Standardized Reporting and Management of Suspicious Findings on Chest CT Imaging Is Associated With Improved Lung Cancer Diagnosis in an Observational Study

Thomas H. Urbania, MD; Jennifer R. Dusendang, MPH; Lisa J. Herrinton, PhD; Stacey Alexeeff, PhD; Douglas A. Corley, MD, PhD, MPH; Sora Ely, MD; Ashish Patel, MD; Todd Osinski, MD; and Lori C. Sakoda, PhD, MPH

BACKGROUND: Follow-up of chest CT scan findings suspicious for lung cancer may be delayed because of inadequate documentation. Standardized reporting and follow-up may reduce time to diagnosis and care for lung cancer.

STUDY DESIGN AND METHODS: We implemented a reporting system that standardizes tagging of chest CT scan reports by classifying pulmonary findings. The system also automates referral of patients with findings suspicious for lung cancer to a multidisciplinary care team for rapid review and follow-up. The system was designed to reduce the time to diagnosis, particularly for early-stage lung cancer. We evaluated the effectiveness of this system, using a quasi-experimental stepped wedge cluster design, examining 99,148 patients who underwent diagnostic (nonscreening) chest CT imaging from 2015 to 2017 and who had not received a chest CT scan in the preceding 24 months. We evaluated the association of the intervention with the incidence of diagnosis and surgical treatment of early-stage (I, II) and late-stage (III, IV) lung cancer within 120 days of chest CT imaging.

RESULTS: Forty percent of patients received the intervention. Among 2,856 patients (2.9%) who received diagnoses of lung cancer, 28% had early-stage disease. In multivariable analyses, the intervention was associated with 24% greater odds of early-stage diagnosis (OR, 1.24; 95% CI, 1.09-1.41) and no change in the odds of late-stage diagnosis (OR, 1.04; 95% CI, 0.95-1.14). The intervention was not associated with the rate of surgical treatment within 120 days.

INTERPRETATION: In this large quasi-experimental community-based observational study, implementation of a system that combines standardized tagging of chest CT scan reports with clinical navigation was effective for increasing the diagnosis of early-stage lung cancer.

CHEST 2020; 158(5):2211-2220

KEY WORDS: CT imaging; diagnostic accuracy; lung cancer

FOR EDITORIAL COMMENT, SEE PAGE 1824

ABBREVIATIONS: CPT = current procedural terminology; HR = hazard ratio; ICD = International Classification of Diseases

AFFILIATIONS: From the Department of Radiology (Drs Urbania and Osinski), the Department of Thoracic Surgery (Drs Ely and Patel), and the Division of Research (Ms Dusendang and Drs Herrinton, Alexeeff, Corley, and Sakoda), Kaiser Permanente Northern California, Oakland, CA.

FUNDING/SUPPORT: This project is funded by the Permanente Medical Group Delivery Science and Applied Research initiative, and in part by K07 CA188142 (to L. C. S.).

CORRESPONDENCE TO: Lisa J. Herrinton, PhD, Division of Research, Kaiser Permanente Northern California, 2000 Broadway, Oakland, CA 94612; e-mail: lisa.herrinton@kp.org

Copyright © 2020 American College of Chest Physicians. Published by Elsevier Inc. All rights reserved.

DOI: <https://doi.org/10.1016/j.chest.2020.05.595>

Lung cancer is most often diagnosed at a late stage, and its 5-year relative survival is only 19.4%.¹ However, when it is diagnosed at an early stage, 5-year relative survival exceeds 50%.¹ With the rising use of chest CT imaging,² it is increasingly important to evaluate and manage pulmonary findings appropriately to minimize unnecessary diagnostic workups while preventing delays in lung cancer diagnosis and treatment.

When radiologic findings are suspicious for lung cancer, effective communication is critical to ensure rapid follow-up. However, the ordering provider may not accurately comprehend the radiologist's narrative report. Demands on provider attention and systemic factors may also deter timely follow-up.³ One strategy for improvement is to standardize reporting of radiologic findings and explicitly link findings to clinical recommendations. This strategy has been used successfully for screening mammography.⁴ However, compared with mammography, standardized reporting of pulmonary findings on chest CT imaging is more

challenging to implement because of its wider range of indications and findings.

