ORIGINAL RESEARCH

Outcomes of Positive and Suspicious Findings in Clinical Computed Tomography Lung Cancer Screening and the Road Ahead

Shawn M. Regis[1,2], Andrea Borondy-Kitts[2], Andrea B. McKee[1,2], Kimberly Rieger-Christ[2,3], Jacob Sands[2,4], Jalil Afnan[5], and Brady J. McKee[2,5†]

[1]Division of Radiation Oncology, [3]Division of Translational Research, and [5]Division of Radiology, Lahey Hospital & Medical Center, Burlington, Massachusetts; [2]Rescue Lung Society, Amesbury, Massachusetts; and [4]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts

Abstract

Rationale: Future optimization of computed tomography (CT) lung cancer screening (CTLS) algorithms will depend on clinical outcomes data.

Objectives: To report the outcomes of positive and suspicious findings in a clinical CTLS program.

Methods: We retrospectively reviewed results for patients from our institution undergoing lung cancer screening from January 2012 through December 2018, with follow-up through December 2019. All exams were retrospectively rescored using Lung-RADS v1.1 (LR). Metrics assessed included positive, probably benign, and suspicious exam rates, frequency/nature of care escalation, and lung cancer detection rates after a positive, probably benign, and suspicious exam result and overall. We calculated time required to resolve suspicious exams as malignant or benign. Results were broken down by subcategories, reason for positive/suspicious designation, and screening round.

Results: During the study period 4,301 individuals underwent a total of 10,897 exams. The number of positive (13.9%), suspicious (5.5%), and significant incidental (6.4%) findings was significantly higher at baseline screening. Cancer detection and false-positive rates were 2.0% and 12.3% at baseline versus 1.3% and 5.1% across subsequent screening rounds, respectively. Baseline solid nodule(s) 6 to <8 mm were the only probably benign findings resulting in lung cancer detection within 12 months. New solid nodules 6 to <8 mm were the only LR category 4A (LR4A) findings falling within the LR predicted cancer detection range of 5–15% (12.8%). 38.5% of LR4A cancers were detected within 3 months.

Conclusions: Modification of the definition and suggested workup of positive and suspicious lung cancer screening findings appears warranted.

Keywords: lung; cancer; screening; CTLS; LCS

†Deceased.

In March 2021, the United States Preventive Services Task Force (USPSTF) expanded their lung cancer screening recommendation to include individuals aged 50 to 80 years with at least a 20 pack-year smoking history who currently smoke or quit within the past 15 years (1). This recommendation will nearly double the population eligible for computed tomography (CT) lung cancer screening (CTLS) to 15 million. Currently, fewer than 5% of qualified Americans are enrolled in a CTLS program (2). Achieving a

50% screening rate for the expanded eligible population could increase CTLS volume at many institutions by 2,000% or more. In anticipation of this dramatic increase in screening volume, institutions performing CTLS and developing reporting systems need to provide updated patient and physician education materials and look for strategies to eliminate unnecessary follow-up imaging and precisely target advanced imaging and invasive procedures.

The USPSTF encourages CTLS-eligible individuals to engage in a shared decision-making (SDM) visit with their ordering healthcare provider (1, 3). Many current CTLS SDM decision aids contain National Lung Screening Trial (NLST)-based information; however, the International Early Lung Cancer Action Program and others demonstrated that increasing the NLST positive nodule size threshold safely improves the performance of CTLS (4–9). Current reporting systems and guidelines, such as Lung-RADS v1.1 (LR) and the National Comprehensive Cancer Network (NCCN) Guidelines Lung Cancer Screening, have adopted this larger positive nodule size threshold, resulting in increased positive predictive values (PPVs) and decreased false-positive rates (FPRs) in clinical practice (10, 11).

Additionally, in contrast to the NLST, LR defines a subset of positive findings termed "suspicious," which warrant consideration of invasive diagnostic procedures and advanced imaging (positron emission tomography [PET]/CT). LR classifies these types of findings as category 4 (LR4), with 3 subcategories (A/B/X). According to LR, LR4A exams are "suspicious," with an estimated malignancy rate of 5–15%, whereas LR4B and LR4X exams are "very suspicious" with an estimated malignancy rate >15%. LR recommends follow-up imaging in 3 months for LR4A exams, with PET/CT a consideration. For LR4B and LR4X findings, additional/advanced imaging and/or tissue sampling is recommended. LR distinguishes suspicious findings from other positive findings termed "probably benign" and designated as category 3 (LR3), with an estimated malignancy rate of <2%. The malignancy estimates and management recommendations for LR categories are primarily based on expert consensus and/or data from research trials such as the NLST (12). Going forward, clinical

outcomes data will be needed to optimize the tools/systems necessary for safe and effective lung cancer screening.

We report the nature and outcomes of positive and suspicious CTLS exams using current clinical standards to provide a reference data point/benchmark for CTLS programs, SDM visits/decision aids, and future CTLS guidelines/reporting systems. A novel CTLS reporting system is proposed on the basis of these clinical outcomes data.

Methods

This retrospective single-center study was institutional review board–approved with waiver of informed consent. We assessed consecutive individuals from our institution who underwent clinical CTLS from January 1, 2012 through December 31, 2018 with follow-up through December 31, 2019. All individuals met NCCN lung cancer screening eligibility criteria as previously described (11). All exams were performed on ⩾64-row multidetector CT scanners at 100 kV and 30–100 mA.

