



Published in final edited form as:

*AJR Am J Roentgenol.* 2022 April ; 218(4): 634–641. doi:10.2214/AJR.21.26927.

## Cancer Risk in Nodules Detected at Follow-up Lung Cancer Screening CT

Mark M Hammer, MD<sup>1</sup>, Suzanne C Byrne, MD<sup>1</sup>

<sup>1</sup>Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115

### Abstract

**Background:** Nodules may have different lung cancer risks when new on follow-up CT versus when present on an earlier CT (“existing” nodules). Diameter-based Lung-RADS and volume-based NELSON categories have shown variable performance in nodule risk assessment.

**Objective:** To assess Lung-RADS and NELSON classifications for nodules detected on follow-up lung cancer screening CT examinations.

**Methods:** This retrospective study included 185 patients (100 women, 85 men; median age, 66 years) who underwent lung cancer screening CT examinations for which a prior CT was available. Stratified random sampling was performed to enrich the sample with suspicious nodules, yielding 50, 45, 47, 30, and 13 nodules with Lung-RADS categories 2, 3, 4A, 4B, and 4X, respectively. Lung-RADS categories were recorded from clinical reports. Nodules' linear measurements were extracted from clinical reports to generate Lung-RADS categories using strict criteria. Two radiologists used a semiautomated tool to obtain nodule volumes, which were used to generate NELSON categories. Lung cancer risk was assessed. ROC analysis was performed. Percentages were weighted by Lung-RADS category frequencies in the underlying screening cohort.

**Results:** Twenty-nine cancers were diagnosed. Weighted cancer risk was 5% for new nodules, 1% for stable existing nodules, and 44% for growing existing nodules. No clinical category 2 nodule was cancer. Using strict Lung-RADS, 34 nodules, including 7 cancers, were downgraded to category 2. AUC for cancer was 0.96 for clinical Lung-RADS, 0.81 for strict Lung-RADS, 0.71–0.84 for NELSON (two readers), and 0.89 for nodule diameter. Clinical Lung-RADS achieved weighted sensitivity and specificity of 100% and 85% in the entire sample, 100% and 41% in new nodules, and 100% and 94% in existing nodules. Optimal diameter threshold was 8 mm for existing nodules, versus 6 mm for new nodules.

**Conclusions:** Lung-RADS, as applied by radiologists in clinical practice, achieved excellent performance on follow-up screening examinations. Strict Lung-RADS downgraded some cancers to category 2. Volumetric assessments had weaker performance than clinical Lung-RADS. New nodules warrant smaller size thresholds than existing nodules.

**Correspondence:** MMH, Department of Radiology, 75 Francis Street, Boston, MA 02115, mmhammer@bwh.harvard.edu, 617-732-6285.

The authors have no financial disclosures.

**Clinical Impact:** The findings provide insight into radiologists' management of nodules detected on follow-up screening examinations.

---

## Introduction

With increasing utilization of CT lung cancer screening, algorithms to triage pulmonary nodules are critically important to provide optimal management of these patients. Various approaches have been described to guide the evaluation of nodules detected in lung cancer screening. These include the American College of Radiology Lung-RADS [1], which is based primarily on linear measurement of nodule size, as well as the algorithm used in the Dutch NELSON trial [2], which is based primarily on nodule volume and growth rate.

The approach to nodules identified on a follow-up lung cancer screening CT is somewhat different from the approach for nodules seen on a baseline CT. Nodules that are new on follow-up CT may have a higher risk of lung cancer than nodules on a follow-up CT that were present on a baseline CT (hereafter, "existing" nodules) [3]. Certain nodule triage algorithms, such as those from the Fleischner Society [4] and the original NELSON trial, use the same size cut-offs for new nodules on follow-up CT as for nodules on baseline CT. Conversely, Lung-RADS recommends a lower size threshold for raising suspicion for cancer for new nodules on follow-up CT than for nodules on baseline CT. For existing nodules at follow-up, both Lung-RADS and NELSON use growth since the earlier examination to determine the need for follow-up: nodules that are either not growing or (for NELSON) slowly growing do not warrant close follow-up.

A number of studies have compared the performance of nodule risk assessment algorithms, for example in the National Lung Screening Trial and NELSON trial cohorts [5–7], yielding inconsistent findings regarding the relative performance of diameter-based Lung-RADS versus volume-based NELSON. However, none of these studies, to our knowledge, specifically focused on nodules detected on follow-up screening CT examinations. Therefore, in this study, we evaluated the utility of both risk assessment algorithms (Lung-RADS and NELSON) for nodules detected on follow-up lung cancer screening CT examinations.

