



Published in final edited form as:

JAMA Intern Med. 2014 June ; 174(6): 871–880. doi:10.1001/jamainternmed.2014.561.

Resource Use and Guideline Concordance in Evaluation of Pulmonary Nodules for Cancer: Too much and too little care

Renda Soylemez Wiener, MD, MPH^{1,2,3}, Michael K. Gould, MD, MS⁴, Christopher G. Slatore, MD, MS^{5,6}, Benjamin G. Fincke, MD^{1,7}, Lisa M. Schwartz, MD, MS^{3,8}, and Steven Woloshin, MD, MS^{3,8}

¹Center for Healthcare Organization & Implementation Research, ENRM VA Hospital, Bedford, MA

²The Pulmonary Center, Boston University School of Medicine, Boston, MA

³The Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth Medical School, Hanover, NH

⁴Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA

⁵Health Services Research & Development, Portland VA Medical Center, Portland, OR

⁶Division of Pulmonary & Critical Care Medicine, Oregon Health & Science University, Portland, OR

⁷Department of Health Policy & Management, Boston University School of Public Health, Boston, MA

⁸VA Outcomes Group, Department of Veterans Affairs, White River Junction, VT

Abstract

Importance—Pulmonary nodules are common, and more will be found with implementation of lung cancer screening. How potentially malignant pulmonary nodules are evaluated may affect

Corresponding Author: Renda Soylemez Wiener, MD, MPH, Center for Healthcare Organization & Implementation Research, ENRM VA Hospital, 200 Springs Road, Bldg 70 (152), Bedford, MA 01730; phone, (781) 687-2872; fax, (781) 687-3106; rwiener@bu.edu. An earlier version of this work was presented at the VA HSR&D National Conference; July 18, 2012; National Harbor, Maryland.

Conflict of interest disclosures: The authors do not have any conflict of interests to disclose.

Previous presentation: An earlier version of this work was presented at the VA HSR&D National Conference; July 18, 2012; National Harbor, Maryland.

The authors have no conflicts of interest. This study was funded by the VA HSR&D (PPO 08-401).

Author Contributions: Dr. Wiener had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Wiener, Schwartz, Woloshin.

Acquisition of data: Wiener.

Analysis and interpretation of data: Wiener, Gould, Slatore, Fincke, Schwartz, Woloshin.

Drafting of the manuscript: Wiener.

Critical revision of the manuscript for important intellectual content: Wiener, Gould, Slatore, Fincke, Schwartz, Woloshin.

Statistical analysis: Wiener.

Obtained funding: Wiener, Schwartz, Woloshin.

Study supervision: Schwartz, Woloshin.

Disclaimer: The views expressed herein do not necessarily represent the views of the funding agencies, the Department of Veterans Affairs, or the United States government.

patient outcomes, healthcare costs, and effectiveness of lung cancer screening programs. Guidelines for evaluating pulmonary nodules for cancer exist, but little is known about how nodules are evaluated in the usual care setting.

Objective—To characterize nodule evaluation and concordance with guidelines

Design—Retrospective cohort study, including detailed chart review from nodule detection through evaluation completion, cancer diagnosis, or study end (12/31/12)

Setting—Fifteen VA hospitals

Participants—300 adult patients with pulmonary nodules

Main outcomes and measures—Resources used for evaluation at any VA facility; guideline-concordant evaluation.

Results—Among 300 patients with pulmonary nodules, 9% (27/300) were ultimately diagnosed with lung cancer: 1/57 (2%) with a nodule \leq 4 mm, 4/134 (3%) with a nodule 5–8 mm, and 20% (22/109) with a nodule $>$ 8 mm. Nodule evaluation entailed 1044 imaging studies, 147 consultations, 76 biopsies, 13 resections, and 21 hospitalizations. Radiographic surveillance (n=277) lasted a median of 13 months, but ranged from $<$ 1 month to 8.5 years. Forty-six patients underwent invasive procedures (range per patient 1–4): 42% (19/46) did not have cancer, and 17% (8/46) experienced complications, including one death. Notably, 15/300 (5%) “fell through the cracks,” without purposeful evaluation or obvious reason for deferral. Among 197 patients with a nodule detected after release of the Fleischner Society guidelines, 45% received care inconsistent with guidelines (18% over-evaluation, 27% under-evaluation). In multivariable analyses, the strongest predictor of guideline-inconsistent care was inappropriate radiologist recommendations (over-evaluation: RR 4.5, [95% CI 2.3–8.7]; under-evaluation: RR 4.3 [2.7–6.8]). Other system factors associated with under-evaluation included receiving care at more than one facility (RR 2.0 [1.5–2.7]) and nodule detection during an inpatient or preoperative visit (RR 1.6 [1.1–2.5]).

Conclusions and Relevance—Pulmonary nodule evaluation is often inconsistent with guidelines, including cases with no work-up and others with prolonged surveillance or unneeded procedures that may cause harm. Systems to improve quality (e.g., aligning radiologist recommendations with guidelines; facilitating communication across providers) are needed before lung cancer screening is widely implemented.

It has been estimated that hundreds of thousands of pulmonary nodules are detected each year,¹ and more will be found now that the US Preventive Services Task Force recommends annual computed tomography (CT) lung cancer screening among high-risk individuals.² Pulmonary nodules present a diagnostic challenge: some are cancers, but most are not. To rule out malignancy, evaluation typically entails radiographic surveillance; some patients also undergo invasive procedures (biopsy and/or surgical resection).³ Unfortunately, surveillance subjects patients to prolonged uncertainty, anxiety, and radiation, while invasive testing can cause physical complications.^{4–10}

Consequently, pulmonary nodule evaluation has potential to create a tremendous burden on individual patients and the healthcare system. In 2003, the American College of Chest Physicians (ACCP) guidelines recommended limiting surveillance to 2 years for most

patients.¹¹ The 2005 Fleischner Society guidelines sought to further reduce the burden of evaluation by recommending that patients at lower risk of cancer (non-smokers or those with smaller nodules) receive fewer tests.¹² The ACCP guidelines were updated in 2007 (and in 2013) to match the Fleischner Society algorithms.^{3, 13} These guidelines apply to incidental and screening-detected nodules.

Despite the existence of guidelines, little is known about how pulmonary nodules are managed in practice. The intensity of evaluation has important implications for patient health, costs, and effectiveness of lung cancer screening programs.^{14, 15} Although the management of screening-detected nodules in clinical trials has been reported,^{16–20} it is unclear whether management in usual care settings will reflect the trials or guideline recommendations. We therefore addressed these 3 questions in a representative sample of 300 Veterans evaluated in 15 VA facilities: First, what resources are used to evaluate potentially malignant pulmonary nodules? Second, is evaluation consistent with guideline recommendations? Finally, what harms are associated with nodule evaluation?