Image interpretation was performed by radiologists credentialed in CTLS reporting with LR (13). LR1 and LR2 exams were negative. LR3 and LR4 (A/B/X) exams were positive. LR4 (A/B/X) exams were suspicious. All patients with LR4 findings were referred for consultation with a pulmonologist. The specific steps taken to investigate a suspicious exam result were left to the consulting pulmonologist and patient without a proscribed follow-up algorithm. Exams from January 2012 through July 2014 were reported with the original version of Lung-RADS developed at Lahey Clinic in 2011; exams from August 2014 through December 2018 were reported with Lung-RADS v1.0 (14, 15). All exams were reclassified to Lung-RADS v1.1 (LR) as needed (10). Significant incidental findings (LR Category S) were defined as new/unknown, unexpected findings warranting dedicated clinical/imaging evaluation before the next CTLS exam (13). Exams with infectious/inflammatory findings were reported as LR2i to enable analysis separate from other LR2 exams. Patients with mediastinal/hilar lymph nodes >1 cm and 1.5 cm without other suspicious findings were categorized as LR3 and LR4X, respectively.

The overall lung cancer detection rate (CDR), PPV (LR3/LR4), suspicious (LR4) predictive value (SPV), probably benign (LR3) predictive value (3PV), FPR, false

negatives, and time to cancer diagnosis were analyzed per screening round, with T0 representing the baseline/prevalence exam and T1, T2, T3, and T4+ reflecting the annual/incidence exams. Metrics were assessed at 12 months unless stated otherwise. The next screening round was defined as the exam performed 11+ months (⩾335 d) after a negative exam. Interval studies performed to resolve positives were not included in this analysis, except in determining which positives were false. Combined incidence (T1+) and overall (T0+) results were reported by LR4 category (A/B/X) and eventual diagnostic outcome (benign vs. lung cancer) after a suspicious exam. True positives required lung cancer diagnosis without an intervening negative exam. All other positives were considered false. False negatives occurred when lung cancer was diagnosed fewer than 365 days after a negative exam. PPV, SPV, and 3PV were defined as the percentage of positive, suspicious, and probably benign exams resulting in lung cancer diagnoses, respectively. False positives divided by negative outcomes constituted FPR (16–18).

Benign outcomes after a suspicious exam result were determined by subsequent negative (LR1/LR2) interval exams or tissue sampling with benign histology. Lung cancer was diagnosed by positive pathology or presumed to be present in patients unable to undergo invasive procedures and meeting institutional criteria for lung cancer including nodule growth, PET/CT assessment, and multidisciplinary consensus. Diagnosis date was defined as the date of positive pathology or, for presumed lung cancer, the date of oncology consultation that first documented clinical stage and diagnosis. Treatment date was defined as date of surgery, stereotactic radiotherapy (SBRT), or first day of chemotherapy or combined-modality therapy.

P values for numerical/continuous variables were calculated using analysis of variance; *P* values for categorical variables were determined using the chi-square test except in cases with fewer than five events, for which Fisher exact test was used. Confidence intervals (95% CI) were calculated using the binomial exact method. For all statistical analyses, the significance level for differences was set at $P \leqslant 0.05$.

Results

A total of 4,301 qualified individuals from our institution underwent a total of 10,897

Table 1. Computed tomography lung cancer screening exam statistics by screening round

	Screening Round						
	T0	T1	T2	T3	T4+	Incidence (T1+)	Overall (T0+)
Exams, *n*	4,301	2,682	1,773	1,144	997	6,596	10,897
Positive (LR3/LR4) findings, % (95% CI)	13.9 (12.9–15.0)	6.9	6.1	5.3	5.4	6.2 (5.6–6.8)	9.2 (8.7–9.8)
Suspicious (LR4) findings, % (95% CI)	5.5 (4.8–6.2)	3.6	3.5	3.2	3.2	3.5 (3.0–3.9)	4.2 (3.9–4.6)
Significant incidental findings, % (95% CI)	6.4 (5.7–7.2)	2.7	1.8	1.8	1.9	2.2 (1.9–2.6)	3.9 (3.5–4.2)
Infectious/inflammatory (LR2i) findings, % (95% CI)	7.4 (6.7–8.2)	7.2	6.6	8.4	5.8	7.0 (6.4–7.6)	7.2 (6.7–7.7)
Incomplete (LR0) exams, *n*	2	1	0	1	0	2	4
Cancers @ 12 mo, *n*	85	37	20	13	14	84	169
CDR @ 12 mo, % (95% CI)	2.0 (1.6–2.4)	1.4	1.1	1.1	1.4	1.3 (1.0–1.6)	1.6 (1.3–1.8)
PPV @ 12 mo, % (95% CI)	13.4 (10.7–16.3)	17.8	17.6	21.3	22.2	18.9 (15.2–23.0)	15.6 (13.4–18.0)
SPV* @ 12 mo, % (95% CI)	32.3 (26.4–38.7)	34.0	30.6	35.1	37.5	33.8 (27.7–40.3)	33.0 (28.8–37.5)
False-positive exams at 12 mo, % (95% CI)	12.1 (11.1–13.1)	5.7	5.0	4.2	4.2	5.0 (4.5–5.6)	7.8 (7.3–8.3)
False-positive rate at 12 mo, % (95% CI)	12.3 (11.3–13.3)	5.7	5.1	4.2	4.3	5.1 (4.6–5.6)	7.9 (7.4–8.5)
False-negative exams, *n*	4	4	1	0	2	7	11
Mean time to diagnosis, d	183	144	251	210	252	203	193
Median time to diagnosis, d	82	87	158	99	114	107	87

Definition of abbreviations: CDR = cancer detection rate; CI = confidence interval; LR0 = Lung-RADS v1.1 category 0; LR2i = exams with infectious/inflammatory findings, not currently defined in Lung-RADS v1.1; LR3 = Lung-RADS v1.1 category 3; LR4 = Lung-RADS v1.1 category 4; PPV = positive predictive value; SPV = suspicious predictive value.
*SPV: percentage of suspicious (LR4A/B/X) exams resulting in lung cancer diagnoses.