## Methods

### Patient Selection

This retrospective HIPAA-compliant study was approved by the institutional review board with waiver of informed consent. We performed an automated search of a database of 5835 patients who underwent lung cancer screening CT between July 2015 and August 2018 within our healthcare network (which included two academic sites and one community site), to identify those who also had a prior chest CT within our network. This search yielded a total of 1150 patients who underwent a total of 2245 CT examinations with a prior examination. Among these examinations, based on the original clinical interpretations, 513 were Lung-RADS category 1, 1435 (83%) were category 2, 166 (10%) were category 3, 80 (5%) were category 4A, 34 (2%) were category 4B, and 14 (1%) were category 4X. For each Lung-RADS category (except for categories 1, 4B, and 4X), we randomly

selected examinations, with a prior specified maximum of 50 examinations per category (aside from category 1, for which no examinations were selected). For categories 4B and 4X, all examinations were initially selected given fewer than 50 examinations available in each category. This initial random selection process yielded 198 examinations in 185 patients. After the initial selection, it was discovered that two category 2 nodules per the clinical reports were misclassified by the search algorithm (one as category 4A, the other as category 4B); these two nodules were reclassified as category 2 based on the clinical reports. After this reclassification, an additional category 4B nodule was randomly selected to replace the reclassified category 4B nodule. Finally, the first follow-up examination was selected for any of the 185 patients who had multiple follow-up examinations. Following this process, the study sample included 185 follow-up screening examinations in 185 patients (100 women, 85 men; median age, 66 years), that detected a nodule (50 Lung-RADS category 2, 45 category 3, 47 category 4A, 30 category 4B, and 13 category 4X). Patient demographics and family history of lung cancer were extracted from the electronic medical record. For examinations with multiple nodules, only the nodule determining the overall assigned category in the report was considered in the analysis.

### **Clinical Lung-RADS and Strict Lung-RADS Categories**

Examinations had been interpreted using Lung-RADS version 1.0, given the time period of the study. A total of 181 examinations were interpreted by fellowship-trained cardiothoracic radiologists at one of the two academic sites; 4 examinations were interpreted by radiologists at the community site. The Lung-RADS categories were extracted automatically from the clinical reports, and then corrected if not matching the clinical report on further manual review. Reports were also manually reviewed to extract nodule attenuation (classified as solid, part-solid, or ground glass), nodule size on the present CT, and nodule size on the prior CT. The clinically recorded nodule size measurements typically represented the mean of nodules' long and short axis diameters, per institutional practice. If the current report did not report the prior nodule size, then the size measurement was obtained from the prior report. These size measurements were then used to retrospectively determine a Lung-RADS category for each nodule, based on size thresholds for new and existing nodules in Lung-RADS version 1.1 ("strict" Lung-RADS). Clinical Lung-RADS categories of 4X were not adjusted because this category depends on additional aspects of nodule appearance, though the images were not reviewed for purposes of this assessment.

### **Retrospective Determination of Nodule Volumes and NELSON Categories**

Two fellowship-trained thoracic radiologists (MH and SB, both with 6 years of post-training experience) independently reviewed the images for each patient. The radiologists accessed the clinical reports to identify the category-determining nodule but were blinded to subsequent imaging and clinical diagnoses. To derive NELSON categories for each nodule, the radiologists classified the nodule as new or existing, assessed its attenuation (solid, part-solid, or ground glass), and determined its volume using a semi-automated segmentation software (Syngo VIA, version VB40, Siemens Healthcare, Erlangen, Germany). For nodules classified as existing at the time of this retrospective review, the radiologists segmented the nodule on both the prior and current examinations. For part-solid nodules, the radiologist separately segmented both the total nodule and the nodule's solid component. Using

the readers' nodule assessments and volumetric measurements, NELSON categories were assigned for each nodule for each reader [2]. NELSON categorizes new nodules (as well as nodules on baseline examinations) as NODCAT 1, 2, 3, or 4, depending on nodule attenuation and volume. NELSON categorizes existing nodules as GROWCAT A, B, or C, depending on volume doubling times (>600 days, 400–600 days, or 0–399 days, respectively).

### Reference Standard

The electronic medical record was manually reviewed through March 2021 to identify any clinical diagnoses of lung cancer attributed to the category-determining nodule for each patient. In patients without histologic confirmation, empiric treatment with radiation therapy was considered to indicate an empiric diagnosis of lung cancer.

