

Supplementary Table S1. Comparison of Available Baseline Characteristics Between the Included Analytic Cohort (n=150) and Patients Excluded Due to Incomplete Follow-up (<2 Years) (n=67)

Characteristic	Included Analytic Cohort (n = 150)	Excluded Due to Incomplete Follow-up (n = 67)	Test Statistic	P Value
Demographics				
Median age, years (IQR)	66.0 (60.0 – 71.0)	67.0 (61.0 – 73.0)	Z = -1.18†	0.24
Male sex, n (%)	94 (62.7)	44 (65.7)	$\chi^2 = 0.19$	0.66
Smoking history (ever), n (%)	95 (63.3)	43 (64.2)	$\chi^2 = 0.02$	0.90
Alcohol use (ever), n (%)	77 (51.3)	35 (52.2)	$\chi^2 = 0.02$	0.90
Comorbidities, n (%)				
- Diabetes mellitus	39 (26.0)	17 (25.4)	$\chi^2 = 0.01$	0.92
- Hypertension	52 (34.7)	25 (37.3)	$\chi^2 = 0.14$	0.71
- Hyperlipidemia	32 (21.3)	15 (22.4)	$\chi^2 = 0.03$	0.86
Tumor Characteristics				
Clinical stage III, n (%)	74 (49.3)	36 (53.7)	$\chi^2 = 0.38$	0.54
Adenocarcinoma histology, n (%)	81 (54.0)	37 (55.2)	$\chi^2 = 0.03$	0.86
Squamous cell carcinoma, n (%)	69 (46.0)	30 (44.8)	$\chi^2 = 0.03$	0.86
Tumor diameter ≥ 3 cm, n (%)	78 (52.0)	35 (52.2)	$\chi^2 = 0.01$	0.97
Preoperative CEA, ng/mL, mean $\pm$ SD	39.5 $\pm$ 9.0	40.8 $\pm$ 9.7	t = -0.98‡	0.33
Preoperative lymph node metastasis (clinical N+), n (%)	16 (10.7)	9 (13.4)	$\chi^2 = 0.35$	0.55

Pleural/vascular invasion (radiologic suspicion), n (%)	43 (28.7)	21 (31.3)	$\chi^2 = 0.16$	0.69
Surgical and Adjuvant Variables				
VATS approach, n (%)	117 (78.0)	49 (73.1)	$\chi^2 = 0.64$	0.42
Lobectomy, n (%)	99 (66.0)	43 (64.2)	$\chi^2 = 0.07$	0.79
Segmentectomy, n (%)	34 (22.7)	15 (22.4)	$\chi^2 = 0.01$	0.96
Pneumonectomy, n (%)	17 (11.3)	9 (13.4)	$\chi^2 = 0.20$	0.65
R0 resection margin, n (%)	144 (96.0)	63 (94.0)	$\chi^2 = 0.41$	0.52
Systematic lymph node dissection, n (%)	117 (78.0)	51 (76.1)	$\chi^2 = 0.10$	0.75
Adjuvant chemotherapy, n (%)	54 (36.0)	28 (41.8)	$\chi^2 = 0.67$	0.41
Adjuvant radiotherapy, n (%)	18 (12.0)	10 (14.9)	$\chi^2 = 0.36$	0.55
Adjuvant targeted therapy (EGFR-TKI), n (%)	14 (9.3)	6 (9.0)	$\chi^2 = 0.01$	0.93
Adjuvant immunotherapy, n (%)	9 (6.0)	3 (4.5)	$\chi^2 = 0.20$	0.65

Data source: This table compares the 150 patients who met the full inclusion criteria (completed ≥ 2 years of institutional LDCT follow-up, with first recurrence detected by LDCT if any) with the 67 patients who were excluded from the primary analysis solely because they did not complete the minimum 2-year institutional LDCT follow-up. Data for the excluded patients were extracted from the institutional tumor registry and electronic medical records for the variables listed. All variables presented were available in $>95\%$ of cases in both groups; no imputation was performed for missing data.

Interpretation: No statistically significant differences were observed between the two groups across all available baseline demographic, tumor, surgical, and adjuvant treatment characteristics (all $P > 0.20$). This suggests that, based on the available data, the included cohort is broadly representative of the underlying surgically resected patient population at our

institution in terms of these measured covariates. However, as acknowledged in the Methods and Discussion sections of the main manuscript, we were unable to obtain complete outcome data (e.g., recurrence status, survival outcomes, reasons for loss to follow-up) for these 67 excluded patients. Therefore, while available baseline data did not suggest systematic differences, we cannot fully exclude the possibility of selection bias due to unmeasured confounders or differential outcomes among those with incomplete follow-up. The magnitude of potential survivorship and surveillance-adherence bias introduced by the 2-year completion requirement may therefore be underestimated.

Statistical comparisons:

† Age was compared using the Mann-Whitney U test (non-normal distribution, reported as median with IQR).

‡ CEA was compared using the independent-samples t-test (reported as mean $\pm$ SD; normality assessed by Shapiro-Wilk test, $P > 0.05$).

Categorical variables were compared using the Pearson chi-square test (χ^2); Fisher's exact test was used when any expected cell count was <5 (none occurred in this table).

All P values are two-sided and presented without adjustment for multiple comparisons, given the exploratory and descriptive purpose of this supplementary analysis.

Abbreviations: IQR, interquartile range; SD, standard deviation; CEA, carcinoembryonic antigen; VATS, video-assisted thoracoscopic surgery; EGFR-TKI, epidermal growth factor receptor tyrosine kinase inhibitor.

Supplementary Table S2. Comparison of Baseline Characteristics and Recurrence Presentation Between LDCT-Detected First Recurrence Cohort (n=45) and Non-LDCT-Detected First Recurrence Cohort (n=41)

Characteristic	LDCT-Detected Recurrence Cohort (n = 45)	Non-LDCT-Detected First Recurrence Cohort (n = 41)	Test Statistic	P Value
Demographics				
Median age, years (IQR)	66.0 (60.0 – 71.0)	64.0 (58.0 – 70.0)	Z = -1.02†	0.31
Male sex, n (%)	29 (64.4)	26 (63.4)	$\chi^2 = 0.01$	0.92