In 2015, Kaiser Permanente Northern California began implementing a reporting system to improve population management of pulmonary findings on diagnostic chest CT imaging consistent with Fleischner Society guidelines.⁵ This system standardizes radiologist reporting of pulmonary findings, using tags that classify the findings. Findings tagged as suspicious for lung cancer are automatically forwarded to a care coordinator and multidisciplinary committee of physicians who meet weekly to review the imaging and coordinate care with the patient's primary care team. We hypothesized that the intervention would result in increased diagnosis and surgery for lung cancer within 120 days of chest CT imaging. Because of the program's focus on speeding management of early-stage lung cancer, we hypothesized a priori that these increases would be more pronounced for early-stage disease.

Methods

Kaiser Permanente provides integrated, comprehensive health care to 4.4 million patients at 21 medical centers in northern California. All care is documented in an electronic medical record (Epic Systems). Radiologists use in-house dictation software that integrates a third-party voice recognition engine with the electronic medical record.

In 2015, the health-care organization launched a broad initiative to regionalize oncology services and support diagnostic workflows. As part of this initiative, an intervention was implemented across radiology centers on a rolling basis. Interventions were also implemented in pulmonology, thoracic surgery, and oncology. In radiology, a multidisciplinary team including radiologists, pulmonologists, and primary care physicians developed an informatics-based reporting system that uses templates and smart-text (dot phrases) to standardize documentation of chest CT scan findings. The system assigns standardized tags to chest CT scan reports, using the prefix "#PUL" to classify pulmonary findings into one of eight categories (Table 1). The dictation software enforces a "hard-stop" requirement for all examination types that include a full chest CT scan. Thus, the radiologist cannot finalize the report unless a #PUL tag or a smart-text phrase such as "No pulmonary nodule" is added to the report. The #PUL5 tag is assigned to reports with findings suspicious for lung cancer. Before the intervention, chest CT scan reports included a free-text summary of findings and impressions. Templates containing Fleischner Society guidelines⁵⁻⁷ were sometimes added to the radiology report, but their use was not forced, and the ordering physician was responsible for follow-up care.

To increase care coordination for patients tagged with #PUL5 (suspicious for lung cancer), algorithms have been written into the electronic medical record to automatically identify these patients and forward them to a worklist that is monitored daily by a care coordinator in the pulmonology department. The coordinator forwards new cases to a daily triage physician, typically a pulmonologist or thoracic surgeon, who uses about 1 h/d to complete

this task. The triage physician performs an initial review. For straightforward cases, the triage physician provides initial management recommendations. More complicated cases are discussed at a weekly multidisciplinary conference that lasts 1 to 2 h, depending on case volume. The multidisciplinary conference includes the care coordinator, pulmonologist, thoracic surgeon, and often a radiologist and oncologist. For some patients, the team takes over management of the case, ordering follow-up imaging studies, arranging biopsy, and scheduling surgery, while for others the team makes recommendations to the primary care physician, with the care coordinator following up to ensure that the plan is implemented. The physicians who are accountable for triage do not use fixed workflows. When we designed the system, we sought to balance the goal of maintaining physician autonomy given historical norms and facilitators and barriers in the local context with the goal of increasing standardization of the workflow. Following discussion, we decided not to mandate standardized workflow criteria beyond the nodule hashtag criteria and the requirement for daily triage and weekly committee meetings.

Also, in 2015, thoracic surgery was consolidated from 16 medical centers to five with the goal of increasing subspecialty expertise and surgical volume. A change made in the thoracic surgery department was fast-tracking of early-stage cases and increasing use of video-assisted thoracic surgery, which sometimes was performed without a prior biopsy.

Design and Study Population

The quasi-experimental, stepped wedge cluster design was implemented in a staggered fashion across 21 radiology centers. In a stepped wedge cluster design, each cluster switches from control to intervention consecutively at different time points.⁸ Each radiology center represented a cluster that switched from control to intervention, that is, from nonstandardized reporting to standardized reporting with population-based follow-up of suspicious findings. The advantage of this design was control for other time trends, including increasing use of chest CT imaging. This quality

TABLE 1] Tags, Descriptions, and Recommendations Used to Code Lung-Specific Findings on Diagnostic Chest CT Imaging^a