### Data Analysis

Characteristics of patients and nodules were described using summary characteristics. Characteristics were compared between new and existing nodules using Wilcoxon and Fisher's exact tests. Results were weighted by the relative frequencies of Lung-RADS categories in the entire lung cancer screening population of 2245 patients from which the study sample was drawn. This weighting was intended to provide representative cancer risks for an overall screening population (e.g., a population in which Lung-RADS category is much more common than categories 3 and 4), given the selection of similar numbers of nodules for each category for the present study sample. The weighting of categories was not performed when summarizing cancer frequency of individual clinical Lung-RADS categories. The frequency of lung cancer was summarized, stratified by combinations of clinical versus strict Lung-RADS, new versus existing nodules, and Lung-RADS category. These frequencies were compared between new and existing nodules for each clinical Lung-RADS category using Fisher's exact test. The weighted cancer frequency was summarized for new, stable existing, and growing existing nodules, and compared using Fisher's exact test. Characteristics of cancers on follow-up CT examinations were summarized, stratified by new, stable existing, and growing existing nodules. Nodules reclassified when using strict Lung-RADS were summarized in terms of the reassignments, reasons for reassignments, and characteristics of cancers among reassigned nodules. Interreader agreement for nodule volumes was assessed using the intraclass correlation coefficient (ICC). Lung cancer risk was also summarized by NELSON category, stratified by reader and new versus existing nodules. Sensitivity and specificity for lung cancer were assessed for clinical Lung-RADS, strict Lung-RADS, NELSON, and nodule diameter, performed for all nodules as well as separately for new and existing nodules; to allow comparison of approaches for the same sets of nodules, sensitivity and specificity for NELSON were computed based on designations of nodules as new or existing in the clinical reports. AUC for lung cancer by the various approaches was also computed for all nodules, but not separately for new and existing nodules given the small number of cancers among new nodules. All AUCs were calculated using the previously described weighting method, and CIs were generated by bootstrapping with 1500 samples. Clinical Lung-RADS and strict Lung-RADS were defined as positive at a category of 3 or greater. NELSON was defined as positive at a category of NODCAT 3 or greater for new nodules or of GROWCAT B or greater for existing

nodules. For nodule diameter, the optimal cut-off was identified using the Youden index. Differences in proportions were evaluated using Fisher's exact test. P values less than .05 were considered statistically significant. Data were initially entered in REDCap [8] and then downloaded and analyzed in JMP Pro (version 15.2, SAS Institute, Cary, NC).

## Results

### Patients and Nodules

Characteristics of the 185 patients and nodules are provided in Table 1. The median interval between the prior and current CT was 12 months (interquartile range, 8 – 15 months; total range, 1 – 96 months). The median nodule diameter was 7 mm (range, 2 – 30 mm). A total of 76% (140) of nodules were solid, 16% (30) were part solid, and 8% (15) were ground glass. A total of 56% (104) of nodules were existing, while 44% (81) were new. New and existing nodules were not significantly different in terms of patient age, patient sex, nodule morphology, or nodule size (all  $p > .05$ ). A total of 16% (29) of nodules were diagnosed as lung cancer (24 histologically; 5 empirically, given treatment with radiation therapy). These malignant nodules had a median diameter of 10 mm (range, 6 – 25 mm).

### Risk of Lung Cancer by Lung-RADS Category

The risk of lung cancer in new and existing nodules, stratified by Lung-RADS category, is provided in Table 2. Based on clinical Lung-RADS, the frequency of lung cancer in new and existing nodules was 0% (0/4) and 0% (0/46) in category 2 nodules ( $p > .99$ ), 10% (3/30) and 7% (1/15) in category 3 nodules ( $p = .59$ ), 4% (1/27) and 30% (6/20) in category 4A nodules ( $p = .02$ ), 0% (0/19) and 67% (8/12) in category 4B nodules ( $p < .001$ ), and 50% (1/2) and 82% (9/11) in category 4X nodules ( $p = .42$ ).

The total frequency of lung cancer in new nodules was 5/81; in stable existing nodules was 10/79; and in growing existing nodules was 14/25. These frequencies corresponded with weighted cancer rates of 5% for new nodules, 1% for stable existing nodules, and 44% for growing existing nodules. Overall, the frequency of cancer was not significantly different between new and stable existing nodules ( $p = .14$ ) or between growing and new nodules ( $p = .05$ ); cancer frequency was significantly higher in growing versus stable existing nodules ( $p = .003$ ). Table 3 compares patient age, patient sex, nodule attenuation, nodule diameter, as well as clinical and strict Lung-RADS categories, for the cancers in new, stable, and growing nodules on follow-up CT.