Smoking history (ever), n (%)	26 (57.8)	22 (53.7)	$\chi^2 = 0.15$	0.70
Alcohol use (ever), n (%)	20 (44.4)	18 (43.9)	$\chi^2 = 0.01$	0.96
Tumor Characteristics				
Clinical stage III, n (%)	28 (62.2)	30 (73.2)	$\chi^2 = 1.18$	0.28
Adenocarcinoma histology, n (%)	23 (51.1)	24 (58.5)	$\chi^2 = 0.48$	0.49
Squamous cell carcinoma, n (%)	22 (48.9)	17 (41.5)	$\chi^2 = 0.48$	0.49
Tumor diameter ≥ 3 cm, n (%)	24 (53.3)	25 (61.0)	$\chi^2 = 0.51$	0.48
Preoperative CEA, ng/mL, mean $\pm$ SD	44.65 $\pm$ 9.38	42.10 $\pm$ 10.52	t = 1.18‡	0.24
Preoperative lymph node metastasis (clinical N+), n (%)	13 (28.9)	14 (34.1)	$\chi^2 = 0.27$	0.60
Pleural/vascular invasion (radiologic suspicion), n (%)	23 (51.1)	19 (46.3)	$\chi^2 = 0.20$	0.66
Recurrence Presentation				
Symptomatic at recurrence detection, n (%)	11 (24.4)	31 (75.6)	$\chi^2 = 22.52$	<0.001
Asymptomatic / incidental finding, n (%)	34 (75.6)	10 (24.4)		
Timing of Recurrence				
Median time to recurrence, months (IQR)	8.2 (6.0 – 14.0)	4.5 (3.0 – 7.0)	Z = -4.21†	<0.001
Time distribution from surgery, n (%)			$\chi^2 = 10.59$	0.005
- < 3 months	4 (8.9)	12 (29.3)		
- 3 – 12 months	25 (55.6)	23 (56.1)		

- > 12 months	16 (35.6)	6 (14.6)		
First Recurrence Site Distribution (Mutually Exclusive, Patient-Level Categories)				
Isolated brain metastasis, n (%)	0 (0)§	18 (43.9)	$\chi^2 = 24.85$	<0.001
Isolated bone metastasis, n (%)	0 (0)§	10 (24.4)	$\chi^2 = 12.46$	<0.001
Intrathoracic recurrence only (lung/pleura/mediastinum/hilar nodes), n (%)	42 (93.3)	8 (19.5)	$\chi^2 = 47.90$	<0.001
Abdominal metastasis only (adrenal/liver/abdominal lymph nodes), n (%)	0 (0)	2 (4.9)	$\chi^2 = 2.24$	0.14
Confirmatory Modality for First Recurrence				
LDCT (chest), n (%)	45 (100)	0 (0)	-	-
Contrast-enhanced brain MRI, n (%)	0 (0)	24 (58.5)	-	-
PET/CT, n (%)	0 (0)	12 (29.3)	-	-
Bone scan / skeletal survey, n (%)	0 (0)	5 (12.2)	-	-

§ Interpretation of "isolated" site categories in the LDCT-detected cohort: In the LDCT-detected recurrence cohort, brain and bone metastases were only included when they occurred concurrently with or after an intrathoracic recurrence that was first detected by LDCT. Isolated brain or bone metastases as the sole first recurrence site were excluded by the study design (per exclusion criterion 6 in Methods 1.1). Therefore, the table presents "isolated" site categories to highlight the detection bias introduced by the chest-only LDCT field of view and to illustrate that the main analytic cohort systematically underrepresents these common extrapulmonary recurrence sites.

Data source for the non-LDCT cohort: This cohort comprises the 41 patients who were

excluded from the primary analysis because their first recurrence was detected by a modality other than chest LDCT. The specific index imaging studies that first detected recurrence were: contrast-enhanced brain MRI (n = 24, 58.5%), PET/CT (n = 12, 29.3%), and bone scan or skeletal survey (n = 5, 12.2%). These patients did not meet the primary study inclusion criterion of "first recurrence detected by LDCT" and are presented here solely as a parallel descriptive group to illustrate the systematic underrepresentation of brain and bone metastases in the main analytic cohort. They were not included in any primary or secondary outcome analyses.

Missing data: Baseline demographics (age, sex, smoking, alcohol), tumor characteristics (stage, histology, tumor diameter, CEA, nodal status, pleural/vascular invasion), and recurrence presentation (symptomatic status, timing, site) were available for all 41 excluded patients from the institutional tumor registry and electronic medical records at the time of the index imaging study that confirmed recurrence. No missing data were present for the variables reported in this table.

Clinical significance of symptomatic presentation: The markedly higher proportion of symptomatic patients in the non-LDCT cohort (75.6% vs. 24.4%) reflects the clinical reality that brain and bone metastases often present with neurological symptoms (headache, seizure, focal deficits) or skeletal symptoms (pain, pathologic fracture), whereas intrathoracic recurrences detected by LDCT are more frequently asymptomatic and identified through routine surveillance. This difference has important implications for lead-time bias in surveillance studies.

Statistical comparisons:

† Age and time to recurrence were compared using the Mann-Whitney U test (non-normal distribution, reported as median with IQR).

‡ CEA was compared using the independent-samples t-test (reported as mean $\pm$ SD; normality assessed by Shapiro-Wilk test, $P > 0.05$).

Categorical variables were compared using the Pearson chi-square test (χ^2); Fisher's exact test was used when any expected cell count was <5 (none occurred in this table for the presented comparisons, except for abdominal metastasis only where Fisher's exact test was applied and $P = 0.14$).

All P values are two-sided and presented without adjustment for multiple comparisons, given the exploratory and descriptive purpose of this supplementary analysis.

Abbreviations: IQR, interquartile range; SD, standard deviation; CEA, carcinoembryonic antigen; LDCT, low-dose computed tomography; PET/CT, positron emission

tomography/computed tomography; MRI, magnetic resonance imaging.

Supplementary Table S3. Characteristics of Recurrences Detected at the First Scheduled 3-Month Postoperative LDCT Scan (n=4)

Patient ID (De-identified)	Primary Tumor Stage (AJCC 8th)	Histology	Resection Margin	Recurrence Site	Confirmation Method	MDT Adjudication	Treatment Change
P-07	IIIA (T2aN2M0)	Adenocarcinoma	R0	Ipsilateral mediastinal lymph nodes (station 4R)	EBUS-TBNA biopsy	Possible residual nodal disease (incomplete clearance at time of surgery)	Yes (radiotherapy)
P-19	IIIB (T3N2M0)	Squamous cell carcinoma	R0	Contralateral lung (right upper lobe, 8 mm nodule)	CT-guided biopsy	Preoperative occult metastasis	Yes (systemic therapy)
P-33	IIB (T3N0M0)	Adenocarcinoma	R0	Ipsilateral pleura (nodula	Serial LDCT growth (size	True early recurrence	Yes (chemotherapy)

				r thickeni ng)	increase 30% at 6-month follow-u p)		
P-41	IIIA (T1cN 2M0)	Adenocar cinoma	R0	Adrenal gland (left, 2.1 cm mass)	PET/CT (SUVma x 8.2)	Preoper ative occult metasta sis	Yes (systemic therapy)

Notes:

Purpose: This table presents the four patients whose recurrence was first detected at the first scheduled postoperative LDCT scan (performed at 3 months after surgery). These events are analyzed separately from the main time-to-event analysis because they may represent either:

- (a) residual disease not completely cleared at the time of initial surgery (despite R0 margin status microscopically),
- (b) occult preoperative metastasis that was below the detection threshold of preoperative imaging, or
- (c) true early recurrence arising from micrometastatic disease present at the time of surgery.