CTLS exams during the study interval. A total of 2,682 (62.4%) had at least one annual exam, and 1,144 (26.5%) completed at least four annual rounds of screening. Average age was 63.1 years, and average smoking history was 48.5 pack-years. Most individuals were male (55.6%), had quit smoking (52.6%), and met NCCN Group 1 eligibility criteria (79.8%) (11). Average quit duration was 9.7 years (ST1), and average follow-up was 3.1 years.

Positive (13.9%; 95% CI, 12.9–15.0%), suspicious (5.5%; 95% CI, 4.8–6.2%), and significant incidental (6.4%; 95% CI, 5.7–7.2%) findings were significantly higher at the baseline round of screening (Table 1). No significant differences existed among annual (T1+) examinations with two exceptions of doubtful clinical significance: significant incidental findings were slightly higher in the T1 round versus the T2 round (2.7% vs. 1.8%), and infectious/inflammatory findings were slightly higher in the T4+ round versus the T3 round (5.8% vs. 8.4%) (Figure 1). Incomplete (LR0) exams were rare, occurring four times (0.04% overall), as were false-negative exams, occurring 11 times (0.1% overall). Nine false negatives occurred within the LR2i group, yielding a 1.2% lung cancer rate at 12 months (95% CI, 0.5–2.2%), with 2/9 (22.2%) small cell lung cancer (SCLC). The overall CDR was 2.0% at baseline (95% CI, 1.6–2.4%) and averaged 1.3% in the annual screening rounds (95% CI, 1.0–1.6%), with 30/199 (15.1%) cancers presumed to be present and treated with SBRT without histologic sampling. The FPR was 12.3% at baseline (95% CI, 11.3–13.3%) and dropped to an average of 5.1% annually thereafter (95% CI, 4.6–5.6%).

Baseline, incidence, and overall PPVs were 13.4% (95% CI, 10.7–16.3%), 18.9% (95% CI, 15.2–23.0%), and 15.6% (95% CI, 13.4–18.0%). The overall 3PV was 0.7% (95% CI, 0.2–1.9%) with only one LR3 finding (solid nodule[s] 6 to <8 mm at baseline) resulting in lung cancer detected within 12 months. The overall SPV was 33.0% (95% CI, 28.8–37.5%) with subcategory SPVs of LR4A 21.2% (95% CI, 15.9–27.4%), LR4B 54.2% (95% CI, 46.3–61.9%), and LR4X 20.5% (95% CI, 12.4–30.8%) (Table 2). Only
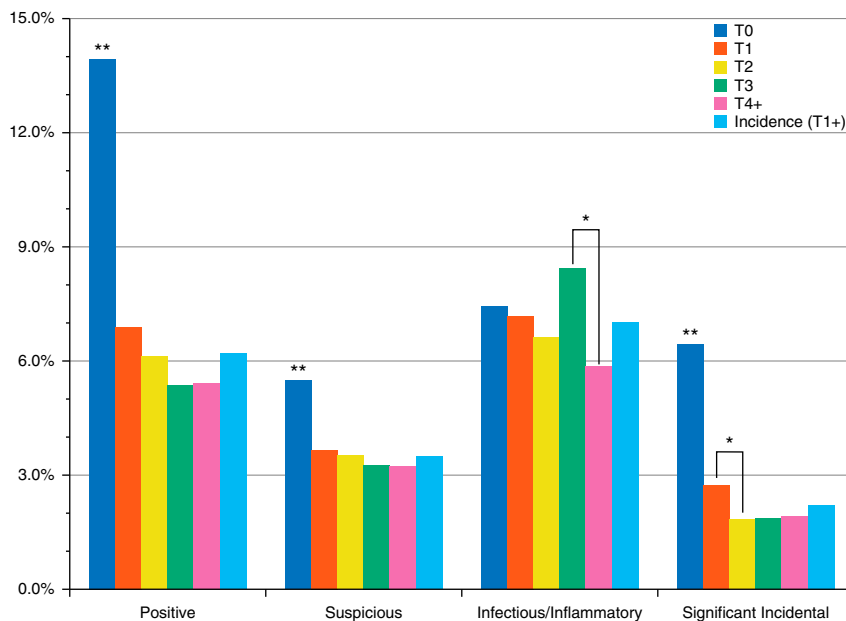
Figure 1. Distribution of computed tomography lung cancer screening exam results by screening round. *Significantly different from one other round. **Significantly different from all other rounds.