Strict Lung-RADS resulted in a change in Lung-RADS category for 59 of the 185 nodules: 1 category 2 nodule was changed to category 3; 29 category 3 nodules were changed to category 2 (18), 4A (7), and 4B (4); 27 category 4A nodules were changed to category 2 (14), category 3 (1), and category 4B (12); and 2 category 4B nodules were changed to category 2. A total of 34 nodules were downgraded from clinical Lung-RADS categories of 3, 4A, or 4B to strict Lung-RADS category 2, for the following reasons: unchanged size since the prior examination (i.e., did not exhibit an increase in size of  $> 1.5$  mm) ( $n = 23$ ); ground glass nodule measuring  $< 30$  mm ( $n = 10$ ); new solid nodule measuring  $< 4$  mm ( $n = 1$ ). As a net result of all adjustments in Lung-RADS categories, 83 nodules were category 2

using strict Lung-RADS, compared with 50 that were category 2 using clinical Lung-RADS. None of the 52 clinical Lung-RADS category 2 nodules were cancer. However, the nodules reclassified as category 2 using strict Lung-RADS included 7 cancers, representing 24% of the 29 cancers in the study cohort. These seven cancers reclassified as Lung-RADS 2 included six unchanged nodules (Fig. 1) and one existing ground glass nodule measuring <30 mm (Fig. 2).

### **Volumetric Analysis Using NELSON**

Interreader agreement for nodule volume, expressed as ICC, was 0.98 (95% CI: 0.97, 0.99). The frequency of lung cancer by NELSON categories are provided in Table 4. The weighted rates of cancer for new nodules were 0% for NODCAT 2 for both readers, 3% and 6% for NODCAT 3 for readers 1 and 2, and 5% for NODCAT 4 for both readers. The weighted lung cancer rates for existing nodules were 1% for GROWCAT A for both readers, 5% and 9% for GROWCAT B for readers 1 and 2, respectively, and 10% and 64% for GROWCAT C for readers 1 and 2.

### **Comparison of Methods for Lung Cancer Detection**

The diagnostic performance for lung cancer detection using the clinical Lung-RADS, strict Lung-RADS, NELSON, and nodule diameter measurements, are given in Table 5. Among all nodules on follow-up CT, the AUC was 0.96 for clinical Lung-RADS, 0.81 for strict Lung-RADS, 0.71–0.84 for NELSON for the two readers, and 0.89 for nodule diameter. The optimal threshold for nodule diameter was 8 mm. The sensitivity and specificity were 100% and 85% for clinical Lung-RADS, 68% and 88% for strict Lung-RADS, 64–75% and 75–88% for NELSON, and 86% and 84% for nodule diameter.

For new nodules, the optimal threshold for nodule diameter was 6 mm. The sensitivity and specificity were 100% and 41% for clinical Lung-RADS, 74% and 38% for strict Lung-RADS, 48–100% and 46% for NELSON, and 100% and 62% for nodule diameter. For existing nodules, the optimal threshold for nodule diameter was 8 mm. The sensitivity and specificity were 100% and 94% for clinical Lung-RADS, 65% and 98% for strict Lung-RADS, 65–71% and 81–96% for NELSON, and 96% and 85% for nodule diameter. The use of a diameter threshold of 8 mm for new nodules resulted in a sensitivity of 61% (3/5) and a specificity of 82% (44/76). The use of a threshold of 6 mm for existing nodules resulted in a sensitivity of 100% (24/24) and specificity of 64% (35/80).

### **Discussion**

We evaluated the performance of the Lung-RADS and NELSON nodule classification schemes for nodules detected at follow-up lung cancer screening CT. The frequency of cancer was 1% in stable existing nodules, 5% in new nodules, and 44% in growing existing nodules. Among both new and existing nodules, the frequency of cancer increased with increasing Lung-RADS categories. Clinical Lung-RADS scores exhibited excellent sensitivity and specificity for cancer in existing nodules and excellent sensitivity in new nodules, though low specificity in new nodules. Volumetric assessment using the NELSON scheme exhibited lower diagnostic performance than clinical Lung-RADS scores. A nodule

diameter cut-off of 8 mm yielded 96% sensitivity in existing nodules, but only 61% in new nodules; rather, a smaller diameter cutoff of 6 mm was optimal in new nodules. The lower size threshold for new versus existing nodules is consistent with smaller thresholds for new nodules in the Lung-RADS recommendations [1].

Strict application of the Lung-RADS criteria (i.e., adjusting nodule categories based on size and growth criteria, aside from category 4X which was not adjusted) resulted in downgrading of many nodules from categories 3 through 4B to category 2, most commonly because they were stable nodules or were pure ground glass nodules measuring <30 mm. Though the clinical Lung-RADS categories had 100% sensitivity for cancer at a threshold category of 3 or higher, the nodules downgraded to category 2 included numerous cancers. The downgrading of malignancies based on strict application of Lung-RADS criteria indicates that the interpreting radiologists had exercised discretion beyond size measurements in the usage of Lung-RADS recommendations, so as to report higher categories for nodules that they believed to be likely malignant. That is, radiologists likely heavily considered additional nodule features beyond size measurements when categorizing nodules. Current practices for assessing stable nodules that do not exhibit entirely benign characteristics (including characteristics other than spiculation) are inconsistent, as radiologists may variably elect to use lower categories than 4X to describe nodules with suspicious features if the nodules are stable or show minimal growth. Our results support 4X as the most appropriate category for nodules that do not exhibit entirely benign characteristics, even if stable. The use of category 4X in this context is consistent with descriptions in the Lung-RADS document.