METHODS

We performed a retrospective cohort study based on review of medical records of patients with indeterminate (not known to be malignant or benign) pulmonary nodules. The White River Junction and ENRM VA Hospital institutional review boards approved this study.

Population

Our goal was to identify a representative cohort of 300 Veterans with “typical” indeterminate pulmonary nodules for which nodule evaluation guidelines would apply (Figure 1). We included patients whose nodule was detected between 1/1/03-12/31/06 because 1) the start date coincides with the publication of the ACCP guidelines recommending surveillance be limited to 2 years,¹¹ providing an upper bound for expected duration of surveillance; 2) the period encompasses the release of the Fleischner Society guidelines,¹² allowing us to assess their impact on intensity of evaluation; 3) these dates allowed an extended period (6–10 years) from nodule detection through end of chart review (12/31/12), allowing us to capture cases of prolonged surveillance.

Figure 1 illustrates the steps used to assemble our cohort. First, we created a VA FileMan algorithm to search for text strings in radiology reports of all chest x-rays and CT scans performed at two VA facilities in 2006. Our algorithm, designed to be more sensitive than specific, searched for “nodul” and “mass,” and then discarded reports containing the phrase “no pulmonary nod.” If the 2006 report (“flagged study”) was a follow-up test for a nodule detected previously, we worked backwards in the VA’s integrated electronic medical record (VistaWeb) until we identified the study corresponding to the first detection of the nodule (“index study”), which may have been performed at another facility.

Steps 2 and 3 were performed iteratively until we reached our target sample of 300. We randomly selected patients from our initial cohort (n=2366) for manual review of the radiology report to confirm presence of a pulmonary nodule. We then conducted a limited

chart review to confirm eligibility. Exclusion criteria were designed to eliminate patients for whom guidelines for nodule evaluation would not apply (Figure 1).

Data Abstraction

We developed a standardized data abstraction form to capture baseline (i.e., at time of index study) patient and nodule characteristics, events during evaluation (change in nodule size, appearance of new nodule, transfer of care to another VA facility), resources used for evaluation at any facility, and patient outcomes (final diagnosis, complications of invasive procedures). Two trained individuals reviewed medical records in duplicate; the lead investigator (RSW) resolved discrepancies.

We recorded all resources used for nodule evaluation: imaging studies (x-ray, CT, positron emission tomography [PET]), consultations, biopsy procedures (bronchoscopy, transthoracic needle biopsy, mediastinoscopy, other [e.g., peripheral lymph node biopsy]), pre-procedure testing (cardiac stress test, pulmonary function tests), hospitalizations, and surgical resection. For each resource used, we graded how likely it was to be related to nodule evaluation: definitely (e.g., indication for CT listed as “follow-up pulmonary nodule”), probably (e.g., pulmonary consultation for “abnormal CT” in a patient with a newly detected nodule; chest x-rays or CT scans performed during a hospitalization associated with nodule evaluation), or possibly (e.g., follow-up CT ordered at an interval consistent with nodule evaluation, but without mention of nodule in indication). Resources graded as “possibly related” were excluded from our analyses. For cases in which no purposeful nodule evaluation occurred, we recorded possible reasons for deferral (e.g., severe comorbidity, patient refusal).

We set pre-determined stopping rules for data abstraction: 1) nodule ruled out (e.g., “nodule” seen on x-ray resolved as nipple shadow on subsequent CT); 2) cancer diagnosed (lung or other primary); 3) patient died or lost from VA system (i.e., no further visits at any VA site). If none of these criteria were met, we reviewed all records through 12/31/12 (6–10 years after nodule detection), recording any resources used for nodule evaluation.

Outcomes

Our main outcomes were resources used for nodule evaluation and the proportion that received evaluation consistent with the Fleischner Society guidelines.¹² Concordance with guideline recommendations was determined by two experts in pulmonary nodule evaluation³; RSW reviewed every case to make the initial determination, and borderline cases were resolved through discussion with MKG. We also categorized radiologist recommendations as being concordant with, more intensive than, or less intensive than guideline recommendations. More intensive evaluation (“over-evaluation”) could include more frequent testing than recommended, more prolonged surveillance than recommended, or performance of tests that are not recommended (e.g., PET or biopsy for nodules <8 mm); conversely less intensive evaluation (“under-evaluation”) encompassed delays or failures to obtain recommended tests. Because the 2005 Fleischner Society and 2007 ACCP guidelines were ambiguous in the recommended duration of surveillance for sub-solid (i.e., ground

glass or part-solid) nodules, we used nodule size to determine the appropriate duration of surveillance for all nodules, regardless of attenuation.

As secondary outcomes, we assessed the impact of the release of the Fleischner Society guidelines on intensity of nodule evaluation and factors associated with evaluation intensity.

Analyses

All analyses were performed in STATA 10.1 (College Station, TX). Because guidelines for nodule evaluation are based on nodule size, we reported baseline patient characteristics and outcomes stratified by nodule size at the time of the index study, comparing medians and proportions using the Kruskal-Wallis, Chi-square, or Fisher's exact test as appropriate. For all analyses, we used two-sided alpha 0.05 as the threshold for statistical significance.

Resource use—We calculated the number and proportion of patients who underwent each type of evaluation, the median number and interquartile range (IQR) or full range of tests or services performed per patient, and the total number performed among all patients.

Concordance of evaluation with guideline recommendations—For our primary analysis, we calculated the proportions of patients who received guideline-concordant evaluation, over-evaluation, and under-evaluation among the subset of patients with a pulmonary nodule detected after release of the Fleischner Society guidelines (index study on or after 11/1/2005). As a secondary analysis to examine the impact of guidelines on evaluation intensity, we compared the proportion that received over-evaluation (using the Fleischner Society algorithm as our benchmark) among patients with a nodule detected before versus after release of these guidelines.

Factors associated with evaluation intensity—Using bivariate logistic regression, we explored factors associated with intensity of evaluation, including: patient characteristics (age, tobacco use, chronic obstructive pulmonary disease [COPD], symptoms suggestive of lung cancer), nodule characteristics (index size, ground-glass attenuation [i.e., subsolid nodules], spiculation, upper lobe location), and evaluation characteristics (index x-ray [vs CT], index study during a pre-operative or inpatient visit, nodule evaluation at more than one facility, development of a new nodule during surveillance, radiologist recommendations for guideline-inconsistent evaluation [e.g., recommendation for more intensive evaluation, as compared with guideline-consistent, less intensive evaluation, or no explicit recommendation], patient refusal). Factors associated with evaluation intensity on bivariate analysis (based on $p < 0.20$) were included in multivariable models. Because outcomes were common ($> 10\%$), we estimated relative risks and 95% confidence intervals (CI) using a Poisson regression model clustered by facility.