Table 2. Suspicious predictive values for Lung-RADS v1.1 category 4 exams

Screening Round	LR4A		LR4B		LR4X		All LR4	
	Lung Cancers	SPV*	Lung Cancers	SPV	Lung Cancers	SPV	Lung Cancers	SPV
Baseline (T0)	25	22.1%	44	57.1%	7	15.6%	76	32.3%
Annual (T1+)	20	20.2%	47	51.7%	10	26.3%	77	33.8%
Overall (T0+)	45	21.2%	91	54.2%	17	20.5%	153	33.0%

Definition of abbreviations: LR4 = Lung-RADS v1.1 category 4; SPV = suspicious predictive value.
*SPV: percentage of suspicious (LR4A/B/X) exams resulting in lung cancer diagnoses.

one LR4A finding (new solid nodule 6 to <8 mm) fell within the LR4A predicted SPV range of 5–15% (12.8%; 95% CI, 4.3–27.4%), with the remaining LR4A findings SPVs ranging from 16.7% (endobronchial nodule; 95% CI, 0.4–64.1%) to 30.0% (part-solid nodule[s] ≥6 mm with solid component 6 to <8 mm; 95% CI, 6.7–65.2%). All LR4B findings had SPVs significantly exceeding 15%, ranging from a minimum of 46.2% (growing solid nodule ≥8 mm; 95% CI, 30.1–62.8%) to a maximum of 75.0% (part-solid nodule with a solid component ≥8 mm; 95% CI, 34.9–96.8%). Lymph nodes with minimum dimension ≥15 mm had a below-expected SPV of 5% (95% CI, 0.1–24.9%). Excluding these lymph nodes increased the LR4X SPV to 25.4% (95% CI, 15.3–37.9%).

A total of 62.8% and 93.3% of lung cancers detected after a suspicious exam were diagnosed within 3 months and 12 months, respectively (LR4A, 38.5% and 86.5%; LR4B, 76.6% and 96.8%; LR4X, 61.1% and 94.4%) (Figure 2). Overall, 65.2% of lung cancers were diagnosed within 6 months and 84.8% within 1 year of the CTLS exam (ST2). The overall (T0+) mean and median times to lung cancer diagnosis after a suspicious exam were 193 and 87 days, respectively.

The rate of PET/CT and invasive procedures performed to evaluate a suspicious exam before cancer diagnosis was significantly lower among individuals eventually determined to have benign findings (41.4% and 13.4%) versus those diagnosed with cancer (79.9% and 90.9%) (ST3). The overall rate that an individual without lung cancer in the study population underwent an invasive procedure was 0.9%.

LR4X patients diagnosed with cancer had the highest percentage of SCLC (16.7%) and squamous cell (27.8%) histology (ST4) and largest average cancer size (2.6 cm) (ST5). Overall survival for cancers in the LR4X group was inferior to those found in both the LR4A and LR4B groups (SF1).

Discussion

Clinical versus NLST Outcomes
Revisions in the definition of a positive CTLS exam since publication of the NLST in 2011 and the overall higher CDR observed in clinical CTLS practice have rendered certain NLST metrics, such as the FPR and PPV, outdated as far as informing eligible individuals considering CTLS enrollment what to realistically expect. Using LR, we found an overall (T0+) CTLS FPR of 7.9%, compared with 23.5% in the NLST. Accordingly, the PPVs of baseline (13.4%) and incidence (T1+) exams (18.9%) were significantly higher than the NLST (3.8% and 3.5%, respectively) (3). The rate of significant incidental findings (category S) was similarly lower for our study population than the NLST both at baseline (6.4% vs. 10.2%) and at annual follow-up (2.9% vs. 7.5%). Our results suggest that sole reliance on the NLST as a source of information for SDM aides could depress CTLS enrollment by underestimating the net benefit of CTLS in current clinical practice.

Specialist Consult for Suspicious Findings
Despite the high rate of malignancy among individuals with suspicious (LR4) exams (33%), we found that advanced imaging and invasive procedures were largely restricted to those ultimately diagnosed with lung cancer, with only 13.4% and 41.8% of individuals with a suspicious (LR4) exam eventually determined to have benign findings undergoing an invasive procedure or PET/CT, respectively. This resulted in an overall rate of invasive procedures and PET/CT among all individuals without lung cancer of 0.9% and 1.8%, respectively. The impact of directing individuals with suspicious exam findings to undergo a specialist consultation (in our case with pulmonary medicine) in lieu of recommending additional imaging

such as PET/CT may play an important role in limiting unwarranted workup and procedures and deserves further study (ST3).

Compiling CTLS Program Metrics: Follow-Up Time
The mean time to lung cancer diagnosis after a suspicious exam in our study was more than 6 months (193 d); however, the median time to diagnosis is closer to 3 months (87 d) (Table 1). This difference primarily results from a small number of cases that required an extended period of time to resolve because of challenging lesions or nonadherent patients lost to follow-up who subsequently represented for evaluation. We observed an even longer period of time to establish a benign diagnosis after a suspicious exam, with a mean range of 150–259 days and a median range of 101–208 days. These numbers help inform the decision about when certain quality metrics, including PPV, SPV, and CDR, can be meaningfully computed. ST2 reviews how these metrics change with lengthening follow-up times of 6, 9, and 12 months. Overall, 35% of lung cancers diagnosed after a suspicious exam in our study required more than 6 months to diagnose, and up to 15% were diagnosed after 12 months. CTLS programs should consider allowing 12 months of follow-up to capture the majority of detected lung cancer when compiling quality assurance metrics with an understanding that these metrics will slightly underestimate actual cancer prevalence/detection. However, follow-up time beyond a year would interfere with timely quality assurance review and may be difficult to interpret, given our observed 1.3% incidence rate of lung cancer and mandate for individuals with negative exams to return for repeat imaging in 12 months.