While use of category 4X in the described fashion helps maximize cancer detection, whether maximal sensitivity should be sought when using Lung-RADS is unclear. Pure ground glass nodules exhibit indolent behavior, with a very low risk of metastatic disease or recurrence after resection, and follow-up rather than immediate therapy of such nodules is supported by cost-effectiveness analysis [9]. The management of slowly growing solid nodules is controversial. Surgeons may prefer early definitive therapy with surgical resection of solid nodules that are suspicious for lung cancer. However, the NELSON algorithm views growth rate as a central determinant of when to intervene for lung nodules, including those that are solid [2, 10].

The volumetric NELSON algorithm, based on nodule growth, performed poorer than did clinical Lung-RADS categories. This likely reflects the presence of slow-growing malignancies in our cohort, which are treated in our clinical practice and thus were grouped with other cancers for the present analysis. However, as previously noted, whether treating such malignancies results in improved patient outcomes is unclear. Nevertheless, no demonstrable benefit was observed as a result of using the potentially time-consuming volumetric measurements in NELSON. Indeed, measurement of nodule volumes requires specialized software and additional time for nodule segmentation, which may present a barrier to implementing lung cancer screening in some radiology practices. Linear measurements are supported by Lung-RADS and may be obtained with far less effort, thus being advised for routine clinical care.

Our study has several limitations. First, it is a retrospective analysis with a small sample size. In particular, the number of cancers was small. Second, rather than including consecutive nodules, a sample enriched with higher category nodules was generated. However, the bulk of excluded nodules were category 1 or 2, and thus had no nodules or very low risk nodules and would have been unlikely to increase the number of cancers in the analysis. In the analysis, percentages were weighted to reflect the distribution of categories across all nodules with a follow-up lung cancer screening examination available. Third, since pathologic proof was not obtained for the benign nodules, some such nodules may represent undiagnosed cancers. However, as these patients were in a lung cancer screening program, they would be expected to have received regular follow-up, and any clinically significant lung cancer would likely have been diagnosed during the course of the study period. Of note, clinical follow-up was performed for more than two years since the final follow-up CT in the study sample. Fourth, the exact reasons for why particular Lung-RADS categories were assigned at the time of clinical interpretation are unknown. Finally, our analysis was performed at a single healthcare network, spanning two academic sites and one community site. Given the variability and radiologist discretion in aspects of Lung-RADS application, findings may differ at other institutions.

In conclusion, on follow-up lung cancer screening examinations, new nodules had intermediate cancer risk between stable existing and growing existing nodules. Clinical Lung-RADS exhibited excellent diagnostic performance for cancer detection on follow-up CT, aside from low specificity in existing nodules. A smaller optimal threshold diameter was identified for new than for existing nodules. Volumetric assessment exhibited poorer performance than did clinical Lung-RADS categories. Strict application of Lung-RADS criteria resulted in downgrading of numerous malignant nodules (typically on the basis of stability or pure ground glass appearance), supporting use of category 4X for nodules with suspicious features, even if stable. The findings provide insight into radiologists' management of nodules detected on follow-up lung cancer screening examinations.

## Acknowledgments

The authors did not receive any funding for this article.

## References

1. American College of Radiology (2019) Lung-RADS<sup>®</sup> Version 1.1. <https://www.acr.org/-/media/ACR/Files/RADS/Lung-RADS/LungRADSAssessmentCategoriesv1-1.pdf?la=en>. Accessed 21 Jun 2019
2. Xu DM, Gietema H, de Koning H, et al. (2006) Nodule management protocol of the NELSON randomised lung cancer screening trial. *Lung Cancer* 54:177–184. 10.1016/j.lungcan.2006.08.006 [PubMed: 16989922]
3. Walter JE, Heuvelmans MA, Oudkerk M (2017) Small pulmonary nodules in baseline and incidence screening rounds of low-dose CT lung cancer screening. *Transl Lung Cancer Res* 6:
4. MacMahon H, Austin JHM, Gamsu G, et al. (2005) Guidelines for Management of Small Pulmonary Nodules Detected on CT Scans: A Statement from the Fleischner Society. *Radiology* 237:395–400. 10.1148/radiol.2372041887 [PubMed: 16244247]
5. White CS, Dharaiya E, Dalal S, et al. (2019) Vancouver Risk Calculator Compared with ACR Lung-RADS in Predicting Malignancy: Analysis of the National Lung Screening Trial. *Radiology* 291:205–211. 10.1148/radiol.2018181050 [PubMed: 30667335]