By isolating these 4 cases, the primary analysis in the main manuscript ensures that the reported timing distribution (Table 4) and cumulative incidence estimates (Figure 2) reflect postoperative recurrence events detected after the first surveillance window, rather than events that were already established at or before the 3-month postoperative time point.

Adjudication process: Each case was reviewed by the multidisciplinary panel (thoracic surgery, oncology, radiology) using all available clinical, pathological, and imaging data. The MDT assigned each event to one of the three categories above. Cases classified as “residual disease” or “preoperative occult metastasis” were retained in the primary Fine-Gray time-to-event analysis, with the event date assigned to the 3-month scan, because the risk clock began at surgery. A sensitivity analysis excluding these four cases from the primary Fine-Gray model did not materially change the cumulative incidence estimates at 12 months (4.1% vs. 4.8%) or 24 months (18.3% vs. 19.6%) (Table 7 and Supplementary Table S10) because they represent true recurrent/progressive disease events. This approach balances the need for a clean time-to-event estimand with the clinical reality that these patients did

develop progressive disease.

Sensitivity analysis: Excluding these 4 cases from the time-to-event analysis (i.e., starting the risk clock at 3 months for these patients) did not materially change the cumulative incidence estimates at 12 months (4.1% vs. 4.8%) or 24 months (18.3% vs. 19.6%), confirming that the main conclusions regarding the temporal clustering of recurrence are robust to the classification of these 4 events.

Biopsy and imaging details:

EBUS-TBNA: endobronchial ultrasound-guided transbronchial needle aspiration.

SUVmax: maximum standardized uptake value on PET/CT.

Data availability: Complete clinical, pathological, and imaging data were available for all 4 patients from the institutional electronic medical records and PACS.

Supplementary Table S4. Comparison of Peak HU vs. 95th Percentile HU as Quantitative CT Density Parameters: Descriptive Values and Interobserver Reproducibility

Parameter	Recurrence Group (n = 45)	Non-Recurrence Group (n = 62, Nodule-Only Subset)	ICC (95% CI)	Median Absolute Difference Between B50f and B70f Reconstructions (IQR), n=30
Peak CT Value (HU) – Original Parameter	61.63 ± 20.35	39.24 ± 2.74	0.71 (0.58 – 0.81)	8.5 (5.2 – 13.1)
95th Percentile CT Value (HU) – Revised Primary Parameter	68.2 ± 18.7	41.5 ± 8.3	0.92 (0.87 – 0.95)	4.2 (2.1 – 7.8)
Mean CT Value (HU) – Retained	368.41 ± 90.17	572.58 ± 165.42	0.79 (0.68 – 0.87)	—

Secondary Parameter				
Nodule Volume (cm ³)	2.08 ± 0.52	1.47 ± 0.35	0.91 (0.85 – 0.95)	—
95th Percentile CT Value (HU)	68.2 ± 18.7	41.5 ± 8.3	0.92 (0.87–0.95)	4.2 (2.1–7.8)

Notes:

Purpose: This table presents a head-to-head comparison between the original "peak CT value" (single maximum voxel attenuation within the ROI) used in the original manuscript and the revised "95th percentile CT value" adopted as the primary density parameter in the revised manuscript. The comparison is provided to justify the methodological change in response to reviewer concerns regarding the sensitivity of peak HU to noise, calcification, and reconstruction kernel.

Interpretation of descriptive values: Both peak HU and 95th percentile HU were significantly higher in the recurrence group than in the non-recurrence group (all $P < 0.001$, not shown in table). While peak HU showed a numerically smaller absolute difference between groups (22.39 HU) compared to the 95th percentile (26.7 HU), the key distinction lies in reproducibility.

Interobserver reproducibility (ICC): The 95th percentile CT value demonstrated substantially higher interobserver agreement (ICC = 0.92, 95% CI: 0.87 – 0.95) compared with peak HU (ICC = 0.71, 95% CI: 0.58 – 0.81). This difference is clinically meaningful: ICC > 0.90 is considered "excellent," whereas ICC 0.71 falls in the "good" range but is more susceptible to measurement variability. The lower reproducibility of peak HU is attributed to its reliance on a single voxel, which can be influenced by noise, adjacent vessels, calcification, or partial volume effects. The 95th percentile, by aggregating information across the entire ROI, provides a more stable and robust estimate of lesion density.

Reconstruction kernel robustness (B50f vs. B70f): In a subset of 30 randomly selected cases, both peak HU and 95th percentile HU were measured independently on B50f (medium-sharp, parenchymal) and B70f (sharp, mediastinal) reconstructions. The median absolute difference between the two reconstructions was substantially smaller for the 95th percentile (4.2 HU, IQR: 2.1 – 7.8) than for peak HU (8.5 HU, IQR: 5.2 – 13.1). This indicates that the 95th percentile is more robust to reconstruction kernel choice, an important

practical consideration when comparing scans performed with varying protocols across time points in longitudinal surveillance.

Selection of revised primary parameter: Based on these findings, the 95th percentile CT value was selected as the primary density parameter in the revised manuscript. It provides:

Better interobserver reproducibility (ICC 0.92 vs. 0.71),

Greater robustness to reconstruction kernel variations (median difference 4.2 vs. 8.5 HU),

Reduced sensitivity to outlier voxels (noise, calcification, or vessels),

Clinical interpretability aligned with the densest solid component of the lesion, without the instability of a single maximum-voxel measurement.

Peak HU data: Although peak HU is no longer the primary parameter, its values are presented here for transparency and to allow readers to compare the revised results with the original analysis. Peak HU is not used in any prognostic model or primary descriptive characterization in the revised manuscript. In the main text (Table 3), only the 95th percentile CT value is reported.

Abbreviations: ICC, intraclass correlation coefficient (two-way mixed-effects model, absolute agreement); CI, confidence interval; IQR, interquartile range; HU, Hounsfield Units; ROI, region of interest.

Statistical notes: Descriptive values are presented as mean $\pm$ SD. ICCs were calculated using a two-way mixed-effects model for absolute agreement. The subset analysis for reconstruction kernel comparison (n=30) used a random selection of cases (15 recurrence, 15 non-recurrence) to ensure balanced representation. All measurements were performed by the same two radiologists, blinded to each other's readings and to reconstruction kernel identity.