RESULTS

In this representative sample of 300 Veterans with an indeterminate pulmonary nodule, 57 (19%) had a nodule ≤ 4 mm on the index study, 134 (45%) had a 5–8 mm nodule, and 109 (36%) had a nodule > 8 mm (Table 1). The sample was typical of the VA population: mostly men who had smoked. Most nodules were incidentally detected and did not have features

associated with malignancy. Ultimately, 27 (9%) patients were diagnosed with lung cancer; the likelihood of cancer was significantly associated with nodule size ($p < 0.001$).

Resource use

These 300 patients underwent nodule evaluation at 15 VA facilities around the country (AZ, CA, CT, FL, GA, KY, MA, ME, NH, RI, TX, VT). More than 10% ($n=34$) underwent evaluation at more than one facility. Counting only faculty-level clinicians (not residents or fellows), over 300 clinicians at 15 sites were involved in guiding nodule evaluation, including 87 radiologists, 114 primary care providers, 51 pulmonologists, and 54 other clinicians (e.g., various consultants, inpatient clinicians who ordered follow-up testing).

Twenty-three patients (8%) received no apparent purposeful nodule evaluation. In 8 of these cases, a likely rationale was evident, such as severe comorbidities, patient refusal, or clinician notes dismissing the nodule as clinically insignificant (e.g., 1–2 mm nodule). However, the other 15 patients (5% of cohort) had no obvious reason for lack of follow-up, seemingly “falling through the cracks.” Fortunately, none of these 23 patients were determined to have lung cancer (at least through 12/31/12).

Among the 277 patients who received at least one follow-up test, nodule evaluation entailed substantial resource use, including 1044 imaging studies (292 chest x-rays, 710 chest CT scans, 42 PET scans), 147 consultations (101 pulmonary, 25 thoracic surgery, 21 other), 22 pre-invasive tests (17 pulmonary function tests, 5 cardiac stress tests), 76 biopsies (46 bronchoscopies, 11 transthoracic needle biopsies, 8 mediastinoscopies, 11 other biopsies), 13 resections (6 wedge resections, 7 lobectomies), and 21 hospitalizations (Table 2 and **eTable**). The median number of tests for nodule evaluation was 2 (IQR 1–5, range 1–32) among patients with benign nodules and 8 (IQR 4–14, range 2–24) among patients who were diagnosed with lung cancer ($p=0.0001$).

Most patients (277/300, 92%) underwent at least one follow-up imaging study (Table 2). The median duration of surveillance was 13 months (IQR 3–33 months, range <0.5 months–8.5 years). After excluding patients whose nodules were ruled out (i.e., no longer present) on a subsequent imaging study, median duration of surveillance was 26 months (IQR 10–40 months) and did not differ by baseline nodule size ($p=0.22$). The median duration of surveillance was 11 months (range <1–51 months) among the 27 patients ultimately diagnosed with lung cancer.

In our sample, 15% (46/300) underwent invasive testing (Table 2); 42% (19/46) did not have lung cancer. Among patients who underwent biopsy, the median number of biopsies was 1 (range 1–4), but 20% (9/46) underwent 3 or more biopsy procedures before a diagnosis was established (Table 2 and **eTable**). Thirteen patients underwent surgical resection for presumptive malignancy; 31% (4/13) had benign pathology. Eight patients (17%) who underwent invasive procedures suffered a total of 11 complications, including 7 pneumothoraces (5 serious enough to require hospitalization), 2 hemorrhages (one led to hospitalization), and 2 pneumonias (both required hospitalization; one patient died). Among patients who underwent invasive procedures for a benign nodule, 4/19 (21%) experienced complications.

Figure 2 depicts the evaluation process and outcomes of 197 patients whose nodules were detected after release of the Fleischner Society guidelines. This figure plots the duration of surveillance, which extended well beyond the recommended 2-year period in many cases, and highlights the complexity of evaluation among patients who required more than simple surveillance; in many cases, these patients shuttled back and forth between surveillance, PET scan, and invasive testing, including multiple biopsies.

Concordance with guidelines and factors associated with evaluation intensity

Among the 197 patients with a nodule detected after release of the Fleischner Society guidelines, 55% received guideline-concordant care, 18% received over-evaluation, and 27% received under-evaluation (Figure 3). Over-evaluation was inversely associated with baseline nodule size (44% for nodules ≤ 4 mm, 15% for 5–8 mm nodules, 11% for nodules >8 mm, $p=0.001$; **eFigure**). The Fleischner Society guidelines appear to have achieved the stated purpose of reducing the burden of nodule evaluation¹²: over-evaluation was far more common among patients with a nodule detected before (versus after) publication of the guidelines (57% vs 18%, $p<0.001$). Radiologist recommendations were often consistent with guidelines (81%); when radiologist recommendations deviated from guidelines, they were far more likely to recommend more intensive (16%) than less intensive (2%) evaluation.

Regardless of whether radiologist recommendations were consistent with guidelines, the intensity of nodule evaluation reflected the intensity of radiologist recommendations in over 60% of cases (Figure 3). In multivariable analyses (Table 3), radiologist recommendations for overly intensive evaluation (RR 4.5, 95% CI 2.3–8.7) and nodule detection by CT rather than x-ray (RR 1.8, 95% CI 1.2–2.7) were both significantly associated with over-evaluation. Meanwhile, older age (RR 1.02 per year, 95% CI 1.00–1.03), nodule detection during an inpatient or preoperative visit (RR 1.6, 95% CI 1.1–2.5), nodule evaluation at more than one facility (RR 2.0, 95% CI 1.5–2.7), radiologist recommendations for less intensive evaluation (RR 4.3, 95% CI 2.7–6.8), and patient refusal (RR 2.9, 95% CI 1.3–6.3) were all significantly associated with under-evaluation.

DISCUSSION

To our knowledge, this is the first study to describe pulmonary nodule evaluation in the United States in the usual care setting (i.e., not in the context of a lung cancer screening study^{17–20} or dedicated pulmonary nodule clinic²¹). The strength of our study is the characterization of complete episodes of pulmonary nodule evaluation over a period of several years and their relationship to guideline recommendations. We found that evaluation of pulmonary nodules for malignancy consumed substantial resources and was often inconsistent with guideline recommendations. Many patients (18%) received over-evaluation, including cases of prolonged surveillance and multiple biopsies, exposing them to unneeded radiation (which confers a small but cumulative risk of radiation-induced cancers)⁶ and the potential for physical⁹ and emotional^{5, 8, 10} harms. Meanwhile, others (27%) received less intensive evaluation than guidelines recommend – or no work-up at all, exposing them to the possibility of delayed cancer diagnosis.