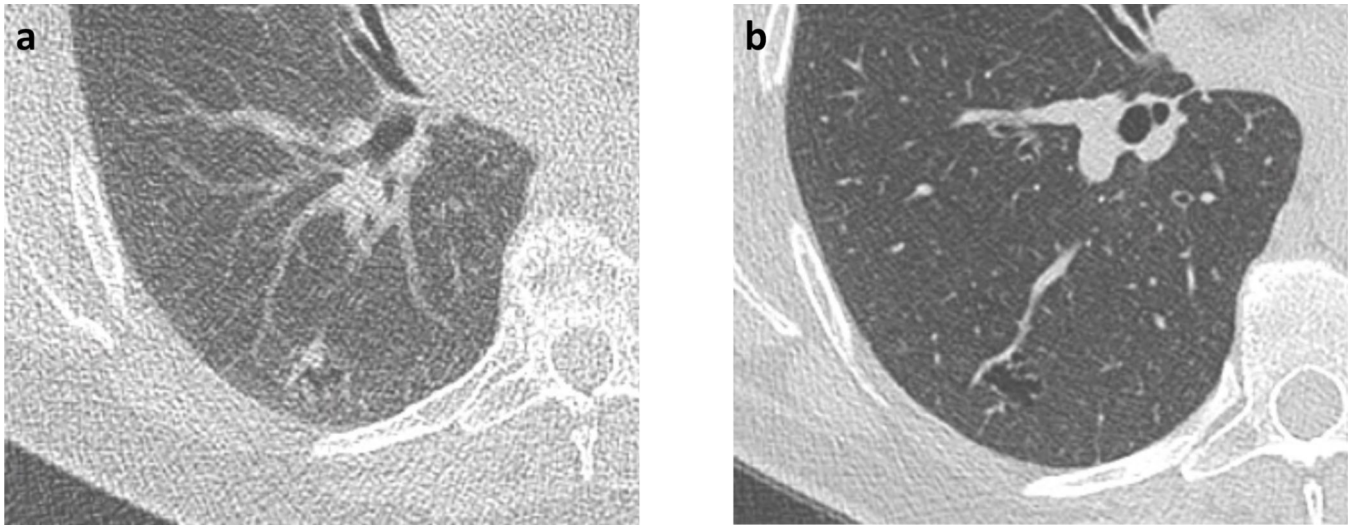
6. Hammer MM, Palazzo LL, Kong CY, Hunsaker AR (2019) Cancer Risk in Subsolid Nodules in the National Lung Screening Trial. *Radiology* 293:441–448. 10.1148/radiol.2019190905 [PubMed: 31526256]
7. Horeweg N, van Rosmalen J, Heuvelmans MA, et al. (2014) Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. *Lancet Oncol* 15:1332–1341. 10.1016/S1470-2045(14)70389-4 [PubMed: 25282285]
8. Harris PA, Taylor R, Thielke R, et al. (2009) Research Electronic Data Capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 42:377–381. 10.1016/j.jbi.2008.08.010 [PubMed: 18929686]
9. Hammer MM, Eckel AL, Palazzo LL, Kong CY (2021) Cost-Effectiveness of Treatment Thresholds for Subsolid Pulmonary Nodules in CT Lung Cancer Screening. *Radiology* 204418. 10.1148/radiol.2021204418
10. de Koning HJ, van der Aalst CM, de Jong PA, et al. (2020) Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. *N Engl J Med* 382:503–513. 10.1056/NEJMoa1911793 [PubMed: 31995683]

