: 10.3322/caac.21660
2. Bray F, Laversanne M, Sung H, *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2024;74(3):229-263.
doi: 10.3322/caac.21834
3. Thaker PH, Han LY, Kamat AA, *et al.* Chronic stress promotes tumor growth and angiogenesis in a mouse model of ovarian carcinoma. *Nat Med.* 2006;12(8):939-944.
doi: 10.1038/nm1447
4. Wu GX, Raz DJ, Brown L, Sun V. Psychological burden associated with lung cancer screening: a systematic review. *Clinical Lung Cancer.* 2016;17(5):315-324.
doi: 10.1016/j.clcc.2016.03.007
5. Kim-Fuchs C, Le CP, Pimentel MA, *et al.* Chronic stress accelerates pancreatic cancer growth and invasion: a critical role for beta-adrenergic signaling in the pancreatic microenvironment. *Brain Behav Immun.* 2014;40:40-47.
doi: 10.1016/j.bbi.2014.02.019
6. Ding J, Li A, Hu Z, Wang L. Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks. In: Descoteaux M *et al.*, eds. Medical Image Computing and Computer Assisted Intervention — MICCAI 2017. Lecture Notes in Computer Science, vol 10435. Cham: Springer; 2017:559-567.
doi: 10.1007/978-3-319-66179-7_64
7. Zhu W, Liu C, Fan W, Xie X. Deep Lung: deep 3D dual path nets for automated pulmonary nodule detection and classification. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Tahoe, NV, USA: IEEE; 2018:673-681.
doi: 10.1109/WACV.2018.00079
8. Tang H, Zhang C, Xie X. NoduleNet: decoupled false positive reduction for pulmonary nodule detection and segmentation. In: Shen D, *et al.*, eds. Medical Image Computing and Computer Assisted Intervention — MICCAI 2019. Lecture Notes in Computer Science, vol 11769. Cham: Springer; 2019:266-274.
doi: 10.1007/978-3-030-32226-7_30
9. Dou Q, Chen H, Yu L, Qin J, Heng PA. Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Trans Biomed Eng.* 2017;64(7):1558-1567.
doi: 10.1109/TBME.2016.2613502
10. Chen J, Lu Y, Yu Q, *et al.* TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. *arXiv.* Preprint posted online 2021.
doi: 10.48550/ARXIV.2102.04306
11. Cao H, Wang Y, Chen J, *et al.* Swin-Unet: Unet-like pure transformer for medical image segmentation. In: European Conference on Computer Vision (ECCV) Workshops. LNCS vol 13803. Cham: Springer; 2023:205-218.
doi: 10.1007/978-3-031-25066-8_9
12. Hatamizadeh A, Tang Y, Nath V, *et al.* UNETR: transformers for 3D medical image segmentation. In: 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, HI, USA: IEEE; 2022:1748-1758.
doi: 10.1109/WACV51458.2022.00181
13. Setio AAA, Ciompi F, Litjens G, *et al.* Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Trans Med Imaging.* 2016;35(5):1160-1169.
doi: 10.1109/TMI.2016.2536809
14. Setio AAA, Traverso A, de Bel T, *et al.* Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. *Med Image Anal.* 2017;42:1-13.
doi: 10.1016/j.media.2017.06.015
15. Reddy CNK, Siva Kumar M. Deep Learning Framework on Early Detection of Lung Cancer with Medical Imaging based on the Cloud. In: 2026 Third International Conference on Networking and Communications (ICNWC). Chennai, India: IEEE; 2026:1-7.
doi: 10.1109/icnwc68145.2026.11517965
16. Kumari P, Goel L. Emerging computational intelligence-based techniques for lung cancer diagnosis and classification on chest CT scan images: a comprehensive survey. *Artif Intell Rev.* 2026;59:44.
doi: 10.1007/s10462-025-11374-9
17. Gal Y, Ghahramani Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *arXiv.* Preprint posted online 2015.
doi: 10.48550/arXiv.1506.02142
18. Lakshminarayanan B, Pritzel A, Blundell C. Simple and Scalable Predictive Uncertainty Estimation using Deep

- Ensembles. *arXiv*. Preprint posted online 2016.
doi: 10.48550/ARXIV.1612.01474
19. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep*. 2017;7(1):42717.
doi: 10.1038/srep42717
20. Pires DEV, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem*. 2015;58(9):4066-4072.
doi: 10.1021/acs.jmedchem.5b00104
21. Xiong G, Wu Z, Yi J, *et al*. ADMETLab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res*. 2021;49(W1):W5-14.
doi: 10.1093/nar/gkab255
22. Nan M, Chiru C, Florea AM. MTL-LungNet: A Unified Model for Nodule Segmentation and Malignancy Classification. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023 Workshops (Lecture Notes in Computer Science). Cham, Switzerland: Springer; 2025:213-226.