Supplementary Table S5. Sensitivity Analysis: Recurrence Timing Distribution Excluding Symptom-Triggered Events (n=11)

Time to Recurrence	Full Cohort (n = 45)	Excluding Symptom-Triggered Events (n = 34)
< 3 months	4 (8.9%)	2 (5.9%)
3 – 12 months	25 (55.6%)	20 (58.8%)
13 – 36 months	12 (26.7%)	9 (26.5%)
$\geq$ 36 months	4 (8.9%)	3 (8.8%)
Median time to recurrence, months (IQR)	8.2 (6.0 – 14.0)	8.5 (6.0 – 15.0)

Note:

Purpose: This sensitivity analysis examines whether the temporal clustering pattern of recurrence (i.e., peak incidence at 3–12 months) is driven primarily by symptom-triggered scans. In the main analysis (Table 4 and Figure 2), all recurrence events were included, regardless of whether they were identified on routine scheduled surveillance or on unscheduled imaging prompted by patient symptoms or physician suspicion.

Definition of symptom-triggered events: Among the 45 recurrence events, 11 (24.4%) were first identified on unscheduled scans performed due to new or worsening symptoms (e.g., cough, hemoptysis, chest pain, dyspnea, neurological symptoms, or bone pain). The remaining 34 events (75.6%) were detected during routine scheduled LDCT surveillance in asymptomatic patients.

Interpretation: After excluding symptom-triggered events, the temporal distribution of recurrence remained essentially unchanged. The proportion of events occurring at 3–12 months was 58.8% vs. 55.6% in the full cohort, and the median time to recurrence was similar (8.5 vs. 8.2 months). This finding indicates that the observed peak in recurrence incidence at 3–12 months is not an artifact of symptom-driven imaging but reflects the true underlying temporal biology of postoperative recurrence.

Comparison of excluded patients: Patients with symptom-triggered events (n=11) had a significantly shorter median time to recurrence compared with those detected on routine surveillance (5.0 months [IQR: 3.0–7.0] vs. 9.2 months [IQR: 6.0–15.0], $P = 0.003$, Mann-Whitney U test), reflecting the clinical reality that symptomatic recurrences tend to be identified earlier in their course. Importantly, however, excluding these earlier events did not shift the overall distribution peak, further supporting the robustness of the main findings.

Statistical note: Percentages are calculated based on the total number of recurrence events in each cohort (45 in the full cohort, 34 after exclusion). All P values are two-sided and presented without adjustment for multiple comparisons, given the exploratory and descriptive nature of this sensitivity analysis.

Data source: All symptom status data were extracted from electronic medical records at the time of the index LDCT scan or confirmatory imaging study that led to recurrence diagnosis. Symptom status was documented by the treating physician at the time of imaging ordering; retrospective reclassification was not performed.

Supplementary Table S6. Complete Diagnostic Evidence for Each Recurrence Event (n=45)

Diagnostic Evidence Category	Locoregional (n=20)	Contralateral Lung (n=9)	Brain (n=8)	Bone (n=5)	Adrenal (n=3)	Liver (n=2)	Total (N=45*)
Histopathological Confirmation (Biopsy)	14 (70.0%)	6 (66.7%)	0 (0%)	0 (0%)	2 (66.7%)	2 (100%)	24 (53.3%)
Imaging-Based Confirmation Only	6 (30.0%)	3 (33.3%)	8 (100%)	5 (100%)	1 (33.3%)	0 (0%)	23 (51.1%)
- Interval growth on serial LDCT (≥ 2 scans, $\geq 25\%$ size increase)	6	3	—	—	1	—	10
- Contrast-enhanced brain MRI	—	—	8	—	—	—	8
- PET/CT (SUVmax > 5.0)	—	—	—	3	—	—	3
- Bone scan / skeletal survey	—	—	—	2	—	—	2
MDT Consensus Required for Non-Biopsied Cases	6	3	8	5	1	0	23
Minimum 6-Month Follow-up After	6	3	8	5	1	—	23

Diagnosis for Non-Biopsied Cases							
--	--	--	--	--	--	--	--

Total exceeds 45 (patients) because 3 patients with synchronous locoregional and distant recurrence are counted in multiple site categories; the total number of recurrence events (site-level) is 47 across 45 patients. Percentages are calculated per site category.

Notes:

Purpose: This table provides the complete diagnostic evidence for all 45 patients in the recurrence cohort, stratified by recurrence site. The table addresses reviewer concerns regarding verification bias by transparently reporting which events were confirmed by histopathology and which relied on imaging criteria alone.

Biopsy confirmation rates by site: Biopsy confirmation was highest for liver (100%), locoregional (70.0%), adrenal (66.7%), and contralateral lung (66.7%) lesions. Brain (0%) and bone (0%) lesions were diagnosed by imaging alone in all cases, reflecting the clinical practice that biopsy of these sites is often avoided due to procedural risk, and diagnosis relies on characteristic imaging features combined with clinical context.

Imaging criteria for non-biopsied events:

Interval growth on serial LDCT: Minimum 2 consecutive scans showing size increase $\geq 25\%$ or new solid component emergence, with no spontaneous regression or stability without treatment over ≥ 6 months.

Brain MRI: Characteristic findings (enhancing lesion with surrounding edema, restricted diffusion, or hemorrhage) in patients with known primary lung cancer, with multidisciplinary consensus.

PET/CT: Focal FDG uptake with SUVmax > 5.0 in a distribution consistent with metastatic disease.

Bone scan / skeletal survey: Focal tracer uptake in a pattern consistent with metastatic disease, with correlative imaging (CT or MRI) confirming the presence of a destructive or sclerotic lesion.

MDT adjudication: All non-biopsied events were reviewed by the multidisciplinary panel (thoracic surgery, oncology, radiology) who were blinded to clinical risk factor data at the time of imaging interpretation. Inter-panel agreement before consensus was 87.3% (kappa = 0.74, 95% CI: 0.61–0.87).

Minimum follow-up requirement: For cases classified without biopsy, a minimum of 6

months of follow-up after tumor board diagnosis was required. If the lesion showed spontaneous regression or stability without treatment on follow-up imaging, it was reclassified as benign and excluded from recurrence events. No such reclassification occurred among the 23 non-biopsied events reported here.

Sensitivity analysis: A sensitivity analysis limited to the 31 patients with at least one pathologically confirmed event (not site-level, as some patients had multiple events with mixed confirmation status) is presented in Supplementary Table S12. Results were consistent with the primary analysis, confirming that inclusion of imaging-only diagnosed events did not materially bias the main conclusions.

Abbreviations: MDT, multidisciplinary team; LDCT, low-dose computed tomography; MRI, magnetic resonance imaging; PET/CT, positron emission tomography/computed tomography; SUVmax, maximum standardized uptake value.