Tag and Description	Recommendation
#PUL0 or smart-text "No pulmonary nodule," "Lungs: Normal," or "Pulmonary Nodules/Masses: None": No pulmonary nodule	No further action
#PUL1: Benign nodule, benign pattern of calcification, definite fat, stable after surveillance	No further action
#PUL2: Pulmonary nodule that is likely infectious, inflammatory, or part of diffuse nodular lung disease	The need for any follow-up imaging depends on the broader clinical context
#PUL3: Subsolid nodule. Pure ground-glass and part ground-glass/part solid nodules. Upgrade to #PUL5 if (1) a solid component is initially ≥ 4 mm; (2) solid component grows to ≥ 4 mm; (3) a new solid component develops; (4) the overall nodule grows by ≥ 2 mm	Follow-up CT scan scheduled by the ordering provider according to Fleischner Society guidelines, based on nodule and patient risk factors for lung cancer
#PUL4: Indeterminate solid nodule ≤ 8 mm	Follow-up CT scan scheduled by the ordering provider according to Fleischner Society guidelines, based on nodule and patient risk factors for lung cancer
#PUL5: Suspicious for malignancy; solid nodule > 8 mm; growth of a nodule by ≥ 2 mm; pure ground-glass nodule that develops a solid component on follow-up; partly solid nodule with a solid component ≥ 4 mm; other finding suspicious for a primary thoracic malignancy	The CT report is automatically forwarded to an expert committee for review and follow-up
#PUL6: Known lung cancer or suspected metastatic nodule; established diagnosis of lung cancer; nodules detected in patients with a known primary malignancy that are considered most likely metastatic	Follow-up by ordering provider
#PULX: Technically limited: factors such as significant motion, artifacts, excluded lung portions, etc., prevent exclusion of pulmonary nodules	Follow-up by ordering provider

^aThe accrual period for the present analysis ended in 2017. The recommendations for #PUL3 and #PUL4 were adapted from the 2005 Fleischner Society guidelines for solid nodules and the 2013 Fleischner Society guidelines for subsolid nodules. In 2018, the system was updated to reflect the 2017 Fleischner Society guidelines.

improvement evaluation was determined by the national compliance officer of the Kaiser Foundation Research Institute as not meeting the regulatory definition of human subjects research.

The study population included all Kaiser Permanente Northern California members aged ≥ 18 years who underwent nonscreening chest CT imaging (current procedural terminology [CPT] codes 71250, 71260, 71275) from January 1, 2015, to July 31, 2017, and had at least 2 years of continuous health plan enrollment preceding their first chest CT scan, hereafter referred to as the index CT scan. We excluded those who had a chest CT scan within 2 years of their index CT scan and who possibly were under surveillance for a prior nodule. We further excluded those previously diagnosed with lung cancer (International Classification of Diseases [ICD]-O-3 site codes C34.0-C34.9) more than 120 days before their index CT scan. Primary diagnoses of lung cancer were identified using our institution's cancer registry. We retained those with a lung cancer diagnosis received within the 120 days preceding their index CT scan, because in the cancer registry, the diagnosis date is based on the earliest clinical evidence of cancer,^{9,10} which includes chest radiography leading to chest CT imaging.

Outcomes

The main outcome was a new diagnosis of early- or late-stage primary lung cancer within 120 days of the index CT scan. Cancer diagnoses were ascertained from the cancer registry and included all cases,

including those that were and were not pathologically confirmed. Stage at diagnosis was classified using the American Joint Committee on Cancer 7th edition criteria and was classified as early (I, II) or late (III, IV). Two secondary outcomes included (1) days from index CT scan to pathologic diagnosis within 120 days and (2) days from index CT scan to initial surgical treatment within 120 days. The interval of 120 days was selected a priori on the basis of considerations of the clinical workflow. In a sensitivity analysis, we also examined 180 days. Primary diagnoses of lung cancer are confirmed both clinically and pathologically. The time from index CT scan to clinical diagnosis is not very informative, because patients may receive a clinical diagnosis before the index CT scan, based on symptoms or radiography. Therefore, we used the date of pathologic diagnosis when it was available, but used the date of clinical diagnosis for those without pathologic confirmation. The date of pathologic diagnosis was determined from the collection date of the first biopsy or surgical specimen. The date of clinical diagnosis was obtained from the cancer registry. The date of surgery was determined on the basis of diagnostic and procedure codes (e-Appendix 1).

Covariables

Patient age, sex, race/ethnicity, and smoking history were ascertained from electronic clinical and administrative records. The Charlson Comorbidity Index was derived from diagnoses recorded during the 12 months before the index CT date. Nonextremity radiography, CT, MRI, and ultrasound examinations that were performed during the

TABLE 2] Patient Characteristics Before and After Implementation of a Chest CT Imaging Reporting System: Kaiser Permanente Northern California, January 2015 to July 2017, N = 99,148^a

Characteristic	%	
	Preimplementation (n = 59,739)	Postimplementation (n = 39,409)
Lung cancer		
No	97.1	97.2
Yes	2.9	2.8
Year		
2015	57	7
2016	34	47
2017 ^b	9	46
Age, y		
18-49	16	17
50-59	17	17
60-69	25	24
70-79	24	23
≥ 80	18	18
Race/ethnicity		
African-American	8	9
Asian/Pacific Islander	12	13
Hispanic	13	15
White	61	58
Other/missing	6	6
Sex		
Female	54	54
Male	46	46
Charlson Comorbidity Index		
0	29	26
1	20	19
2-3	25	26
≥ 4	26	29
Smoking history		
Never	51	52
Ever	49	48
Missing	< 1	< 1
History of cancer^c		
No	80	81
Yes	20	19

^aBecause of the large sample size, *P* values were significant with differences smaller than 1% and therefore are not shown.