Our findings are similar to the results of the only other research assessing pulmonary nodule evaluation in the usual care setting: a French group found tremendous variation in nodule evaluation²² and associated resource use.²³ Similarly, in physician surveys evaluation choices were highly variable.²⁴ Even in a dedicated pulmonary nodule clinic, 45% of patients did not complete the recommended duration of surveillance, highlighting the difficulty of achieving guideline-concordant care.²¹

When exploring reasons for non-concordance with guidelines, we found radiologist recommendations to be the strongest predictor of evaluation intensity. In our sample, radiologist recommendations were inconsistent with guidelines in 18% of cases (16% more intensive, 2% less intensive). Other studies have found an even higher rate of non-concordance between guidelines and radiologist recommendations, which may reflect the fact that VA facilities are typically academic affiliates. In national surveys, 39–73% of radiologists recommended care that differed from guidelines in response to pulmonary nodule vignettes.^{25, 26} Similarly, in audits of radiology reports, 21–66% recommended nodule evaluation inconsistent with guidelines.^{27, 28} Mirroring our findings, these studies found that radiologists erred on the side of recommending over-evaluation.^{25, 26, 28, 29} Standardizing radiology report templates to include the Fleischner Society algorithm for nodule evaluation may help improve concordance with guideline recommendations.^{30, 31}

The other modifiable systems factors associated with receipt of inappropriate evaluation (in this case, under-evaluation) were initial nodule detection in the inpatient or preoperative setting – in other words, having a nodule identified by a provider who would not be the one to direct subsequent nodule evaluation – and undergoing evaluation at more than one facility. Both likely reflect a failure in communication between care teams, a factor cited by the Institute of Medicine as one of the largest barriers to high-quality medical care.³² To improve the quality of nodule evaluation and reduce delays in lung cancer diagnosis, systems should be implemented to notify not only the ordering provider, but also the primary care provider, when a pulmonary nodule is identified. Another successful model is to appoint a dedicated clinician (often a mid-level provider) who is notified whenever a new pulmonary nodule is detected, maintains a registry of these patients, and coordinates their evaluation.³³ Finally, during care transitions (e.g., inpatient to outpatient, one facility to another), explicit summaries of ongoing healthcare issues and pending action items may help avoid delays or gaps in care.^{34, 35}

The appropriateness of nodule evaluation will affect the cost-effectiveness and risk-benefit trade-offs of CT lung cancer screening.^{14, 15} In the NLST, 42% of invasive procedures were performed in patients with benign nodules,³⁶ identical to the rate in our study. Similarly, in the NLST, 44% of patients who underwent resection had a benign nodule,³⁶ whereas in our study 31% of patients who underwent resection had a benign nodule. Thus, at least in this sample, triage of patients for invasive testing in these VA facilities reflects similar decision-making as described in studies of lung cancer screening. The inefficiencies we observed related primarily to imaging, including both over- and under-utilization. While the per unit cost of imaging studies is lower than those associated with invasive testing, the aggregate cost of overuse of imaging may be substantial.³⁷

Both over- and under-evaluation may have important implications for patient outcomes. Many patients in our study underwent multiple biopsies, and complications following invasive procedures were more common in our study than in the NLST¹⁷: 17% vs 10%. Of note, our 17% complication rate closely approximates the 16% rate of complications following transthoracic needle biopsy of pulmonary nodules in 4 states.⁹ The NLST's 10% complication rate is likely unrealistically low, reflecting both the inclusion of highly skilled comprehensive cancer centers as study sites and the healthy participant bias common to trials. Although our numbers were too small to reliably estimate effects on patient outcomes, 27% of our sample received under-evaluation. Delays or outright failures to obtain appropriate tests may prolong the time to lung cancer diagnosis (which may or may not affect outcomes such as resectability and lung cancer mortality).

Our study has limitations. We evaluated 300 episodes of pulmonary nodule evaluation in the VA system, which may not represent care at other sites. While we found substantial variation in nodule evaluation, with instances of both under- and over-evaluation, it is possible that there is less variation in the single payer, academically affiliated VA system, which has an integrated electronic medical record, than in the broader community. Although we were able to capture evaluation conducted at any VA site, we may have missed evaluation studies that were performed in the private sector and never documented in the VA record. Because we were limited to the information in the record, we had to make inferences when notes were not explicit. This may have resulted in mis-attribution of resources, including both failure to count resources used for nodule evaluation and ascribing resources to nodule evaluation that were intended for another purpose. Any mis-attribution may also have affected whether episodes of nodule evaluation were classified as concordant with guidelines. In particular, there may have been reasons that were not documented that would explain why some patients received no apparent nodule evaluation (e.g., verbal communication between the radiologist and treating clinician that a small nodule appeared to be an intrapulmonary lymph node); any such cases would have resulted in erroneous conclusions that the patient “fell through the cracks” when in fact evaluation was purposely deferred. While we found radiologists' recommendations to be a very strong predictor of care received, it should be noted that both the radiologists' recommendations and the care received were categorized in relationship to the guidelines. Finally, while we targeted nodules identified between 2003–2006 to capture a period during which guidelines were introduced and to allow a prolonged period for follow-up, care may differ for nodules detected in more recent years.

Overall we observed a substantial burden of pulmonary nodule evaluation at the patient and healthcare system levels. This study raises questions about the expected cost-effectiveness and risk-benefit trade-offs of lung cancer screening in the usual care setting. Systems to improve efficiency and safety of nodule evaluation are needed – especially before wide-scale adoption of lung cancer screening. Possible solutions that warrant further exploration are inclusion of the Fleischner Society algorithm in the radiologist reports describing a pulmonary nodule, and improved systems to communicate findings of a new nodule and current stage of evaluation between care teams.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Rick Hines, Amanda Donaher, and Jim Rothlender, MD, who created the VA FileMan algorithms to assemble our study cohorts; as well as Terry Kneeland, MPH and Elaine Hickey, RN, who were instrumental in data abstraction for the initial chart review. We would like to recognize the contribution of our colleagues in the VA Outcomes Group and the writers' group of the Center for Healthcare Organization & Implementation Research, whose voluntary feedback enhanced both our thinking and the presentation of our results.

Funding / Support: This study was supported by the VA HSR&D (PPO 08-401) and with resources from the White River Junction VA Medical Center, White River Junction, VT; ENRM VA Hospital, Bedford, MA; and Portland VA Medical Center, Portland, OR. Dr. Wiener is also supported by a career development award from the National Cancer Institute (K07 CA138772), and Dr. Slatore is supported by a VA HSR&D Career Development Award.