Cancer Outcome among Suspicious Findings
Although individuals with LR4X findings are less likely to be diagnosed with cancer than
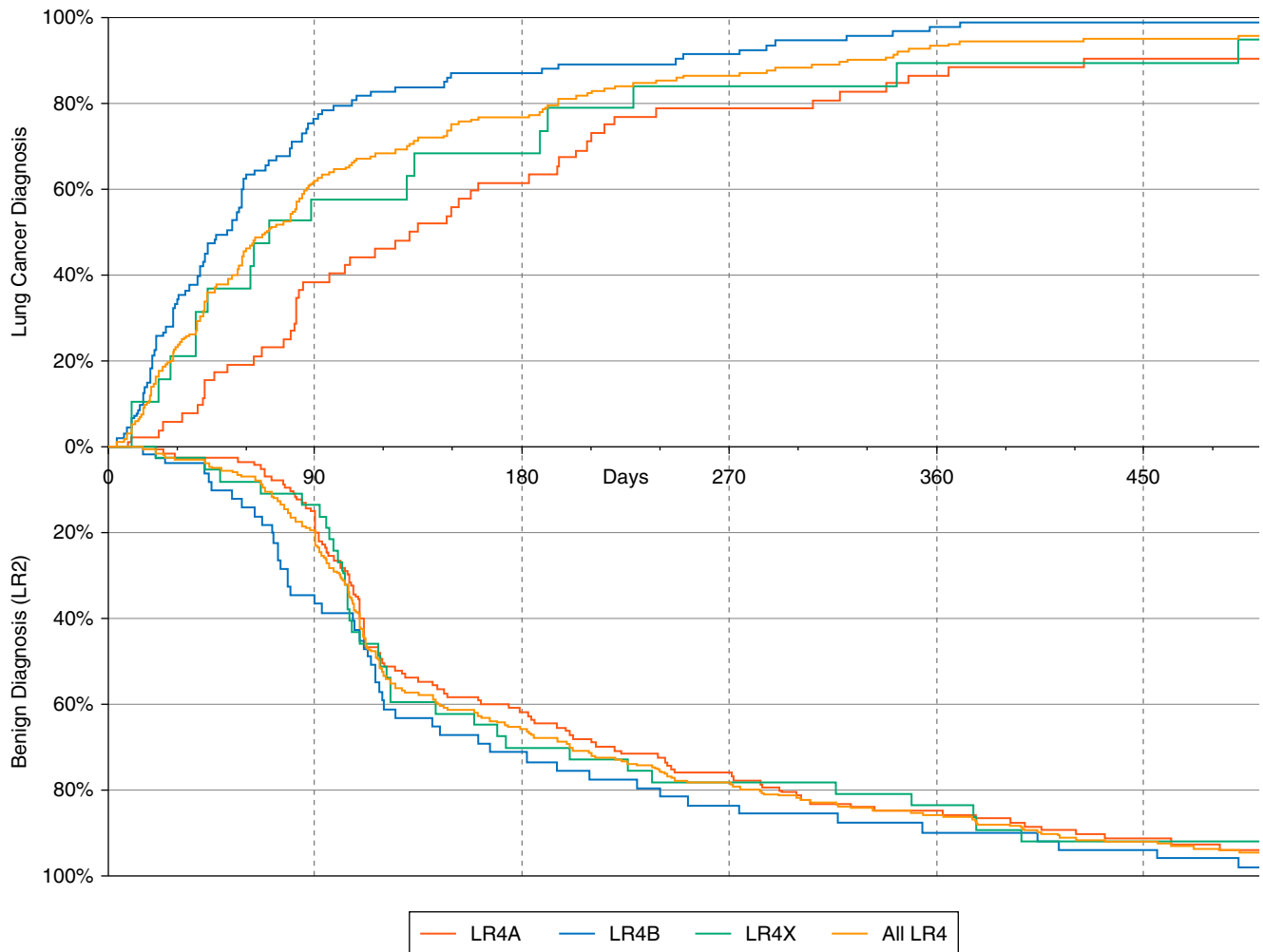
Figure 2. Overall time from suspicious (Lung-RADS v1.1 category 4 [LR4]) exam to lung cancer (top) or benign (bottom) diagnosis.

those with LR4B findings (SPV 20.5% vs. 54.2%), when cancer is diagnosed the LR4X group has the poorest survival. This is likely due to a higher rate of aggressive histologies (squamous and SCLC) in the LR4X group (ST4). Presumably, these histologies more commonly present with "additional suspicious features" not defined by LR4A or LR4B. However, a host of benign conditions can also display "additional suspicious features," lowering the SPV of LR4X versus LR4B.

LR Discrepancies/Gaps
Only one of the LR3 category findings resulted in lung cancer detected within 12 months despite standard follow-up imaging obtained at 6 months. Reclassification of the other LR3 category findings as LR2 may reduce unnecessary interval imaging (Table 3). All but one of the LR4A category findings had a higher SPV

than the LR predicted rate of 5–15%, with nearly 40% of LR4A lung cancers in our cohort diagnosed within 3 months. This suggests a standard 3-month follow-up recommendation for LR4A may be too liberal (Table 3). The SPV of LR4X cases (excluding lymph nodes > 15 mm) of 24.5% more closely approximates the LR4A rate of 21.2% than the LR4B rate of 54.2%, questioning LR descriptively differentiating LR4X (very suspicious) from LR4A (suspicious). The fact that LR4A, LR4X, and LR4B all have SPVs in excess of 15% would appear to obviate the need for any subcategorization of LR4.