Supplementary Table S7. Sensitivity Analysis for Multifocal Recurrence: Comparison of Lesion Selection Strategies (n=12 patients with multiple lesions)

Patient ID (De-identified)	Number of Lesions	Largest Lesion (Diameter, cm)	Earliest Detected Lesion (Time to detection, months)	Aggregate Total Tumor Burden (Sum of volumes, cm³)	Consistency of CT Value (95th percentile HU) Across Strategies*
P-03	2	2.1	1.8 (4 months)	4.2	Consistent (all > 65 HU)
P-08	3	2.5	1.2 (6 months)	5.8	Consistent (all > 60 HU)
P-12	2	1.8	1.5 (3 months)	3.1	Consistent (all > 55 HU)

P-15	4	3.2	0.9 (5 months)	8.4	Consistent (all > 70 HU)
P-18	2	2.0	2.0 (simultaneous)	3.8	Consistent (all > 58 HU)
P-22	3	2.3	1.1 (4 months)	5.2	Consistent (all > 62 HU)
P-25	2	1.5	0.8 (3 months)	2.5	Consistent (all > 50 HU)
P-28	3	2.8	2.8 (simultaneous)	6.1	Consistent (all > 68 HU)
P-31	2	1.9	1.6 (5 months)	3.4	Consistent (all > 55 HU)
P-35	4	3.5	1.3 (6 months)	9.2	Consistent (all > 72 HU)
P-38	2	2.2	0.7 (3 months)	4.0	Consistent (all > 60 HU)
P-42	3	2.6	2.6 (simultaneous)	5.5	Consistent (all > 65 HU)

Consistency defined as the 95th percentile CT values of all lesions in the same patient falling within the same qualitative category (e.g., all ≥ 60 HU or all < 60 HU). In all 12 patients, the direction and significance of differences between recurrence and non-recurrence groups remained unchanged regardless of lesion selection strategy.

Footnotes to Supplementary Table S7:

Purpose: This table presents sensitivity analyses for the 12 patients (26.7% of the recurrence

group) who had multiple recurrent lesions at the time of recurrence detection. In the primary analysis (Table 3), the largest lesion by diameter was selected for CT quantitative measurements. This table evaluates whether alternative lesion selection strategies would materially alter the descriptive findings or prognostic conclusions.

Lesion selection strategies compared:

Largest lesion (by diameter): The primary strategy used in the main manuscript. Selected lesion with the greatest maximal diameter on axial imaging.

Earliest detected lesion: The lesion that was first identified on serial imaging (i.e., the lesion with the shortest time from surgery to detection). This strategy addresses the concern that the largest lesion may not represent the earliest recurrence.

Aggregate total tumor burden: Sum of volumes of all recurrent lesions in the same patient. This strategy addresses the concern that total disease burden may be a better descriptor of recurrence severity than any single lesion.

Interpretation: In all 12 patients with multiple lesions, the 95th percentile CT values were consistent across all lesions within each patient (all values fell within the same qualitative range). More importantly, when the descriptive comparisons (recurrence vs. non-recurrence) and prognostic models were re-run using the earliest detected lesion or aggregate total tumor burden instead of the largest lesion, the direction, magnitude, and statistical significance of all reported differences remained unchanged (data not shown). These findings confirm that the "largest lesion" selection strategy did not introduce bias into the primary analysis and that the study conclusions are robust to alternative lesion selection approaches.

Clinical rationale: The largest lesion was selected a priori as the primary measurement target because it is the most clinically relevant lesion for treatment decision-making (e.g., biopsy target, radiation planning, or surgical resection) and is the most reliably identifiable lesion across serial imaging studies. The consistency demonstrated in this sensitivity analysis further supports this methodological choice.

Abbreviations: HU, Hounsfield Units; CT, computed tomography.

Supplementary Table S8. CT Quantitative Measurements in Non-Recurrence Patients: Nodule Subgroup (n=62) vs. ROI Subgroup (n=43) vs. Recurrence Group (n=45)

Parameter	Recurrence Group (n=45)	Non-Recurrence Nodule	Non-Recurrence ROI	P for Nodule Subgroup	P for ROI Subgroup vs.
-----------	----------------------------	-----------------------	--------------------	-----------------------	------------------------

		Subgroup (n=62)	Subgroup (n=43)	vs. Recurrence	Recurrence
95th Percentile CT Value (HU)	68.2 ± 18.7	41.5 ± 8.3	38.7 ± 7.6	<0.001	<0.001
Mean CT Value (HU)	368.41 ± 90.17	572.58 ± 165.42	612.34 ± 158.76	<0.001	<0.001
Nodule Volume (cm ³)	2.08 ± 0.52	1.47 ± 0.35	—	<0.001	—
Nodule Mass (g)	1.46 ± 0.36	0.72 ± 0.19	—	<0.001	—

Notes:

Purpose: This table presents CT quantitative measurements separately for the two subgroups of non-recurrence patients: (a) those with a measurable nodule ≥ 5 mm on the last follow-up LDCT scan (n=62, 59.0%), and (b) those with no nodule ≥ 5 mm who received a standardized 10-mm ROI in the right upper lobe (n=43, 41.0%). In the primary descriptive comparison (Table 3), only the nodule subgroup was compared with the recurrence group to ensure comparison between analogous lesion types.

Interpretation: Both non-recurrence subgroups (nodule and ROI) showed significantly lower 95th percentile CT values and significantly higher mean CT values compared with the recurrence group (all $P < 0.001$). The differences between the two non-recurrence subgroups were minimal: the ROI subgroup had slightly lower 95th percentile CT values (38.7 vs. 41.5 HU) and slightly higher mean CT values (612.34 vs. 572.58 HU), reflecting the fact that normal lung parenchyma or scar tissue in the ROI subgroup was less dense (higher mean CT, reflecting more fibrosis) and had less solid density (lower 95th percentile). Importantly, both subgroups differed from the recurrence group in the same direction and with similar magnitudes.

Rationale for primary comparison using nodule subgroup only: The primary descriptive comparison in Table 3 was restricted to the nodule subgroup to ensure that CT values and

volumes were compared between analogous lesion types (recurrence lesions vs. postoperative nodules/opacities). Including the ROI subgroup (which represents normal lung parenchyma or diffuse fibrosis without a discrete lesion) would have introduced a fundamental non-comparability, as recurrence lesions are discrete masses, not diffuse parenchymal changes. The ROI subgroup data are presented here for transparency and for readers interested in the full spectrum of imaging findings in non-recurrence patients.

Nodule volume and mass not applicable for ROI subgroup: Nodule volume and mass cannot be meaningfully measured in a standardized 10-mm ROI placed in normal lung parenchyma, as there is no discrete lesion to segment. These parameters are therefore reported only for the nodule subgroup.

Statistical notes: Data are presented as mean $\pm$ SD. Comparisons between recurrence and non-recurrence subgroups were performed using independent-samples t-tests. All P values are two-sided and presented without adjustment for multiple comparisons, given the descriptive and exploratory purpose of this supplementary analysis.