Role of the sponsor: This study was funded by the Department of Veterans Affairs (VA HSR&D PPO 08-401). The funding organization had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

References

- Ost D, Fein AM, Feinsilver SH. Clinical practice. The solitary pulmonary nodule. *N Engl J Med*. 2003; 348(25):2535–2542. [PubMed: 12815140]
- Moyer VA. Screening for Lung Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. 2013 Dec 31. [Epub ahead of print]. 10.7326/M13-2771
- Gould MK, Donington J, Lynch WR, et al. Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*. May; 2013 143(5 Suppl):e93S–120S. [PubMed: 23649456]
- Humphrey L, Deffebach M, Pappas M, et al. Screening for Lung Cancer with Low-dose Computerized Tomography: A Systematic Review to Update the U.S. Preventive Services Task Force Recommendation. *Ann Intern Med*. 2013; 159(6):411–20. [PubMed: 23897166]
- Slatore CG, Press N, Au DH, Curtis JR, Wiener RS, Ganzini L. What the heck is a “nodule”? A qualitative study of veterans with pulmonary nodules. *Ann Am Thorac Soc*. 2013; 10(4):330–335. [PubMed: 23952851]
- Smith-Bindman R, Miglioretti DL, Johnson E, et al. Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems, 1996–2010. *JAMA*. Jun 13; 2012 307(22):2400–2409. [PubMed: 22692172]
- Tukey MH, Wiener RS. Population-based estimates of transbronchial lung biopsy utilization and complications. *Respir Med*. Nov; 2012 106(11):1559–1565. [PubMed: 22938740]
- Wiener RS, Gould MK, Woloshin S, Schwartz LM, Clark JA. “The thing is not knowing”: Patients' perspectives on surveillance of an indeterminate pulmonary nodule. *Health Expect*. 2012 Dec 16. [Epub ahead of print]. 10.1111/hex.12036
- Wiener RS, Schwartz LM, Woloshin S, Welch HG. Population-based risk of complications after transthoracic needle lung biopsy of a pulmonary nodule: an analysis of discharge records. *Ann Intern Med*. 2011; 155(3):137–144. [PubMed: 21810706]
- Wiener RS, Gould MK, Woloshin S, Schwartz LM, Clark JA. “What do you mean, a spot?": A qualitative analysis of patients' reactions to discussions with their physicians about pulmonary nodules. *Chest*. 2013; 143:672–677. [PubMed: 22814873]
- Tan BB, Flaherty KR, Kazerooni EA, Iannettoni MD. The solitary pulmonary nodule. *Chest*. Jan; 2003 123(1 Suppl):89S–96S. [PubMed: 12527568]

12. MacMahon H, Austin JH, Gamsu G, et al. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. *Radiology*. 2005; 237(2):395–400. [PubMed: 16244247]
13. Gould MK, Fletcher J, Iannettoni MD, et al. Evaluation of patients with pulmonary nodules: when is it lung cancer?: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest*. 2007; 132:108S–130S. [PubMed: 17873164]
14. Silvestri GA. Screening for lung cancer: it works, but does it really work? *Ann Intern Med*. 2011; 155(8):537–539. [PubMed: 21893614]
15. McMahon PM, Kong CY, Bouzan C, et al. Cost-effectiveness of computed tomography screening for lung cancer in the United States. *J Thorac Oncol*. Nov; 2011 6(11):1841–1848. [PubMed: 21892105]
16. Nair A, Hansell DM. European and North American lung cancer screening experience and implications for pulmonary nodule management. *Eur Radiol*. Dec; 2011 21(12):2445–2454. [PubMed: 21830100]
17. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. Aug 4; 2011 365(5):395–409. [PubMed: 21714641]
18. Pinsky PF, Marcus PM, Kramer BS, et al. Diagnostic procedures after a positive spiral computed tomography lung carcinoma screen. *Cancer*. 2005; 103(1):157–163. [PubMed: 15529306]
19. Libby DM, Smith JP, Altorki NK, Pasmantier MW, Yankelevitz D, Henschke CI. Managing the small pulmonary nodule discovered by CT. *Chest*. Apr; 2004 125(4):1522–1529. [PubMed: 15078769]
20. Crestanello JA, Allen MS, Jett JR, et al. Thoracic surgical operations in patients enrolled in a computed tomographic screening trial. *J Thorac Cardiovasc Surg*. Aug; 2004 128(2):254–259. [PubMed: 15282462]
21. Veeramachaneni NK, Crabtree TD, Kreisel D, et al. A thoracic surgery clinic dedicated to indeterminate pulmonary nodules: too many scans and too little pathology? *J Thorac Cardiovasc Surg*. Jan; 2009 137(1):30–35. [PubMed: 19154897]
22. Alzahouri K, Velten M, Arveux P, Woronoff-Lemsi MC, Jolly D, Guillemin F. Management of SPN in France. Pathways for definitive diagnosis of solitary pulmonary nodule: a multicentre study in 18 French districts. *BMC Cancer*. 2008; 8:93. [PubMed: 18402653]
23. Lemonnier I, Baumann C, Jay N, et al. Does the availability of positron emission tomography modify diagnostic strategies for solitary pulmonary nodules? An observational study in France. *BMC Cancer*. 2009; 9:139. [PubMed: 19426566]
24. Prosch H, Strasser G, Oschatz E, Schober E, Schneider B, Mostbeck GH. Management of patients with small pulmonary nodules: a survey of radiologists, pulmonologists, and thoracic surgeons. *AJR Am J Roentgenol*. Jul; 2006 187(1):143–148. [PubMed: 16794168]
25. Esmaili A, Munden RF, Mohammed TL. Small pulmonary nodule management: a survey of the members of the Society of Thoracic Radiology with comparison to the Fleischner Society guidelines. *J Thorac Imaging*. 2011; 26(1):27–31. [PubMed: 20622690]
26. Eisenberg RL, Bankier AA, Boiselle PM. Compliance with Fleischner Society guidelines for management of small lung nodules: a survey of 834 radiologists. *Radiology*. 2010; 255(1):218–224. [PubMed: 20308458]
27. Lacson R, Prevedello LM, Andriole KP, et al. Factors associated with radiologists' adherence to Fleischner Society guidelines for management of pulmonary nodules. *J Am Coll Radiol*. Jul; 2012 9(7):468–473. [PubMed: 22748786]
28. Masciocchi M, Wagner B, Lloyd B. Quality review: Fleischner criteria adherence by radiologists in a large community hospital. *J Am Coll Radiol*. May; 2012 9(5):336–339. [PubMed: 22554631]
29. Jeudy J, White CS, Munden RF, Boiselle PM. Management of small (3–5-mm) pulmonary nodules at chest CT: global survey of thoracic radiologists. *Radiology*. Jun; 2008 247(3):847–853. [PubMed: 18413885]
30. Woloshin S, Schwartz L, Dann EW, Black WC. Using Radiology Reports to Encourage Evidence-based Practice in the Evaluation of Small Incidentally Detected Pulmonary Nodules: A Preliminary Study. *Ann Am Thorac Soc*. 2014 in press.