^bJanuary 1 to July 31, 2017.

^cIncludes all cancers except nonmelanoma skin cancer and lung carcinoma. Patients with a history of lung carcinoma were excluded from the study.

120 days before index CT scan were identified on the basis of CPT and internal procedure codes (see the online article).

Statistical Analysis

Following a quasi-experimental stepped wedge cluster design, the timing of the index CT scan, that is, before vs after implementation

of the intervention, determined classification into preimplementation (control) and postimplementation groups.⁸ The dichotomous outcome of lung cancer diagnosis was modeled using generalized estimating equations with a logit link to calculate the OR and 95% CI associated with implementation of the intervention with sandwich-type variance estimates to account for clustering. Models

TABLE 3] Adjusted ORs and 95% CIs for the Association of Implementation of a Chest CT Imaging Reporting System With Lung Cancer Diagnosis Within 120 Days: Kaiser Permanente Northern California, January 2015 to July 2017, N = 99,148^a

Characteristic	No Lung Cancer: %	Early Stage: %	Late Stage: %	All Stages ^b		Early Stage		Late Stage	
	(n = 96,292)	(n = 795)	(n = 2,027)	Adjusted OR	95% CI	Adjusted OR	95% CI	Adjusted OR	95% CI
Implementation									
Pre	60	58	63	1.00	Ref.	1.00	Ref.	1.00	Ref.
Post	40	42	37	1.09	1.00-1.18	1.24	1.09-1.41	1.04	0.95-1.14
Age, y									
18-49	17	2	2	1.00	Ref.	1.00	Ref.	1.00	Ref.
50-59	17	9	10	3.57	2.88-4.43	3.80	1.86-7.74	3.55	2.80-4.50
60-69	25	28	27	6.82	5.58-8.34	8.26	4.80-14.22	6.44	5.08-8.18
70-79	23	37	35	9.68	7.92-11.83	11.53	6.60-20.29	9.13	7.24-11.52
≥ 80	18	25	25	10.00	8.21-12.20	11.59	6.72-19.98	9.58	7.47-12.30
Race/ethnicity									
African-American	8	7	7	0.87	0.74-1.03	0.94	0.74-1.20	0.87	0.69-1.10
Asian/Pacific Islander	12	13	14	1.52	1.33-1.75	1.42	1.17-1.73	1.57	1.34-1.85
Hispanic	14	8	7	0.74	0.67-0.81	0.81	0.64-1.01	0.70	0.60-0.81
White	60	68	67	1.00	Ref.	1.00	Ref.	1.00	Ref.
Other/missing	6	5	5	0.85	0.74-0.98	0.79	0.61-1.03	0.90	0.74-1.08
Sex									
Female	54	57	51	1.00	Ref.	1.00	Ref.	1.00	Ref.
Male	46	43	49	0.82	0.76-0.89	0.70	0.59-0.83	0.88	0.79-0.99
Charlson Comorbidity Index									
0	28	17	22	1.00	Ref.	1.00	Ref.	1.00	Ref.
1	20	24	22	1.07	0.97-1.18	1.32	1.06-1.63	1.00	0.90-1.11
2-3	25	35	25	0.82	0.73-0.93	1.19	0.94-1.50	0.71	0.64-0.80
≥ 4	27	25	30	0.72	0.63-0.84	0.71	0.53-0.94	0.73	0.65-0.83
Smoking history									
Never	52	24	21	1.00	Ref.	1.00	Ref.	1.00	Ref.
Ever	48	76	79	3.66	3.17-4.22	3.32	2.72-4.03	3.81	3.21-4.51
History of cancer^c									
No	80	79	83	1.00	Ref.	1.00	Ref.	1.00	Ref.
Yes	20	21	17	0.76	0.69-0.84	0.96	0.81-1.14	0.70	0.61-0.81

^aThe logistic regression model included every variable shown in Table 3, as coded in Table 3, as well as calendar period (6-month interval) as a fixed effect and radiology center as a random effect. Patients whose data on smoking history (1%) or sex (< 0.01%) was missing were excluded from the logistic regression analysis.