**Key Finding:**

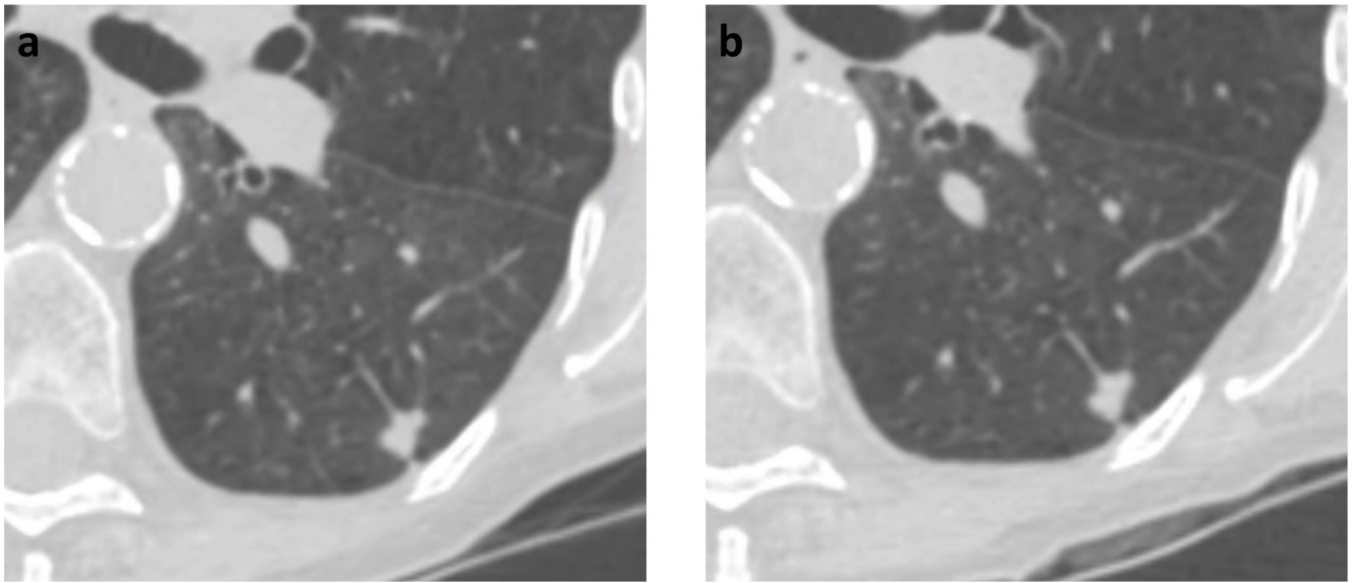
Cancer risk on follow-up screening examinations was 5% for new nodules, 1% for stable existing nodules, and 44% for growing existing nodules. AUC was 0.96 for clinical Lung-RADS categories versus 0.71–0.84 for volumetric NELSON categories. Application of strict Lung-RADS criteria downgraded 7 of 29 cancers to Lung-RADS category 2.

**Importance:**

Lung-RADS, as applied by radiologists in clinical practice, had excellent performance for cancer risk assessment on follow-up screening examinations. Strict Lung-RADS criteria downgrades some malignancies.



**Figure 1.** 58-year-old woman with right lower lobe nodule on (A) present lung cancer screening CT, and (B) lung cancer screening CT performed 2-years prior. The nodule was described as a growing pure ground glass nodule (associated with a cystic lesion), measuring 15 mm in mean diameter. The nodule was categorized as Lung-RADS 3 by the clinical report, though as Lung-RADS 2 by strict application of size criteria. Subsequent wedge resection demonstrated lepidic-predominant adenocarcinoma.



**Figure 2.**

77-year-old woman with left lower lobe nodule on (A) present lung cancer screening CT, and (B) lung cancer CT performed 3 months earlier. The nodule was described as a stable solid nodule measuring 8 mm in mean diameter. The nodule was categorized as Lung-RADS 4A by the clinical report, though as Lung-RADS 2 by strict application of size criteria. Subsequent biopsy demonstrated adenocarcinoma.

**Table 1.**

## Patient and nodule characteristics

Characteristic	All Nodules (n=185)	New Nodules (n= 81)	Existing Nodules (n=104)	P (New vs Existing)
Age (y), median (range)	66 (55 – 79)	65 (56 – 79)	67 (55 – 78)	.35
Sex				
Female	100 (54)	44 (54)	56 (54)	.53
Male	85 (46)	37 (46)	48 (46)	
Nodule attenuation				.44
Solid	140 (76)	65 (80)	75 (72)	
Part-solid	30 (16)	10 (12)	20 (19)	
Ground glass	15 (8)	6 (7)	9 (9)	
Nodule diameter (mm), median (range)	7 (2 – 30)	7 (2–29)	7 (2 – 30)	.78
Clinical Lung-RADS category				<.001
2	50 (27)	4 (5)	46 (44)	
3	45 (24)	30 (37)	15 (14)	
4A	47 (25)	27 (33)	20 (19)	
4B	30 (16)	18 (22)	12 (12)	
4X	13 (7)	2 (2)	11 (11)	
Lung cancer <sup>a</sup>	29 (3)	5 (5)	24 (2)	.60

Unless otherwise indicated, values represent number of patients, with percentage in parentheses.

<sup>a</sup>Weighted percentages based on relative frequencies of clinical LungRADS categories in entire lung cancer screening population from which study sample is drawn

**Table 2.**

Risk of lung cancer, stratified by Lung-RADS category

Category	Clinical Lung-RADS			Strict Lung-RADS <sup>a</sup>		
	All Nodules	New Nodules	Existing Nodules	All Nodules	New Nodules	Existing Nodules
2	0/50 (0)	0/4 (0)	0/46 (0)	7/83 (1)	1/10 (3)	7/73 (1)
3	4/45 (9)	3/30 (10)	1/15 (7)	0/18 (0)	0/18 (0)	0/0 (0)
4A	7/47 (15)	1/27 (4)	6/20 (30)	2/27 (9)	2/24 (10)	0/3 (0)
4B	8/30 (27)	0/18 (0)	8/12 (67)	10/44 (22)	1/27 (10)	9/17 (39)
4X	10/13 (77)	1/2 (50)	9/11 (82)	10/13 (77)	1/2 (50)	9/11 (82)

Values represent numerators and denominators, with percentage in parentheses.