31. Kahn CE Jr, Heilbrun ME, Applegate KE. From guidelines to practice: how reporting templates promote the use of radiology practice guidelines. *J Am Coll Radiol.* Apr; 2013 10(4):268–273. [PubMed: 23332496]
32. Institute of Medicine. *Crossing the quality chasm: A new health system for the 21st century.* Washington, D.C: National Academies Press; 2001.
33. Holden WE, Lewinsohn DM, Osborne ML, et al. Use of a clinical pathway to manage unsuspected radiographic findings. *Chest.* May; 2004 125(5):1753–1760. [PubMed: 15136387]
34. Health Quality Ontario. *Electronic tools for health information exchange: an evidence-based analysis.* *Ont Health Technol Assess Ser.* 2013; 13(11):1–76.
35. Rao S, Brammer C, McKethan A, Buntin MB. Health information technology: transforming chronic disease management and care transitions. *Prim Care.* Jun; 2012 39(2):327–344. [PubMed: 22608869]
36. Church TR, Black WC, Aberle DR, et al. Results of initial low-dose computed tomographic screening for lung cancer. *N Engl J Med.* May 23; 2013 368(21):1980–1991. [PubMed: 23697514]
37. Dunnick NR, Applegate KE, Arenson RL. The inappropriate use of imaging studies: a report of the 2004 Intersociety Conference. *J Am Coll Radiol.* May; 2005 2(5):401–406. [PubMed: 17411843]

Goal: To identify 300 patients with a “typical” indeterminate lung nodule detected within past 10 years

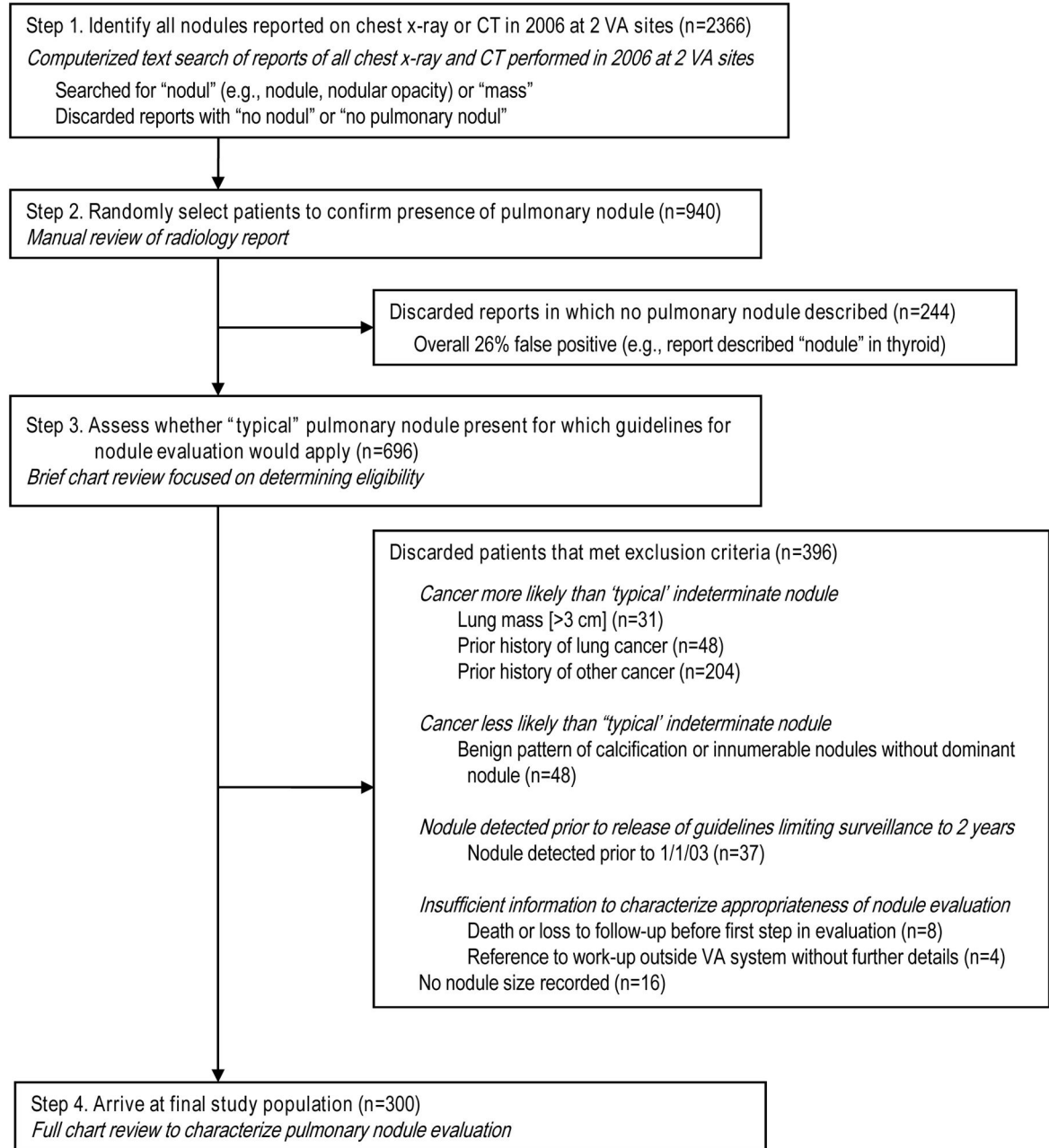
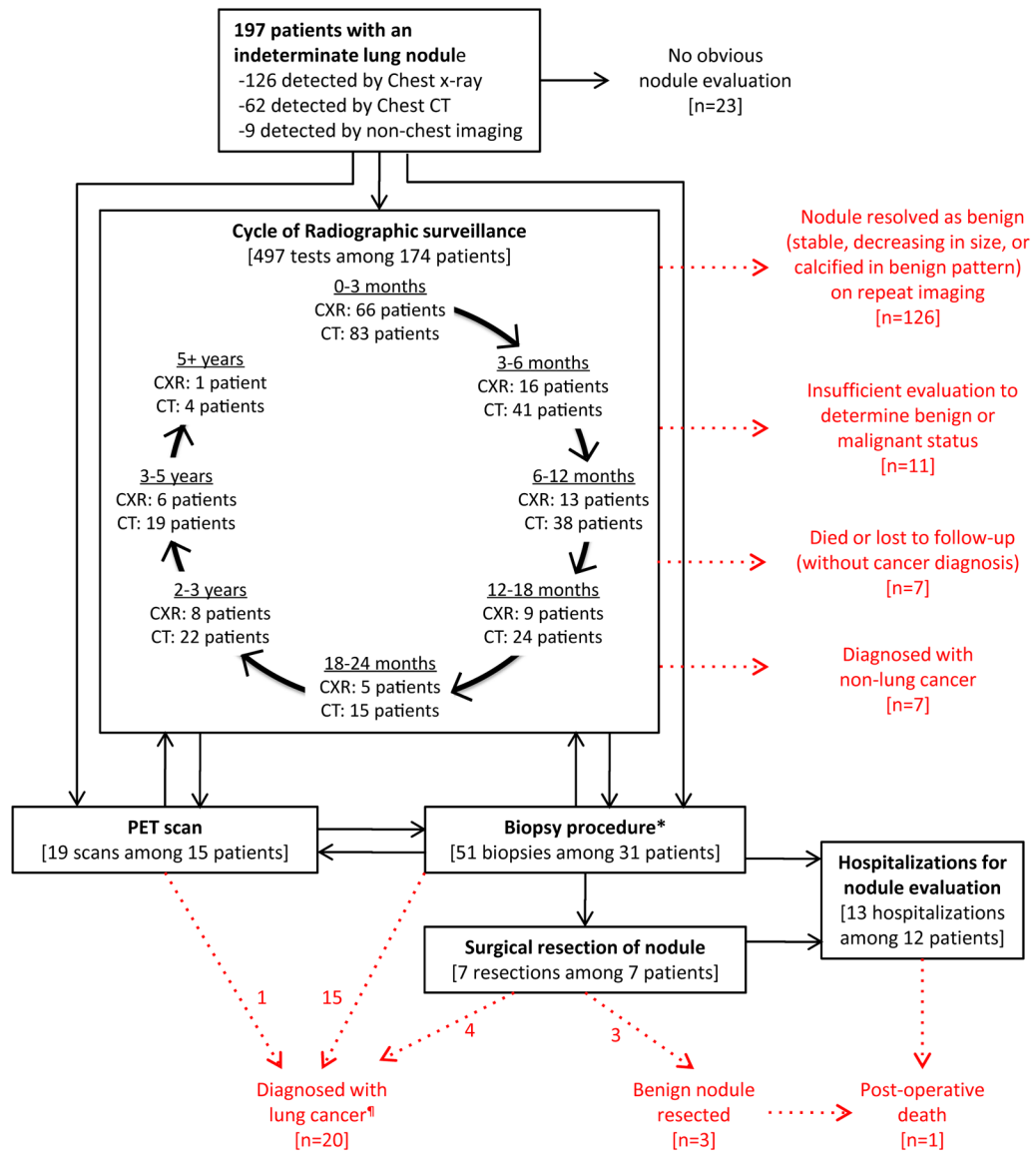