How to classify findings characteristic of infection/inflammation and enlarged lymph nodes in the lung cancer stations (mediastinal/hilar) is not sufficiently addressed by LR. Despite our institutional policy of delaying CTLS for 12 weeks after resolution of symptoms and/or treatment of

pulmonary infection and multiple reminders of the need to be asymptomatic before imaging, we observed a remarkably consistent LR2i rate of 6–8% across all rounds of screening, similar to other clinical reports (19). These findings are more common than both probably benign and suspicious findings. LR2i most commonly reflects multifocal areas of tree-in-bud nodularity that resolve on follow-up imaging or wax and wane over multiple exams. Although there were nine false-negative cancers within the LR2i group, this represents an overall malignancy rate of just over 1%. At 12 months we found no lung cancer among individuals with mediastinal/hilar lymph nodes between 10 and 15 mm in the absence of other positive findings. For mediastinal/hilar lymph nodes 15 mm or greater in short axis, the malignancy rate of 5% was the lowest rate of all individual suspicious findings assessed. Finally, no

Table 3. Observed versus estimated lung cancer detection rates at 12 months for individual computed tomography lung cancer screening findings (all scans, T0+)

LR Category		Category Findings	Cases (n)	Prevalence (%)	Lung Cancer at 12 mo (n)	Observed (%)	95% CI (%)	Estimated by LR (%)	Observed vs. Estimated
Negative	1	No lung nodules, nodule(s) w/fat or characteristically benign calcifications	2,784	25.5	1	<0.1	0–0.2	<1	√
Benign	2	Solid nodule(s) < 6 mm at baseline, solid nodule(s) < 4 mm new, part-solid nodule(s) < 6 mm total diameter at baseline, nonsolid nodule(s) (GGN) < 30 mm at baseline, nonsolid nodule(s) (GGN) slowly growing, nodules unchanged for >3 mo	6,320	58.0	1	<0.1	0–0.09	<1	√
	2i	Findings characteristic of infection/inflammation*	782	7.2	9	1.2	0.5–2.2	N/A	N/A
		Total LR2	7,102	65.2	10	0.1	0.07–0.3	<1	√
Probably benign	3	Lymph nodes 10–15 mm without pulmonary nodules*	92	0.8	0	0.0	0–3.9	N/A	N/A
		Solid nodule(s) ≥ 6 to <8 mm at baseline	203	1.9	4	2.0	0.5–5.0	1–2	√
		New solid nodule(s) 4 to <6 mm	145	1.3	0	0.0	0–2.5		↓
		Part solid nodule(s) ≥ 6 mm total diameter w/solid component < 6 mm	47	0.4	0	0.0	0–7.6		↓
		New part-solid nodule(s) < 6 mm total diameter	35	0.3	0	0.0	0–10.0		↓
		Nonsolid nodule(s) (GGN) ≥ 30 mm at baseline CT	1	0.0	0	0.0	0–97.5†		↓
		New GGN	21	0.2	0	0.0	0–16.1		↓
		Total LR3	544	5.0	4	0.7	0.2–1.9		↓
Suspicious	4A	Solid nodule(s) ≥ 8 to <15 mm at baseline	80	0.7	20	25.0	16.0–35.9	5–15	↑
		Growing nodule(s) < 8 mm	23	0.2	4	17.4	5.0–38.8		↑
		New solid nodule(s) 6 to <8 mm	39	0.4	5	12.8	4.3–27.4		√
		Part solid nodule(s) ≥ 6 mm w/solid component ≥ 6 to <8 mm	10	0.1	3	30.0	6.7–65.2		↑
		Part solid nodule(s) w/new or growing <4-mm solid component	54	0.5	12	22.2	12.0–36.0		↑
		Endobronchial nodule(s)	6	0.1	1	16.7	0.4–64.1		↑
		Total LR4A	212	1.9	45	21.2	15.9–27.4		↑
Very suspicious	4B	Solid nodule(s) ≥ 15 mm at baseline	35	0.3	25	71.4	53.7–85.4	>15	√
		New solid nodule(s) ≥ 8 mm	62	0.6	30	48.4	35.5–61.4		√
		Growing solid nodule(s) ≥ 8 mm	39	0.4	18	46.2	30.1–62.8		√
		Part-solid nodule(s) with solid component ≥ 8 mm	8	0.1	6	75.0	34.9–96.8		√
		Part-solid nodule(s) w/new or growing ≥ 4 mm solid component	24	0.2	12	50.0	29.1–70.9		√
		Total LR4B	168	1.5	91	54.2	46.3–61.9		√
	4X	Lymph nodes ≥ 15 mm without pulmonary nodules*	20	0.2	1	5.0	0.1–24.9	N/A	N/A
		LR 3 or 4 nodules w/findings increasing cancer risk	63	0.6	16	25.4	15.3–37.9	>15	√
		Total LR4X	83	0.8	17	20.5	12.4–30.8		√

Definition of abbreviations: CI = confidence interval; CT = computed tomography; GGN = ground-glass nodules; LR = Lung-RADS v1.1; N/A = not applicable.
*Not currently part of LR.
†Unreliable CI.

nonsolid (ground glass) nodules regardless of size, new presentation, or growth resulted in lung cancer detection within 12 months.

Proposed Reporting System
A proposed reporting system, the RLS System, which accounts for these clinical observations and addresses LR gaps, is provided (Figure 3). The RLS System recommendations for follow-up imaging are based on the clinically observed cancer diagnosis rates within the updated categories. As compared with LR, some

findings would not require interval imaging, such as new ground-glass nodules. Other findings would prompt specialist evaluation rather than confirmation imaging at 3 months, such as a part-solid nodule with a >6 mm solid component due to a higher likelihood of cancer than predicted by LR. Based on this single-center data set, the proposed RLS System would reduce the overall number of scans performed in a lung cancer screening program, while prompting more rapid workup of those most likely to

have a new lung cancer. This has the potential to improve both patient outcomes and cost effectiveness.

Limitations
The generalizability of our results is limited, as our data are drawn from a single institution screening a relatively homogeneous patient population. Fifteen percent of cancers diagnosed included those presumed to be present in patients deemed medically unsafe for biopsy and empirically treated with SBRT (20).