^bPatients with missing lung cancer stage information were included in the full analysis but excluded from stage-specific analyses.

^cIncludes all cancers, except nonmelanoma skin cancer and lung carcinoma. Patients with a history of lung carcinoma were excluded from the study.

were adjusted for each 6-month time period of the study, patient age, sex, race/ethnicity, Charlson Comorbidity Index, history of cancer (other than lung or nonmelanoma skin cancer), and smoking status and accounted for clustering of patients in radiology centers. Analyses were conducted first for the outcome of lung cancer

diagnosis, and then separately for the outcomes of early- and late-stage lung cancer diagnosis.

Among patients receiving diagnoses of lung cancer, we fit cumulative incidence curves for pathologic diagnosis and surgery before and

after implementation of the intervention, using an Aalen-Johansen estimator to account for competing risks due to death.¹¹ We used Cox proportional hazards regression with Fine and Gray competing risks to test the hypothesis that implementation of the reporting system was associated with pathologic diagnosis and surgical treatment.¹² Follow-up (in days) began on the date of the index CT scan and ended on the date of the outcome, health plan disenrollment, death, or follow-up end at 120 days, whichever occurred earliest. Hazard ratios (HRs) and 95% CIs were adjusted for 6-month time period, patient age, sex, race/ethnicity, Charlson

Comorbidity Index, history of cancer, and smoking status. We accounted for clustering of patients by radiology center, using a working covariance and robust sandwich estimators to account for unknown covariance. Because analyses indicated that the proportional hazards assumption was violated for the association of the intervention with incidence of pathologic diagnosis, we estimated separate HRs for diagnosis within and more than 21 days after CT imaging.¹³ These analyses were performed both for the overall population and for subgroups defined by disease stage.

Results

During the study period, from January 2015 to July 2017, chest CT imaging use increased in our system by 19%, from 31 to 36 per 1,000 person-years. We identified 125,189 adult patients who underwent a chest CT imaging examination during this period. After excluding 25,374 patients (20%) who had a chest CT scan during the preceding 24 months, 625 patients (< 1%) who had a history of lung cancer, and 42 patients (< 1%) who had a history of cancer that could not be dated, our study population included 99,148 patients, of whom 60% (n = 59,739) had their index CT scan before and 40% (n = 39,409) had their index CT scan after implementation of the intervention. Postimplementation patients were disproportionately nonwhite (42% vs 39%) and had larger Charlson comorbidity scores (≥ 4 ; 29% vs 26%) (Table 2). In addition, 82% of the preimplementation patients and 81% of the postimplementation patients had a preceding radiograph, while 41% and 42% of the two groups had a preceding nonchest image, including head/neck CT scan, abdominal/pelvic CT scan, abdominal/pelvic radiograph, head/neck MRI, or whole body imaging, respectively (e-Table 1).

During observation, 2,856 patients received new diagnoses of lung cancer (with or without pathologic diagnosis), of whom 1,751 received diagnoses preimplementation (2.9% of 59,739) and 1,105 received diagnoses postimplementation (2.8% of 39,409). Lung cancer was detected at an early stage in 26% of preimplementation cases and in 30% of postimplementation cases. Using general estimating equations (Table 3), implementation of the intervention was associated with a 24% increase in the odds of early-stage lung cancer diagnosis within 120 days (OR, 1.24; 95% CI, 1.09-1.41) and a 4% increase in the odds of late-stage lung cancer diagnosis within 120 days (OR, 1.04; 95% CI, 0.95-1.14). The results were very similar when we examined lung cancer cases diagnosed within 180 days (early stage: OR, 1.22; 95% CI, 0.96-1.55; late stage: OR, 0.92; 95% CI, 0.69-1.25; all stages: OR, 1.07;

95% CI, 0.90-1.28). When we used 3-month intervals and continuous time in place of 6-month intervals to account for time trends, we found no difference in the effect estimate or CI (6-month interval: OR, 1.09; 95% CI, 1.00-1.18; 3-month interval: OR, 1.08; 95% CI, 1.00-1.17; continuous: OR, 1.08; 95% CI, 1.00-1.18).