Figure 1.
Creation of final study population.



*Biopsy procedures included bronchoscopy, transthoracic needle biopsy, mediastinoscopy, and other biopsy (e.g., supraclavicular lymph node resection)

[†]One patient presumed to have cancer based on clinical presentation and PET findings alone
 Abbreviations: CT: computed tomography; CXR: chest x-ray; PET: positron emission tomography

Figure 2. Downstream evaluation and outcomes of 197 patients with a pulmonary nodule detected after release of Fleischner Society guidelines

*Biopsy procedures included bronchoscopy, transthoracic needle biopsy, mediastinoscopy, and other biopsy (e.g., supraclavicular lymph node resection)

[†]One patient presumed to have cancer based on clinical presentation and PET findings alone
 Abbreviations: CT: computed tomography; CXR: chest x-ray; PET: positron emission tomography

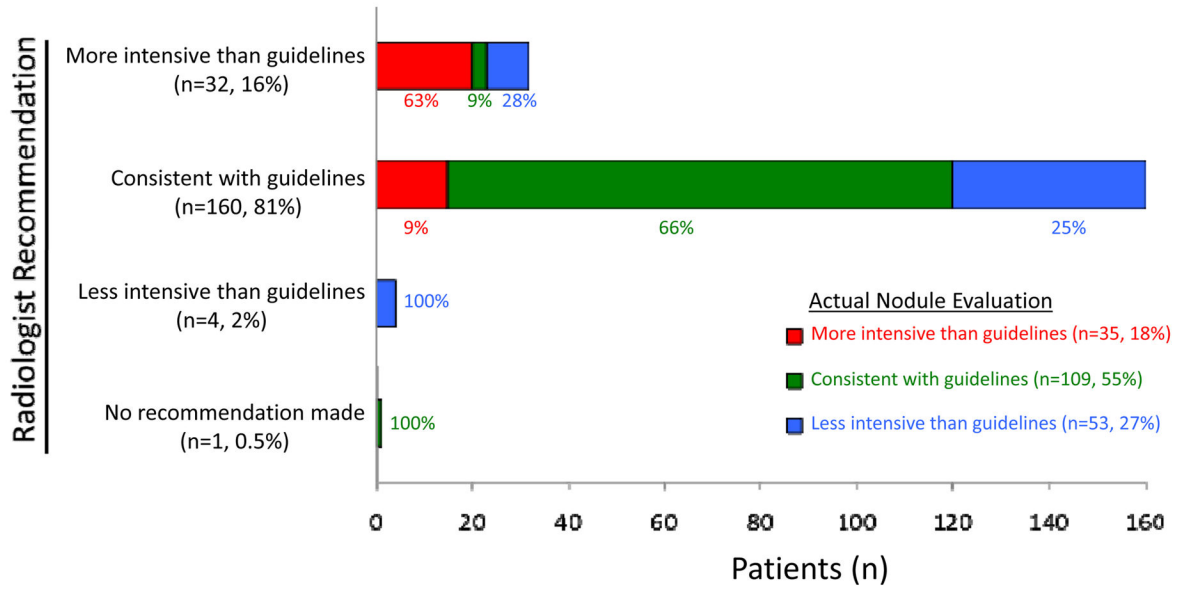


Figure 3. Relationship of radiologist recommendations and actual care received as compared to recommendations in the Fleischner Society Guidelines (among 197 patients with pulmonary nodule detected after guidelines released).