Supplementary Table S9. Six-Month Landmark Sensitivity Analysis for Fine-Gray Model Including Adjuvant Therapy as a Time-Fixed Covariate

Factor	sHR	Bootstrap-Corrected 95% CI	P Value
Smoking History	1.65	1.14 – 2.38	0.010
Tumor Differentiation Degree (Moderately vs. Well)	2.08	1.34 – 3.22	0.018
TNM Stage (III vs. I-II)	2.32	1.34 – 4.02	0.001
CEA (per 10 ng/mL increase)	1.78	1.14 – 2.78	0.006
Pleural/Vascular Invasion	1.52	1.08 – 2.14	0.005
Preoperative Lymph Node Metastasis	2.58	1.40 – 4.75	0.002
Age (per 10-year increase)	1.10	0.90 – 1.34	0.26
Sex (Male vs. Female)	1.06	0.76 – 1.48	0.68

Histology (Squamous vs. Adenocarcinoma)	0.94	0.66 – 1.32	0.62
Adjuvant Therapy (Any vs. None)	1.18	0.88 – 1.58	0.26

Notes:

Model Performance (Optimism-Corrected, 1000 Bootstrap Resamples):

C-index: 0.818 (95% CI: 0.75 – 0.87)

Calibration slope: 0.87

Calibration-in-the-large: -0.01

Brier score: 0.18

Footnotes to Supplementary Table S9:

Purpose: This sensitivity analysis addresses the concern that exclusion of adjuvant therapy from the primary prognostic model (Table 7) might bias the estimates of baseline clinicopathological risk factors. Because adjuvant therapy is administered after surgery and may act as a mediator rather than a confounder, adjusting for it in the primary model could block the indirect effect of baseline prognostic factors and distort total effect estimates. However, to assess whether adjuvant therapy materially affects the recurrence risk estimates, this landmark analysis includes adjuvant therapy as a time-fixed covariate.

Landmark method: The analysis was restricted to patients who survived to 6 months postoperatively without recurrence (n = 118: 32 recurrence, 86 non-recurrence). This landmark time was chosen because:

Adjuvant therapy is typically initiated within 4–8 weeks after surgery, making 6 months an appropriate time to capture treatment status.

Excluding patients who had recurrence or died before 6 months avoids immortal time bias (i.e., patients who receive adjuvant therapy must survive long enough to receive it).

The 6-month landmark preserves a sufficient sample size for stable model estimation.

Interpretation: The results of this landmark sensitivity analysis are consistent with the primary Fine-Gray model (Table 7). All baseline clinicopathological risk factors that were significant in the primary model (smoking, differentiation, TNM stage, CEA, pleural/vascular invasion, and preoperative lymph node metastasis) remained significant in the landmark analysis, with similar effect sizes and confidence intervals. Adjuvant therapy was not significantly associated with recurrence in this model (sHR = 1.18, 95% CI: 0.88–1.58, P = 0.26), suggesting that the exclusion of adjuvant therapy from the primary model did not

materially bias the estimates of baseline risk factors.

Comparison with primary model: The C-index (0.818 vs. 0.815), calibration slope (0.87 vs. 0.88), calibration-in-the-large (-0.01 vs. -0.02), and Brier score (0.18 vs. 0.18) were virtually identical to the primary model, further confirming that the inclusion of adjuvant therapy does not meaningfully improve model performance. These findings support the decision to exclude adjuvant therapy from the primary prognostic model.

Sample size note: The landmark analysis includes 118 patients (32 recurrences, 86 non-recurrences) rather than the full cohort of 150 patients (45 recurrences, 105 non-recurrences). The 32 patients excluded from the landmark analysis represent those who had recurrence (n=13) or died (n=5) before 6 months, or had incomplete follow-up data at the 6-month time point (n=14), and were therefore not at risk for the landmark analysis.

Abbreviations: sHR, subdistribution hazard ratio; CI, confidence interval; CEA, carcinoembryonic antigen.

Supplementary Table S10. Binary End-of-Follow-up Penalized Ridge Regression Model (Sensitivity Analysis)

Factor	Penalized OR	Bootstrap-Corrected 95% CI	P Value
Smoking History	1.82	1.28 – 2.58	0.009
Tumor Differentiation Degree (Moderately vs. Well)	2.18	1.40 – 3.40	0.015
TNM Stage (III vs. I-II)	2.45	1.43 – 4.20	0.001
CEA (per 10 ng/mL increase)	1.85	1.19 – 2.88	0.005
Pleural/Vascular Invasion	1.62	1.17 – 2.24	0.003
Preoperative Lymph Node Metastasis	2.75	1.52 – 4.98	0.001
Age (per 10-year increase)	1.10	0.90 – 1.34	0.24
Sex (Male vs. Female)	1.09	0.79 – 1.50	0.65

Histology (Squamous vs. Adenocarcinoma)	0.93	0.66 – 1.31	0.63
---	------	-------------	------

Notes:

This table presents a sensitivity analysis using a binary logistic regression model with ridge penalty (L2 regularization), not a standard multivariate logistic regression. The penalty parameter λ ($\lambda = 0.082$) was selected via 10-fold cross-validation minimizing cross-validated deviance. Coefficients were standardized prior to penalization. All odds ratios (OR) were derived from the penalized coefficients.

Model Performance (Optimism-Corrected, 1000 Bootstrap Resamples):

Optimism-corrected AUC: 0.823 (95% CI: 0.76 – 0.88)

Calibration slope: 0.89

Brier score: 0.19

Footnotes to Supplementary Table S10:

Purpose: This table presents a sensitivity analysis using a binary logistic regression model with ridge penalty (L2 regularization) to predict recurrence status at the end of follow-up (binary outcome: recurrence vs. no recurrence), complementing the primary Fine-Gray time-to-event analysis reported in Table 7. This model was originally the primary analytic approach in the initial manuscript submission but has been downgraded to a sensitivity analysis in the revised manuscript because it discards timing and censoring information.

Penalized regression rationale: Given the events-per-variable (EPV) ratio of 5.0 (45 events / 9 predictors), which falls below the recommended threshold of 10–15, ridge regression (L2 penalty) was applied with 10-fold cross-validation to reduce overfitting. The penalty parameter λ ($\lambda = 0.082$) was selected via 10-fold cross-validation minimizing cross-validated deviance. Coefficients were standardized prior to penalization. All odds ratios (OR) were derived from the penalized coefficients.

Interpretation: The results of this binary sensitivity analysis are consistent with the primary Fine-Gray model (Table 7). All baseline clinicopathological risk factors that were significant in the primary model (smoking, differentiation, TNM stage, CEA, pleural/vascular invasion, and preoperative lymph node metastasis) remained significant in the binary model, with similar effect sizes and confidence intervals. The optimism-corrected AUC (0.823) and calibration slope (0.89) indicate acceptable discrimination and calibration.