doi: 10.1007/978-3-032-11402-0_16
23. Armato SG III, McLennan G, Bidaut L, *et al*. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans. *Med Phys*. 2011;38(2):915-931.
doi: 10.1118/1.3528204
24. Heerink WJ, de Bock GH, de Jonge GJ, Groen HJM, Vliegenthart R, Oudkerk M. Complication rates of CT-guided transthoracic lung biopsy: meta-analysis. *Eur Radiol*. 2017;27(1):138-148.
doi: 10.1007/s00330-016-4357-8
25. Carleton RN. Fear of the unknown: one fear to rule them all? *J Anxiety Disord*. 2016;41:5-21.
doi: 10.1016/j.janxdis.2016.03.011
26. Lustberg T, van Soest J, Gooding M, *et al*. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. *Radiother Oncol*. 2018;126(2):312-317.
doi: 10.1016/j.radonc.2017.11.012
27. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 1997;23(1-3):3-26.
doi: 10.1016/s0169-409x(96)00423-1
28. Schyman P, Liu R, Desai V, Wallqvist A. vNN web server for ADMET predictions. *Front Pharmacol*. 2017;8:889.
doi: 10.3389/fphar.2017.00889
29. Dong J, Wang NN, Yao ZJ, *et al*. ADMETlab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. *J Cheminform*. 2018;10(1):29.
doi: 10.1186/s13321-018-0283-x
30. Eckerling A, Ricon-Becker I, Sorski L, Sandbank E, Ben-Eliyahu S. Stress and cancer: mechanisms, significance and future directions. *Nat Rev Cancer*. 2021;21(12):767-785.
doi: 10.1038/s41568-021-00395-5
31. Sood AK, Bhatti R, Kamat AA, *et al*. Stress hormone-mediated invasion of ovarian cancer cells. *Clin Cancer Res*. 2006;12(2):369-375.
doi: 10.1158/1078-0432.CCR-05-1698
32. McNitt-Gray MF, Armato III SG, Meyer CR, *et al*. The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation. *Acad Radiol*. 2007;14(12):1464-1474.
doi: 10.1016/j.acra.2007.07.021
33. Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE; 2017:5987-5995.
doi: 10.1109/CVPR.2017.634
34. Dai J, Qi H, Xiong Y, *et al*. Deformable convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE; 2017:764-773.
doi: 10.1109/ICCV.2017.89
35. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010;31(2):455-461.
doi: 10.1002/jcc.21334
36. Ardila D, Kiraly AP, Bharadwaj S, *et al*. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med*. 2019;25(6):954-961.
doi: 10.1038/s41591-019-0447-x
37. Huang H, Yang F, Xiao T, Tang D, Wang Z, Gao W. Enhancing lung nodule classification with triple union efficient attention and neural architecture search. *J Imaging Inform Med*. 2026;1-5.
doi: 10.1007/s10278-025-01807-w
38. Mok TS, Wu YL, Ahn MJ, *et al*. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. *N Engl J Med*. 2017;376(7):629-640.
doi: 10.1056/NEJMoa1612674

39. Chen J, Pan Y, Huang D, *et al.* Psychological distress in patients with detected pulmonary nodules. *Transl Lung Cancer Res.* 2026;15(2):34.
doi: 10.21037/tlcr-2025-aw-1153
40. Byrne MM, Weissfeld J, Roberts MS. Anxiety, Fear of Cancer, and Perceived Risk of Cancer following Lung Cancer Screening. *Med Decis Making.* 2008;28(6):917-925.
doi: 10.1177/0272989x08322013
41. Mishel MH. Uncertainty in Illness. *Image J Nurs Scholarsh.* 1988;20(4):225-232.
doi: 10.1111/j.1547-5069.1988.tb00082.x
42. Degner LE, Sloan JA. Decision making during serious illness: what role do patients really want to play? *J Clin Epidemiol.* 1992;45(9):941-950.
doi: 10.1016/0895-4356(92)90110-9
43. Riba MB, Donovan KA, Ahmed K, *et al.* NCCN Guidelines® Insights: Distress Management, Version 2.2023. *J Natl Compr Canc Netw.* 2023;21(5):450-457.
doi: 10.6004/jnccn.2023.0026
44. Chang CH, Lee CH, Ko JC, *et al.* Effect of β -Blocker in Treatment-Naïve Patients With Advanced Lung Adenocarcinoma Receiving First-Generation EGFR-TKIs. *Front Oncol.* 2020;10:583529.
doi: 10.3389/fonc.2020.583529
45. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol.* 2001;2(2):127-137.
doi: 10.1038/35052073
46. Wang T, Sun J, Gu D, Shen S, Zhou Y, Wang Z. Dyadic effects of social support, illness uncertainty on anxiety and depression among lung cancer patients and their caregivers: a cross-sectional study. *Supportive Care Cancer.* 2023;31(7):402.
doi: 10.1007/s00520-023-07876-3