Table 1

Baseline patient characteristics and outcomes, stratified by size of nodule at time of detection

	Nodule Size			Overall [n=300]	P-value
	4 mm [n=57]	5-8 mm [n=134]	>8 mm [n=109]		
<i>Characteristics at nodule detection</i>					
Patient characteristics					
Age, years, mean (SD)	64 (12)	66 (11)	67 (11)	66 (11)	NS
Male, %	93%	93%	95%	94%	NS
Smoker, current or former, %	82%	85%	89%	86%	NS
COPD, %	58%	53%	54%	54%	NS
Nodule detected on x-ray (vs CT scan)	23%	59%	72%	57%	<0.001
Reason initial imaging obtained					NS
Symptoms suggestive of lung cancer (e.g. hemoptysis, weight loss)	18%	10%	13%	13%	
Symptoms not likely related to nodule	33%	44%	45%	42%	
No symptoms (e.g., preoperative film)	49%	45%	42%	45%	
Nodule characteristics					
Ground glass, %	11%	15%	13%	13%	NS
Spiculated, %	0%	4%	26%	11%	<0.001
Upper lobe location, %	35%	33%	40%	36%	NS
<i>Outcomes by end of data abstraction period*</i>					
Nodule resolved as benign	46 (81%)	104 (78%)	69 (63%)	219 (73%)	
Patient diagnosed with lung cancer	1 (2%)	4 (3%)	22 (20%)	27 (9%)	
Patient diagnosed with other cancer	0 (0%)	5 (4%)	3 (3%)	8 (3%)	
Patient died without cancer diagnosis	2 (4%)	6 (4%)	5 (5%)	13 (4%)	
Patient lost from VA system	1 (2%)	1 (1%)	0 (0%)	2 (1%)	

* Does not include 31 patients who had incomplete nodule evaluation such that benign or malignant status could not be determined

Abbreviations: COPD, chronic obstructive pulmonary disease; CT, computed tomography; SD, standard deviation; VA: Veterans Administration

Table 2

Pulmonary nodule evaluation versus final lung cancer outcome, stratified by baseline nodule size

	Nodule Size			Overall [n=300]
	4 mm [n=57]	5–8 mm [n=134]	>8 mm [n=109]	
<i>Patients without any purposeful nodule evaluation</i>	7 (12%)	10 (7%)	6 (6%)	23 (8%)
<i>Patients who underwent nodule evaluation</i>	50 (88%)	124 (93%)	103 (95%)	277 (92%)
<i>Non-invasive work-up [n=277]</i>				
Radiographic surveillance				
Median (range) duration of surveillance, months	28 (0.5–90)	12 (0.1–102)	6 (0.1–90)	13 (0.1–102)
Chest x-ray				
Patients who underwent test, n (%)	11 (22%)	54 (44%)	60 (58%)	125 (45%)
Median (range) tests per patient	0 (0–2)	0 (0–6)	1 (0–23)	0 (0–23)
Total number of tests performed in group	14	80	198	292
Chest CT scan				
Patients who underwent test, n (%)	44 (88%)	90 (73%)	81 (79%)	215 (78%)
Median (range) tests per patient	3 (0–9)	2 (0–16)	2 (0–11)	2 (0–16)
Total number of tests performed in group	144	299	267	710
PET scan				
Patients who underwent test, n (%)	0 (0%)	7 (6%)	26 (25%)	33 (12%)
Median (range) tests per patient	0 (0–0)	0 (0–2)	0 (0–4)	0 (0–4)
Total number of tests performed in group	0	9	33	42
Patients referred for specialist consultation, n (%)	12 (24%)	37 (30%)	50 (49%)	99 (36%)
Median (range) consultations per patient	0 (0–1)	0 (0–2)	0 (0–7)	0 (0–7)
Total number of consults in group	12	43	92	147
<i>Invasive work-up [n=46]</i>				
Patients who underwent pre-invasive testing, n (%)	1 (2%)	4 (3%)	12 (12%)	17 (6%)
Total number of tests in group	1	4	17	22
Biopsy procedures*				
Patients who underwent any biopsy, n (%)	2 (4%)	10 (8%)	34 (33%)	46 (17%)
Median (range) number of biopsies per patient	0 (0–2)	0 (0–1)	0 (0–4)	0 (0–4)
Total biopsies performed in group	3	10	63	76
Surgical resection				
Wedge resection, n (%)	0 (0%)	1 (1%)	6 (6%)	7 (3%)
Lobectomy, n (%)	0 (0%)	0 (0%)	6 (6%)	6 (2%)
Patients hospitalized for invasive evaluation, n (%)	0 (0%)	1 (1%)	17 (17%)	18 (7%)
Median (range) hospitalizations per patient	0 (0–0)	0 (0–1)	0 (0–3)	0 (0–3)
Total number of hospitalizations in group	0	1	20	21
<i>Outcome: Patients diagnosed with lung cancer</i>	1 (2%)	4 (3%)	22 (20%)	27 (9%)

* Biopsy procedures included bronchoscopy, transthoracic needle biopsy, mediastinoscopy, and other biopsy (e.g., supraclavicular lymph node resection)

Abbreviations: CT: computed tomography; PET: positron emission tomography

Table 3

Factors associated with intensity of pulmonary nodule evaluation for the 197 patients with a nodule detected after release of Fleischner Society guidelines

	RR (95% CI) Bivariate model	RR (95% CI) Multivariate model
<i>Factors associated with more intensive care than recommended by guidelines</i>		
Age (years)	0.98 (0.95–1.01), p=0.17	NS
Current or former tobacco use	NS	*
Chronic obstructive pulmonary disease (COPD)	NS	*
Symptoms suggestive of lung cancer	NS	*
Subcentimeter nodule	2.0 (0.9–4.5), p=0.08	NS
Ground glass attenuation	NS	*
Spiculated nodule	NS	*
Upper lobe location	NS	*
Nodule detected on CT scan (as opposed to x-ray)	3.3 (1.7–6.5), p=0.001	1.8 (1.2–2.7), p=0.004
Nodule detected during inpatient or preoperative visit	0.26 (0.06–1.03), p=0.06	NS
Nodule evaluation received at more than one facility	2.1 (1.0–4.5), p=0.06	NS
New nodule detected during evaluation	2.0 (1.1–3.9), p=0.03	NS
Radiologist recommendation	6.7 (3.6–12.4), p<0.001	4.5 (2.3–8.7), p<0.001
<i>Factors associated with less intensive care than recommended by guidelines</i>		
Age (per year)	1.02 (1.00–1.04), p=0.05	1.02 (1.00–1.03), p<0.001
Current or former tobacco use	NS	*
Chronic obstructive pulmonary disease	NS	*
Symptoms suggestive of lung cancer	NS	*
Subcentimeter nodule	NS	*
Ground glass attenuation	NS	*
Spiculated nodule	NS	*
Upper lobe location	NS	*
Nodule detected during inpatient or preoperative visit	1.5 (1.0–2.3), p=0.07	1.6 (1.1–2.5), p=0.03
Nodule evaluation received at more than one facility	2.1 (1.3–3.2), p=0.001	2.0 (1.5–2.7), p<0.001
New nodule detected during evaluation	1.4 (0.9–2.1), p=0.11	NS
Radiologist recommendation	3.4 (2.7–4.2), p<0.001	4.3 (2.7–6.8), p<0.001
Patient refusal	3.2 (2.3–4.3), p<0.001	2.9 (1.3–6.3), p=0.006

* Not included in multivariate model as association was not significant in bivariate model

Abbreviations: CI: confidence interval; CT: computed tomography; NS: not significant; RR: relative risk