A pathologic specimen was obtained within 120 days of the index CT scan for 77% of all patients with lung cancer, 92% of those with early-stage cancer, and 72% of those with late-stage cancer (Table 4). A similar proportion of pre- and postimplementation patients received a pathologic diagnosis (77%, 77%). However, more postimplementation patients received their first pathologic diagnosis from a surgical specimen rather than a biopsy specimen (15% vs 9%) and they had a slightly longer time to pathologic diagnosis (median: pre, 12 days; post, 15 days). In unadjusted analyses, the cumulative incidence of pathologic diagnosis differed over time between pre- and postimplementation patients (Fig 1). Adjusted analyses using Cox proportional hazards analysis confirmed this observation. During the first 21 days after CT imaging, the intervention was associated with a slower rate of pathologic diagnosis, whereas after 21 days, it was associated with a faster rate of pathologic diagnosis. For patients with early-stage cancer, the HR was 0.69 (95% CI, 0.53-0.90) during the first 21 days and 1.08 (95% CI, 0.94-1.24) after 21 days. For patients with late-stage cancer, it was 0.83 (95% CI, 0.74-0.93) during the first 21 days and 1.13 (95% CI, 0.96-1.33) after 21 days. Among all patients with lung cancer, it was 0.79 (95% CI, 0.71-0.87) during the first 21 days and 1.15 (95% CI, 1.02-1.29) after 21 days.

Surgery occurred within 120 days of the index CT scan for 26% of all patients with lung cancer, 59% of those with early-stage cancer, and 13% of those with late-stage cancer (Table 4). The median time from index CT scan to surgery was 43 days (preimplementation, 46 days; postimplementation, 41 days). In bivariate analysis, surgery was more frequent among postimplementation

TABLE 4] Unadjusted Frequencies of Pathology and Surgery After Index CT Examination Among N = 2,856 Patients With Lung Cancer

Cohort ^a and Utilization	Preimplementation		Postimplementation		P Value
	No.	%	No.	%	
All stages	n = 1,751		n = 1,105		
Lung pathology	1,354	77	850	77	.80
Complete follow-up without pathology	212	12	135	12	.93
Died during follow-up without pathology	183	10	118	11	.85
First pathology before surgery	1,197	68	685	62	< .01
First pathology during surgery	157	9	165	15	< .01
Surgery	406	23	334	30	< .01
Complete follow-up without surgery	921	53	518	47	< .01
Died during follow-up without surgery	412	24	245	22	.40
Early stage	n = 460		n = 335		
Lung pathology	422	92	307	92	.96
Complete follow-up without pathology	30	6	23	7	.85
Died during follow-up without pathology	8	2	5	2	.79
First pathology before surgery	338	73	222	66	.03
First pathology during surgery	84	18	85	25	.02
Surgery	264	57	205	61	.28
Complete follow-up without surgery	184	40	118	35	.17
Died during follow-up without surgery	11	2	12	4	.32
Late stage	n = 1,269		n = 758		
Lung pathology	918	72	532	70	.29
Complete follow-up without pathology	176	14	111	15	.63
Died during follow-up without pathology	173	14	113	15	.43
First pathology before surgery	847	67	453	60	< .01
First pathology during surgery	71	6	79	10	< .01
Surgery	137	11	126	17	< .01
Complete follow-up without surgery	723	57	393	52	.02
Died during follow-up without surgery	398	31	231	30	.68

^aLess than 2% disenrolled from each cohort.

than preimplementation patients (Fig 1). However, in multivariable Cox proportional hazards regression analysis, implementation was not clearly associated with the rate of surgery for early-stage (HR, 1.08; 95% CI, 0.86-1.35) or late-stage (HR, 0.94; 95% CI, 0.75-1.18) cancer.

Discussion

We studied the effectiveness of an intervention that tagged chest CT scan reports and automated referral of patients with findings suspicious for lung cancer to a multidisciplinary care team for evaluation and follow-up. The study included nearly 100,000 patients, of whom 40% were accrued after implementation of this intervention. The intervention was associated with

9% greater odds of diagnosing new lung cancer (OR, 1.09; 95% CI, 1.00-1.18) and 24% greater odds of diagnosing early-stage lung cancer (OR, 1.24; 95% CI, 1.09-1.41) but was not associated with diagnosing late-stage lung cancer within 120 days of CT examination. A possible explanation for the lack of association with late-stage disease is the relative ease of diagnosis and the lack of opportunity for improvement. The community-based setting and large sample size enhanced the informativeness of our findings.