<sup>a</sup>Percentages weighted by relative frequency of clinical Lung-RADS scores in the underlying population.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3-**

Characteristics of 29 cancers on follow-up CT examinations

Characteristic	All (n=29)	New (n=5)	Stable Existing (n=10)	Growing Existing (n=14)
Age (y), median (range)	66 (55–78)	66 (58–74)	71 (63–78)	63 (55–75)
Sex				
Female	17 (59)	2 (40)	7 (70)	8 (57)
Male	12 (41)	3 (60)	3 (30)	6 (43)
Nodule attenuation				
Solid	13 (45)	3 (60)	4 (40)	6 (43)
Part-solid	13 (45)	1 (20)	6 (60)	6 (43)
Ground glass	3 (10)	1 (20)	0 (0)	2 (14)
Nodule diameter (mm), median (range)	10 (6–25)	10 (6–19)	10 (8–25)	10 (6–21)
Clinical Lung-RADS category				
2	0 (0)	0 (0)	0 (0)	0 (0)
3	4 (14)	3 (60)	1 (10)	0 (0)
4A	7 (24)	1 (20)	4 (40)	2 (14)
4B	8 (28)	0 (0)	1 (10)	7 (50)
4X	10 (34)	1 (20)	4 (40)	5 (36)
Strict Lung-RADS category				
2	7 (24)	1 (20)	6 (60)	0 (0)
3	0 (0)	0 (0)	0 (0)	0 (0)
4A	2 (7)	2 (40)	0 (0)	0 (0)
4B	10 (34)	1 (20)	0 (0)	9 (64)
4X	10 (34)	1 (20)	4 (40)	5 (35)

Unless otherwise indicated, values represent number of patients, with percentage in parentheses.

**Table 4.**

Risk of lung cancer by NELSON category

NELSON Category	Reader 1	Reader 2
New		
NODCAT 2	0/10 (0)	0/13 (0)
NODCAT 3	1/46 (3)	3/48 (6)
NODCAT 4	1/17 (5)	1/18 (5)
Existing <sup>a</sup>		
GROWCAT A	10/71 (1)	8/78 (1)
GROWCAT B	5/10 (5)	2/6 (9)
GROWCAT C	12/31 (10)	15/22 (64)

Values represent numerators and denominators, with percentage in parentheses.

<sup>a</sup>Growth categories correspond to volume doubling times of >600 days (category A), 400–600 days (category B), or 0–399 days (category C). Percentages are weighted by relative frequency of clinical Lung-RADS scores in the underlying population.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 5.**

Diagnostic performance of nodule evaluation approaches

Approach	All Nodules			New Nodules		Existing Nodules	
	AUC <sup>a</sup> (95% CI)	Sens <sup>a</sup>	Spec <sup>a</sup>	Sens <sup>a</sup>	Spec <sup>a</sup>	Sens <sup>a</sup>	Spec <sup>a</sup>
Clinical Lung-RADS <sup>b</sup>	0.96 (0.93–0.98)	29/29 (100)	50/156 (85)	5/5 (100)	4/76 (41)	24/24 (100)	46/80 (94)
Strict Lung-RADS <sup>b</sup>	0.81 (0.68–0.90)	22/29 (68)	76/156 (88)	4/5 (74)	9/76 (38)	18/24 (65)	67/80 (98)
NELSON <sup>c</sup>							
Reader 1	0.71 (0.60–0.83)	19/29 (64)	71/156 (75)	3/5 (48)	16/76 (46)	16/24 (71)	55/80 (81)
Reader 2	0.84 (0.73–0.92)	21/29 (75)	100/156 (88)	5/5 (100)	16/76 (46)	16/24 (65)	67/80 (96)
Nodule diameter <sup>d</sup>	0.89 (0.82–0.95)	26/29 (86)	100/156 (84)	5/5 (100)	25/76 (62)	23/24 (96)	56/80 (85)

Unless otherwise indicated, values represent numerators and denominators, with percentage in parentheses.

<sup>a</sup>Weighted based on relative frequencies of clinical LungRADS categories in entire lung cancer screening population from which study sample is drawn

<sup>b</sup>Considered positive at category of 3 or higher

<sup>c</sup>Considered positive at NODCAT 2 or higher, or at GROWCAT B or higher

<sup>d</sup>Based on Youden index from ROC analysis, considered positive at 8 mm in all nodules, 6 mm in new nodules, and 8 mm in existing nodules

Sens = sensitivity. Spec = specificity.