RLS Category[Δ]		Findings			Prevalence[z]			Lung Cancer @ 12 Months	Recommendation
Description	Code				Baseline	Annual	Overall		
Incomplete	0	• Motion/other artifact	• Partial imaging of lungs	• Insufficient dose/Noise	< 0.1%	< 0.1%	< 0.1%	N/A	Repeat CTLS
Negative	1	• No findings	• Definitively benign nodule(s)[a]	• Lymph node(s) < 10 mm[b]	17.9%	16.2%	16.9%	< 0.1%	Annual CTLS
Benign	2	• Solid nodule(s) < 6mm[cg]	• Part-solid nodule(s) w/solid component < 6 mm[cd]	• Lymph node(s) 10 to < 15 mm[b] • Non-solid nodule(s)[h]	64.5%	73.2%	69.7%	< 0.1%	Annual CTLS
		• Nodule(s) or lymph node(s) stable for ≥ 3 months and no growth versus all prior exams[be]							
	2i	• Finding(s) characteristic of infection/inflammation (ie multifocal tree-in-bud nodularity)[fy]			7.4%	7.1%	7.2%	1%	3–12 month CTLS[y] (Antibiotics if appropriate)
Probably Benign	3	• Solid nodule(s) @ baseline 6 to < 8 mm[c]			4.7%	< 0.1%[k]	1.9%	2%	6 month CTLS[ci]
Suspicious	4	All other nodules, lymph nodes, and other findings concerning for lung cancer:[k]			5.5%	3.5%	4.4%	≥5%[x]	Specialist consult
		• Solid nodule(s)[c] ○ Baseline ≥ 8 mm[p] ○ New ≥ 6 mm[mq] ○ Growing ≥ 4 mm[ejr§]	• Part-solid nodule(s)[c] ○ Solid component ≥ 6 mm[gs] ○ New/growing solid component[t]	• Lymph node(s) ≥ 15 mm[blu] • Endobronchial nodule(s)[iv] • RLS 2/3 nodule(s) w/features increasing lung cancer risk[w]					
Lung Cancer	5A[no]	Presumed[o]			< 0.3%	< 0.3%	< 0.3%	~100%	Treatment (Non-surgical)
	5B[n]	Pathology proven			2.1%	1.2%	1.6%	100%	Treatment
Significant Incidental	S[†]	New or unknown, unexpected findings warranting dedicated clinical or imaging evaluation[‡]			6.4%	2.2%	3.9%	<0.1%	Custom

[a]Nodules w/fat or specific patterns of calcification (diffuse, central, laminar, concentric, popcorn)
[b]Mediastinal and hilar stations measured on short axis
[c]All nodules measured as mean diameter and reported in interval 1mm ranges (ie 4–5mm, 9–10mm etc) to respect limited spatial resolution of CTLS
[d]Baseline
[e]Growth defined as an increase in mean diameter of at least 1.5mm
[f]12 weeks waiting period prior to CTLS after resolution of symptoms/treatment of infection
[g]Baseline or new
[h]Baseline, new, or growing
[i]Repeat CTLS after vigorous coughing a consideration
[j]Includes nodules w/recent interval stability but growth in comparison to older exams
[k]For nodules which appear enlarged but do not meet criteria for growth keep RLS Category and follow-up interval the same.
[l]Followup CTLS in 3 months a consideration after specialist consultation
[m]Followup CTLS in 3 months for new solid nodules < 8mm a consideration after specialist consultation
[n]Patient level database recording only – not for exam assessment
[o]Nodule growth, PET/CT obtained, & multidisciplinary consensus finding represents lung cancer

Suspicious predictive value (SPV): percentage of suspicious (RLS 4) exams resulting in lung cancer diagnoses
[p]SPV @ 12 Months 39.1%, n = 115 (95% CI 30.2–48.7%)
[q]SPV @ 12 Months 34.7%, n = 101 (95% CI 25.5–44.8%)
[r]SPV @ 12 Months 35.5%, n = 62 (95% CI 23.7–48.7%)
[s]SPV @ 12 Months 50.0%, n = 18 (95% CI 26.0–74.0%)
[t]SPV @ 12 Months 30.8%, n = 78 (95% CI 20.8–42.2%)
[u]SPV @ 12 Months 5.0%, n = 20 (95% CI 0.1–24.9%)
[v]SPV @ 12 Months 16.7%, n = 6 (95% CI 0.4–64.1%)
[w]SPV @ 12 Months 25.4%, n = 63 (95% CI 15.3–37.9%)
[x]SPV @ 12 Months 5–50%, average 33%
[y]TBD followup by nature of findings: 3 months for new and 6–12 months for chronic, waxing & waning findings
[z]Baseline n = 4301, Annual n = 6596, Overall n = 10897
[§]Micronodules (nodules <4mm) baseline, new or growing are RLS category 2
[†]Append "S" to RLS category number if significant incidental finding(s) present
[‡]Expected findings in the CTLS population including coronary artery calcifications and emphysema not included
[ci]If 6 Month CTLS = RLS category 2, schedule next Annual CTLS 12 months after date of the 6 Month CTLS exam
[Δ]For quality metrics negative = RLS 1, 2, & 2i, positive = RLS 3 & 4, suspicious = RLS 4, cancers diagnosed = RLS 5A, 5B

Figure 3. Proposed reporting system. CI = confidence interval; CTLS = computed tomography lung cancer screening; PET/CT = positron emission tomography/computed tomography; TBD = to be determined.