The stepped wedge design adjusted for inherent differences across radiology centers and for time trends resulting from increased use of CT imaging over time, reduced the incidence of lung cancer over time, and possible improved radiologist performance and CT

Figure 1 – Unadjusted cumulative incidence of pathologic diagnosis and surgery in days from index CT scan, overall and by stage.

imaging resolution. We also examined patient-level differences, but these did not confound the relationship. Among the potential confounding factors examined, only calendar time affected the estimated ORs. During the intervention's roll-out period, the overall rate of chest CT scans increased in our system, whereas the incidence of lung cancer decreased.¹⁴ By adjusting for time in our multivariable models, we minimized these confounding effects. In sensitivity analysis, instead of using 6-month intervals, we adjusted for calendar time by using 3-month intervals and continuous time, but found no difference in the effect estimates.

A key question in evaluating pulmonary findings is how processes and tools should be deployed to optimize care.¹⁵ A 2016 survey of American Thoracic Society members reported that the three most common systems-level resources included (1) adding Fleischner Society guidelines to radiology reports, (2) adding prompts to the ordering provider on radiology reports describing new nodules, and (3) employing clinical staff to coordinate recommended follow-up.^{16,17} Our intervention includes these three components.

A major goal in implementing this reporting system was to more efficiently diagnose and treat patients with lung

cancer at its earliest stage. Overall, our results suggest that the reporting system has been effective in detecting new lung cancer cases, particularly at an early stage, that otherwise would have been detected after 120 days. The introduction of multidisciplinary case management, however, resulted in an initial delay in pathologic diagnosis to 21 days or more, although it had no negative impact on treatment by surgery. This delay can be explained, at least partly, by the increased recommendation for surgery without biopsy confirmation. This improved adherence to guidelines and avoidance of unnecessary biopsies represent a potential benefit of the multidisciplinary committee. Another contributing factor was the added time required to route and evaluate patients through a weekly multidisciplinary committee, compared with direct referral to the ordering or primary care provider for follow-up care.

A potential unintended consequence of the tagging system was an increase in workup for lung cancer in patients without lung cancer. Among the patients with #PUL5 tags, the sensitivity of the tag was 74% (95% CI, 72%-77%), the positive predictive value was 36% (35%-38%), the number needed to diagnose was 2.7 (2.6-2.9), and the specificity and negative

predictive value were > 95%. We did not evaluate the potential increase in physician workload. In addition, some patients with late-stage disease are eligible for targeted therapy and may benefit from expedited multidisciplinary review.

Tagging suspicious findings and automating referral to a regional multidisciplinary committee increases the committee's expertise and likely increases care quality.¹⁸ Physical tags have been used to rapidly triage patients in mass casualty situations, an application that is similar conceptually to the problem of processing large volumes of imaging reports. As a general medical informatics solution, tagging systems are easy to establish and effective when the physicians assigning tags and those acting on the tags share a precise vocabulary, the effort needed to apply the tag is minimal, and the information encoded in the tag is critical to decision-making. Similarly, tags with pelvic ultrasound for follow-up of adnexal masses were effective for supporting risk-based management by physicians, based on ultrasound characteristics.^{19,20}

A key limitation is generalizability. Translation to less integrated settings may require more coordination and support than are available, although the US Department of Veterans Affairs also uses a nodule tracking system.^{21,22} A second limitation relates to comorbidity scores, which differed slightly pre- and postimplementation. Across our studies, we have noted increased coding related to the transition from ICD9 to

ICD10 in September 2015. A consequence is increased comorbidity scores since 2015 that we believe are spuriously related to the change in ICD coding.

Future work includes examining the impact of increased diagnosis of early-stage lung cancer on lung cancer mortality. We also plan to evaluate the effectiveness of tagging subsolid and indeterminate nodules for improving nodule management and outcomes over intermediate and longer periods of follow-up. In addition, we have found opportunities to improve the system through minor design changes and additional radiologist feedback and training to clarify that the definition of #PUL5 includes any finding suspicious for malignancy, not just pulmonary nodules, and #PUL6 provides a tag for patients who have received a confirmed cancer diagnosis, not those in whom cancer appears metastatic on CT imaging.

In conclusion, an informatics-based system in which tags are added to chest CT scan reports to standardize documentation and automate follow-up of findings suspicious for lung cancer was evaluated in a community-based setting, using a quasi-experimental stepped wedge cluster design that accounted for time trends, differences across radiology centers, and patient-level differences. Our results suggest this intervention has been effective in increasing diagnosis but not surgical treatment of early-stage lung cancer within 120 days following chest CT imaging.

Acknowledgments

Author contributions: L. J. H. is the guarantor of this manuscript and its content. J. R. D. had full access to the data and performed the analyses. J. R. D., L. J. H., S. A., and L. C. S. contributed to the study design. All authors meaningfully interpreted the results, and T. H. U., J. R. D., L. J. H., S. A., D. A. C., T. O., and L. C. S. contributed to writing the manuscript. All authors approved the manuscript.