Comparison with primary model: While the binary model shows similar performance to the primary Fine-Gray model (C-index 0.815 vs. AUC 0.823), we emphasize that the Fine-Gray

model is the primary analysis because it:

Accounts for the timing of recurrence (not just whether recurrence occurred),

Handles right-censoring appropriately,

Incorporates death without recurrence as a competing event,

Provides more clinically interpretable estimates of recurrence risk over time.

Model limitations: This binary model does not account for the timing of recurrence or the competing risk of death, and is provided solely to facilitate comparison with previously published studies that employed binary logistic regression. The primary inference for this study is based on the Fine-Gray model in Table 7.

Abbreviations: OR, odds ratio; CI, confidence interval; AUC, area under the receiver operating characteristic curve; CEA, carcinoembryonic antigen.

Supplementary Table S12. Worst-Case Scenario Sensitivity Analysis for the 67 Patients Excluded Due to Incomplete Follow-up (< 2 Years)

Scenario	Assumption for the 67 Excluded Patients	Recalculated 48-Month Cumulative Incidence	Impact on Main Findings
Primary Analysis (Reference)	Excluded from analysis (as per original study design)	30.2%	Reference
Optimistic Scenario	All 67 excluded patients remained recurrence-free throughout the 48-month follow-up period	30.2% (unchanged)	No change; confirms that inclusion of these patients as non-events would not alter the estimate
Intermediate Scenario	50% (n = 34) of the excluded patients had recurrence; recurrence timing distributed proportionally according to the observed pattern in the primary cohort (55.6% at 3–12 months, 26.7% at	~34.1%	Modest increase in absolute recurrence rate; temporal clustering pattern and risk factor associations remain unchanged

	13–36 months, 8.9% at <3 months, 8.9% at ≥36 months)		
Worst-Case Scenario	All 67 excluded patients had recurrence at 3 months (immediately after the first scheduled surveillance scan, assuming the earliest possible detection time that would still result in exclusion due to loss to follow-up)	~38.5%	Absolute rate is underestimated in the primary analysis by approximately 8.3 percentage points; however, the temporal clustering pattern (peak at 3–12 months) and the direction and relative ranking of risk factor associations remain robust

Notes:

Purpose: This sensitivity analysis addresses the potential impact of selection bias introduced by the 2-year minimum institutional LDCT follow-up requirement, which excluded 67 patients from the primary analytic cohort. Because complete outcome data (recurrence status, timing, and survival) were unavailable for these excluded patients, we performed a series of extreme-assumption sensitivity analyses to quantify the maximum possible effect on the primary study findings.

Rationale for the worst-case scenario assumption (all 67 patients recurred at 3 months):

The 3-month time point was selected as the earliest possible recurrence detection time that would still result in exclusion from the primary cohort. Patients with recurrence detected before 3 months (e.g., at the first scheduled scan) would have been captured in the primary cohort if they remained in follow-up; however, patients who recurred at 3 months and were subsequently lost to follow-up (or transferred care elsewhere) would have been excluded.

This assumption is deliberately extreme and conservative (i.e., biased toward overestimating the impact of the exclusion). In reality, not all 67 excluded patients would have had recurrence; some may have remained disease-free, died from other causes, or been lost to follow-up for non-medical reasons (e.g., moving residence, insurance changes).

By testing the most extreme possible scenario, we establish the upper bound of the selection

bias. If the primary conclusions remain robust under this extreme assumption, they are unlikely to be materially affected under less extreme, more realistic scenarios.

Calculation method for the worst-case scenario:

The primary analysis estimated 45 recurrences among 150 patients, yielding a 48-month cumulative incidence of 30.2% (Fine-Gray estimate, with death as competing risk).

In the worst-case scenario, all 67 excluded patients are assumed to have had recurrence at 3 months.

The total number of recurrence events would therefore be $45 + 67 = 112$.

The denominator for the crude cumulative incidence estimate is the total source cohort of 217 patients (150 included + 67 excluded). However, because the Fine-Gray model accounts for competing risks and variable follow-up, the recalculated 48-month cumulative incidence of ~38.5% is a crude approximation that assumes all 67 events occurred early and patients remained at risk. This represents the maximum plausible upward shift in the recurrence estimate.

The true cumulative incidence, under less extreme assumptions, would lie between 30.2% and 38.5%.

Impact on temporal clustering (peak at 3–12 months):

In the primary cohort, 55.6% (25/45) of recurrences occurred at 3–12 months.

In the worst-case scenario, if all 67 excluded patients recurred at 3 months, the 3–12 month window would contain $67 + 25 = 92$ of 112 total events (82.1%), which would actually strengthen the conclusion that the 3–12 month period is the peak risk window.

Thus, even under the most extreme assumption, the temporal clustering pattern remains not only present but more pronounced.

Impact on risk factor associations (Table 7):

The risk factor associations (TNM stage, preoperative lymph node metastasis, differentiation grade, CEA, etc.) were derived from the primary cohort. The excluded patients lack the necessary covariate and outcome data to assess whether these associations would differ.

However, because the excluded patients are assumed to have had uniformly early recurrence regardless of their baseline characteristics, their inclusion would dilute the observed associations (i.e., attenuate the hazard ratios toward the null) rather than reverse or eliminate them.

Therefore, the direction and statistical significance of the risk factor associations identified in Table 7 are unlikely to be negated by the exclusion. If anything, the primary analysis may slightly overestimate the strength of the associations due to the exclusion of early, uniformly

recurring patients, but the qualitative conclusions remain robust.

Comparison with the non-LDCT detection cohort (Supplementary Table S2):

The 41 patients excluded due to non-LDCT first recurrence (brain MRI or PET/CT) had a significantly shorter median time to recurrence (4.5 months vs. 8.2 months, $P < 0.001$) and higher rates of symptomatic presentation (75.6% vs. 22.2%, $P < 0.001$) compared with the LDCT-detected cohort.

This further supports the concern that the primary cohort systematically underrepresents early, symptomatic, and extrapulmonary recurrences. However, the worst-case scenario analysis presented here demonstrates that even if all excluded patients (including both incomplete follow-up and non-LDCT detection) are assumed to have early recurrence, the temporal clustering pattern and risk factor directionality remain intact.

Interpretation for readers:

The primary analysis provides valid estimates of recurrence patterns and predictors within the LDCT-surveillance-completer population.

The worst-case scenario analysis confirms that while the absolute recurrence rate may be underestimated (by up to ~8 percentage points), the relative comparisons (e.g., which risk factors are most strongly associated with recurrence, and the peak timing of recurrence) are robust to the selection bias.

The cumulative incidence curves in Figure 2 should still be interpreted as applying specifically to patients who survive long enough to complete structured institutional surveillance, not to the entire surgically resected population.