Although this raises the potential for overestimating lung cancer prevalence, more than 30% of these patients subsequently experienced pathology-proven regional/distant lung cancer recurrence, a rate exceeding that observed in patients with pathology-proven lung cancer at initial diagnosis (21). However, low numbers of patients treated with SBRT without biopsy preclude assessment of survival differences at this time. Although all patients with suspicious findings were recommended to undergo consultation with a pulmonologist, the next steps in each patient's care were customized on the basis of exam findings, physician experience, patient preference, patient performance level, and a variety of other factors rather than a preset algorithm. This limits the generalizability of our results to institutions without access to lung cancer specialists. Finally, because of the low positive rate of CTLS, particularly during annual rounds of screening, a larger number of exams would be useful to confirm the prevalence and malignancy rates of each LR finding within the probably benign and suspicious categories.

Conclusions

The outcomes of positive and suspicious CTLS findings in clinical practice differ meaningfully from those reported by the NLST and predicted by current CTLS reporting systems. To account for these differences and in anticipation of an oncoming USPSTF eligibility expansion-driven increase in CTLS volume, we propose a novel CTLS reporting system designed to improve clinical outcomes by accelerating the workup of those most likely to have cancer while reducing the overall number of scans performed. A multi-institutional validation study is planned. ∎

References

1 U.S. Preventive Services Task Force. Lung cancer: screening. 2021 [accessed 2021 Jun 20]. Available from: https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/lung-cancer-screening.

2 Fedewa SA, Kazerooni EA, Studts JL, Smith R, Bandi P, Sauer AG, *et al.* State variation in low-dose CT scanning for lung cancer screening in the United States. *J Natl Cancer Inst* 2021;113:1044–1052.

3 Moyer VA; U.S. Preventive Services Task Force. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2014;160:330–338.

4 Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, *et al.*; National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365:395–409.

5 de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, *et al.* Reduced lung-cancer mortality with volume CT screening in a randomized trial. *N Engl J Med* 2020;382:503–513.

6 Pastorino U, Silva M, Sestini S, Sabia F, Boeri M, Cantarutti A, *et al.* Prolonged lung cancer screening reduced 10-year mortality in the MILD trial: new confirmation of lung cancer screening efficacy. *Ann Oncol* 2019;30:1162–1169.

7 McKee BJ, Regis SM, McKee AB, Flacke S, Wald C. Performance of ACR Lung-RADS in a clinical CT lung screening program. *J Am Coll Radiol* 2015;12:273–276.

8 Pinsky PF, Gierada DS, Black W, Munden R, Nath H, Aberle D, *et al.* Performance of Lung-RADS in the National Lung Screening Trial: a retrospective assessment. *Ann Intern Med* 2015;162:485–491.

9 Yip R, Henschke CI, Yankelevitz DF, Smith JP. CT screening for lung cancer: alternative definitions of positive test result based on the national lung screening trial and international early lung cancer action program databases. *Radiology* 2014;273:591–596.

10 Lung CT Screening Reporting & Data System. Lung-RADS Version 1.1. American College of Radiology; 2019 [accessed 2021 Jun 20]. Available from: https://www.acr.org/-/media/ACR/Files/RADS/Lung-RADS/LungRADSAssessmentCategoriesv1-1.pdf?la=en.

11 Wood DE, Kazerooni EA, Baum SL, Eapen GA, Ettinger DS, Hou L, *et al.* Lung Cancer Screening, Version 3.2018, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2018;16:412–441.

12 Chelala L, Hossain R, Kazerooni EA, Christensen JD, Dyer DS, White CS. Lung-RADS Version 1.1: challenges and a look ahead, from the *AJR* Special Series on Radiology Reporting and Data Systems. *AJR Am J Roentgenol* 2021;216:1411–1422.

13 McKee BJ, McKee AB, Kitts AB, Regis SM, Wald C. Low-dose computed tomography screening for lung cancer in a clinical setting: essential elements of a screening program. *J Thorac Imaging* 2015;30:115–129.

14 McKee BJ, McKee AB, Flacke S, Lamb CR, Hesketh PJ, Wald C. Initial experience with a free, high-volume, low-dose CT lung cancer screening program. *J Am Coll Radiol* 2013;10:586–592.

15 Lung-RADS Version 1.0 Assessment Categories. 2014 [accessed 2021 Jun 20]. Available from: https://www.acr.org/-/media/ACR/Files/RADS/Lung-RADS/Lung-RADS_AssessmentCategories.pdf.

16 Sands J, Tammemägi MC, Couraud S, Baldwin DR, Borondy-Kitts A, Yankelevitz D, *et al.* Lung screening benefits and challenges: a review of the data and outline for implementation. *J Thorac Oncol* 2021;16:37–53.

17 Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford Statistical Science Series #31. Oxford: Oxford University Press; 2003.

18 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995;57:289–300.

19 Mendoza DP, Chintanapakdee W, Zhang EW, Gilman MD, Lennes IT, Frank AJ, *et al.* Management and outcomes of suspected infectious and inflammatory lung abnormalities identified on lung cancer screening CT. *AJR Am J Roentgenol* 2021;217:1083–1092.

20 Wegner RE, Ahmed N, Hasan S, Schumacher LY, Van Deusen M, Colonias A. SBRT for early stage lung cancer: outcomes from biopsy-proven and empirically treated lesions. *Lung Cancer Manag* 2018;7:LMT01.

21 Ho H, Williamson C, Regis SM, Stock CT, Quadri SM, McKee BJ, *et al.* Surgery and invasive diagnostic procedures for benign disease are rare in a large low-dose computed tomography lung cancer screening program. *J Thorac Cardiovasc Surg* 2021;161:790–802.e2.