Financial/nonfinancial disclosures: The authors have reported to *CHEST* the following: The investigators are partners and staff of the Permanente Medical Group.

Role of sponsors: The sponsors had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

Additional information: The e-Appendix and e-Table can be found in the Supplemental Materials section of the online article.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(1):7-34.
2. Smith-Bindman R, Kwan ML, Marlow EC, et al. Trends in use of medical imaging in US health care systems and in Ontario, Canada, 2000-2016. *JAMA*. 2019;322(9):843-856.
3. Wiener RS, Gould MK, Slatore CG, Fincke BG, Schwartz LM, Woloshin S. Resource use and guideline concordance in evaluation of pulmonary nodules for cancer: too much and too little care. *JAMA Intern Med*. 2014;174(6):871-880.
4. D'Orsi CJ, Sickles EA, Mendelson EB, et al. *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*. Reston, VA: American College of Radiology; 2013.
5. Naidich DP, Bankier AA, MacMahon H, et al. Recommendations for the management of subsolid pulmonary nodules detected at CT: a statement from the Fleischner Society. *Radiology*. 2013;266(1):304-317.
6. MacMahon H, Austin JH, Gamsu G, et al; Fleischner Society. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. *Radiology*. 2005;237(2):395-400.
7. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner Society 2017. *Radiology*. 2017;284(1):228-243.
8. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials*. 2007;28(2):182-191.
9. National Cancer Institute. SEER Program Coding and Staging Manual 2012, p. 13. https://seer.cancer.gov/archive/manuals/2012/SPCSM_2012_maindoc.pdf. Accessed June 29, 2020.
10. National Cancer Institute. SEER Registrar Staging Assistant: NAACCR Item #390: Date of Diagnosis. <https://staging.seer>.

- cancer.gov/naaccr/item/cs/02.05.50/390/?version=/tnm/home/1.2/. Accessed June 29, 2020.
11. Aalen OO, Johansen S. An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scand J Stat.* 1978;5(3):141-150.
 12. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat.* 1988;16:1141-1154.
 13. SAS Institute. Example 64.12: Model Assessment Using Cumulative Sums of Martingale Residuals. In: *SAS/STAT[®] 9.2 User's Guide*. 2nd ed. Cary, NC: SAS Institute; 2008. https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_phregsect043.htm. Accessed June 29, 2020.
 14. National Cancer Institute. Surveillance, Epidemiology, and End Results Program: Cancer Stat Factors: Lung and Bronchus Cancer. <https://seer.cancer.gov/statfacts/html/lungb.html>. Accessed June 29, 2020.
 15. Slatore CG, Horeweg N, Jett JR, et al; ATS Ad Hoc Committee on Setting a Research Framework for Pulmonary Nodule Evaluation. An official American Thoracic Society research statement: a research framework for pulmonary nodule evaluation and management. *Am J Respir Crit Care Med.* 2015;192(4):500-514.
 16. Simmons J, Gould MK, Iaccarino J, Slatore CG, Wiener RS. Systems-level resources for pulmonary nodule evaluation in the United States: a national survey. *Am J Respir Crit Care Med.* 2016;193(9):1063-1065.
 17. McDonald JS, Koo CW, White D, Hartman TE, Bender CE, Sykes AG. Addition of the Fleischner Society guidelines to chest CT examination interpretive reports improves adherence to recommended follow-up care for incidental pulmonary nodules. *Acad Radiol.* 2017;24(3):337-344.
 18. Cassel CK, Reuben DB. Specialization, subspecialization, and subspecialization in internal medicine. *N Engl J Med.* 2011;364(12):1169-1173.
 19. Suh-Burgmann E, Flanagan T, Osinski T, Alavi M, Herrinton L. Prospective validation of a standardized ultrasonography-based ovarian cancer risk assessment system. *Obstet Gynecol.* 2018;132(5):1101-1111.
 20. Suh-Burgmann EJ, Flanagan T, Lee N, et al. Large-scale implementation of structured reporting of adnexal masses on ultrasound. *J Am Coll Radiol.* 2018;15(5):755-761.
 21. Holden WE, Lewinsohn DM, Osborne ML, et al. Use of a clinical pathway to manage unsuspected radiographic findings. *Chest.* 2004;125(5):1753-1760.
 22. Moseson EM, Wiener RS, Golden SE, et al. Patient and clinician characteristics associated with adherence: a cohort study of veterans with incidental pulmonary nodules. *Ann Am Thorac Soc.* 2016;13(5):651-659.