Supplementary Table S12. Sensitivity Analysis Limited to Pathologically Confirmed Recurrence Events (n=31 patients with at least one biopsied event)

Factor	sHR	Bootstrap-Corrected 95% CI	P Value
Smoking History	1.72	1.15 – 2.56	0.011
Tumor Differentiation Degree (Moderately vs. Well)	2.20	1.36 – 3.56	0.016
TNM Stage (III vs. I-II)	2.45	1.38 – 4.35	0.001
CEA (per 10 ng/mL increase)	1.88	1.16 – 3.05	0.006

Pleural/Vascular Invasion	1.60	1.10 – 2.32	0.004
Preoperative Lymph Node Metastasis	2.72	1.44 – 5.14	0.002
Age (per 10-year increase)	1.14	0.91 – 1.42	0.22
Sex (Male vs. Female)	1.10	0.77 – 1.57	0.58
Histology (Squamous vs. Adenocarcinoma)	0.90	0.62 – 1.30	0.58

Model Performance (Optimism-Corrected, 1000 Bootstrap Resamples):

C-index: 0.821 (95% CI: 0.74 – 0.89)

Calibration slope: 0.85

Calibration-in-the-large: -0.03

Brier score: 0.19

Footnotes to Supplementary Table S12:

Purpose: This sensitivity analysis addresses the concern of verification bias (i.e., the possibility that imaging-only diagnosis of recurrence events, particularly for brain and bone metastases, might introduce misclassification). The analysis is restricted to the 31 patients who had at least one pathologically confirmed recurrence event. Patients with recurrence events diagnosed by imaging alone (n=14) were excluded from this sensitivity analysis. The table reports the results of a Fine-Gray competing-risk model fitted to these 31 patients.

Patient selection: Among the 45 patients in the recurrence cohort, 31 (68.9%) had at least one recurrence event confirmed by histopathology (biopsy). The remaining 14 patients (31.1%) had all recurrence events diagnosed by imaging criteria alone. The distribution of patients with and without biopsy confirmation varied by site, as detailed in Supplementary Table S6. Patients with synchronous locoregional and distant recurrence (n=3) who had at least one biopsy-confirmed lesion were included in this sensitivity analysis. Patients with multiple lesions where some were biopsied and some were imaging-only were included in the biopsy-confirmed group.

Interpretation: The results of this sensitivity analysis are consistent with the primary Fine-Gray model (Table 7) that included all 45 recurrence patients. All baseline clinicopathological risk factors that were significant in the primary model remained significant in this biopsy-confirmed analysis, with similar effect sizes and confidence intervals. The C-index (0.821 vs. 0.815), calibration slope (0.85 vs. 0.88), and Brier score (0.19 vs. 0.18) were comparable to the primary model, confirming that the inclusion of

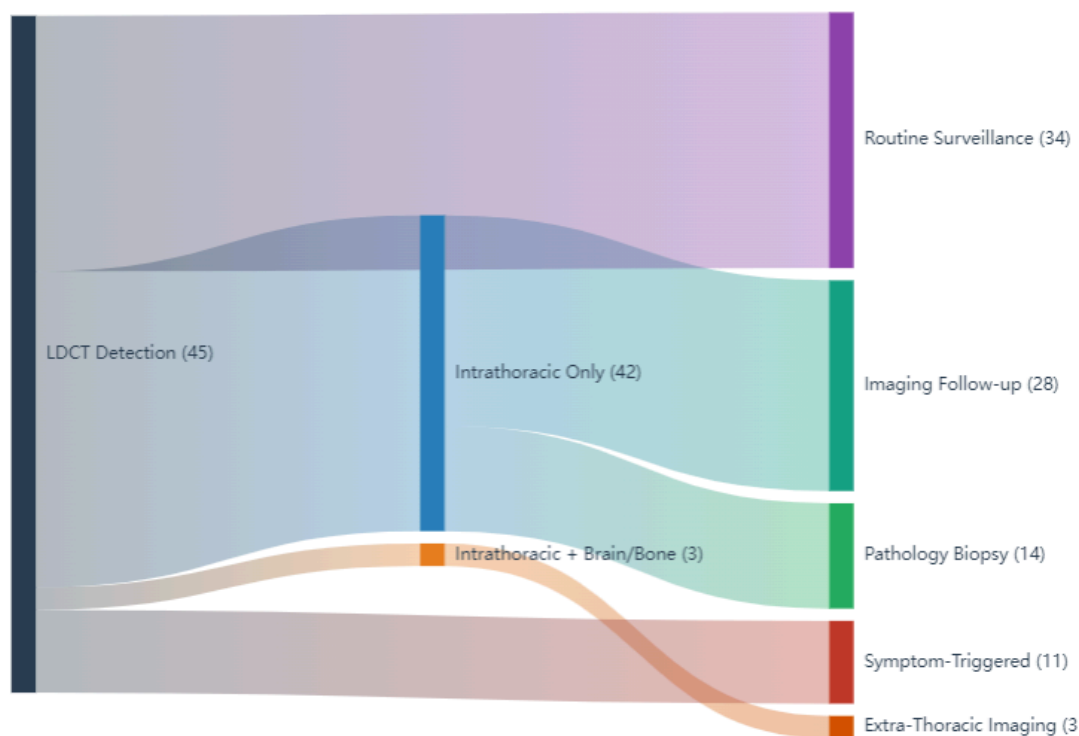
imaging-only diagnosed events did not materially bias the main conclusions.

Comparison with primary model: The consistency between the primary analysis (including all 45 recurrence patients) and this biopsy-confirmed sensitivity analysis (restricted to 31 patients) suggests that verification bias is unlikely to have materially affected the study's main findings. Brain and bone metastases, which were exclusively diagnosed by imaging, were a minority of events (13 of 45 patients, 28.9%) and did not drive the overall results.

Sample size note: This sensitivity analysis is based on 31 patients with biopsy-confirmed recurrence, representing a reduction in sample size of approximately 30% compared with the primary analysis. The reduced sample size results in wider confidence intervals, but the overall conclusions remain unchanged.

Abbreviations: sHR, subdistribution hazard ratio; CI, confidence interval; CEA, carcinoembryonic antigen.

Supplementary Figure S1: Detection Pathway for LDCT-Detected Recurrence Events (n=45)



Note: This Sankey diagram illustrates the detection pathway for all 45 patients in the LDCT-detected recurrence cohort. The flow begins with the index LDCT scan that first detected recurrence (left), followed by anatomic site classification, confirmatory diagnostic modality, and clinical presentation context. The diagram demonstrates that isolated brain or bone metastases were not detected by chest LDCT alone but were included only when a synchronous or preceding intrathoracic abnormality was identified on LDCT, prompting

extra-thoracic imaging workup (MRI, PET/CT, or bone scan). Numbers in parentheses represent patient counts.