The Watch the Spot Trial: A pragmatic trial of more versus less intensive strategies for active surveillance of patients with small pulmonary nodules

Principal Investigator:
Michael K. Gould, MD, MS
Kaiser Permanente Bernard J. Tyson School of Medicine
100 S. Los Robles Ave.
Pasadena, CA 91101
626-564-3926
FAX: 626-564-3694
michael.k.gould@kp.org

Co-Principal Investigators:
Karen Kelly, MD
University of California, Davis Medical Center
UC Davis Comprehensive Cancer Center
2279 45th Street
Sacramento, CA 95817
916-734-5959
karkelly@ucdavis.edu

Diana Miglioretti, PhD
University of California, Davis
Department of Public Health Sciences, UC Davis School of Medicine
One Shields Ave., Med Sci 1C, Room 145, Davis, CA 95616
530-752-7168
dmiglioretti@ucdavis.edu

Rebecca Smith-Bindman, MD
University of California, San Francisco
350 Parnassus Avenue
San Francisco, CA 94117
415-353-4946
rebecca.smith-bindman@ucsf.edu

Funder: Patient-Centered Outcomes Research Institute

NCT02623712

Version 1.0: 5/3/16
Version 2.0: 9/22/16
Version 3.0: 4/6/17

Version 4.0: 10/24/19
Version 4.1 12/11/21 (uploaded to ClinicalTrials.gov on 1/18/22)
Version 4.2 (created on 11-14-22; approved by PCORI on 12/20/22; uploaded to ClinicalTrials.gov on 5/25/23)
Version 4.3 7/21/25 (approved by PCORI on 8/4/25; uploaded to ClinicalTrials.gov on 8/11/25)

Version history:

Version	Amendments/Revisions	Date
Version 1.0	N/A	5-3-16

Version 2.0	• Revised surveillance protocols
	• Revised eligibility criteria
	  ○ Increased the minimum age of eligibility from 21 years to 35 years
	  ○ Removed >10 pulmonary nodules as exclusion criteria
	• Revised data management section
	• Revised data analysis plan
		9-22-16

Version 3.0	• Clarified that primary outcome of stage T1a lung cancer is defined by the American Joint Commission on Cancer (AJCC) classification system, 7th edition (pages 3, 9, 10, 27, 35).
	• Clarified eligibility criteria:
	  ○ Only patients with newly identified nodules are eligible (pages 12, 20, 54)
	  ○ Exclusion of individuals who are non-members at participating integrated health care systems (page 15)
	  ○ Exclusion of patients who previously requested not to be contacted for research purposes (page 15)
	  ○ Exclusion of patients who are unreachable by mail or email and thus cannot be notified of the study (page 15)
	• Clarified that while all 14 sites will enroll and use the assigned protocols for patients with incidentally-detected nodules, only seven sites will use the assigned protocols for patients with screening-detected nodules: HealthPartners, Kaiser Permanente (Colorado, Northwest, Southern California), Marshfield Clinic, Medical University of South Carolina and UC Davis (page 12).
	• Provided additional details about surveillance protocols for patients without risk factors for cancer and for patients with screening-detected nodules (pages 23-26, Tables C1-C3).
	• Clarified that invitations to complete patient surveys will be sent by regular mail or email (page 29).
	• Updated the timeframe for recording of complications (page 37).
		4-6-17

Version 4.0	• Clarified that 2nd radiologist and ordering provider surveys will take place during the follow-up period • Clarified that follow-up patient surveys will take place 13 to 24 months (round 2) and 25 or more months (round 3) after enrollment • Revised months of enrollment to 28	10-21-19

	• Clarified that primary outcome is tumor stage >T1a (AJCC 7th edition) or tumor stage >T1b (AJCC 8th edition) • Described pre-planned secondary analyses for Aim 1 • Refined and clarified power calculations to show that study has at least 90% power to test for non-inferiority with sample size of 35,200 enrolled participants • Revised analysis plan for Aim 2 to focus on comparisons of anxiety and distress between treatment arms within each time period • Updated table describing distribution of subjects by sex/gender, race and ethnicity	
Version 4.1	• Modified protocol to include searching of multiple data sources including local cancer registries and electronic health records to ascertain lung cancer incidence, tumor size and stage, supplementing available data from State cancer registries. • Added sensitivity analysis to examine the effect of using multiple data sources.	12-11-21
Version 4.2	• Modified protocol to clarify the distinction in Aim 1 between primary and secondary outcomes and the main analysis and sensitivity analyses for each outcome. • Clarified to ensure consistency throughout protocol that the secondary outcome of timeliness of care includes both time to diagnosis and time to treatment.	11-14-22
Version 4.3	• Corrected minor error in power calculation for primary outcome	7-21-25

Contents

Abstract

Background: Each year, over 1 million Americans learn that they have a potentially cancerous lung nodule that was found incidentally on a chest x-ray or chest computed tomography (CT) scan that was ordered for some other indication. Now that guidelines from multiple organizations recommend lung cancer screening with low radiation-dose CT for high-risk smokers and former smokers, many more individuals with lung nodules will be identified. In contrast with more advanced presentations of lung cancer that are almost uniformly fatal, cancerous nodules can be cured with surgery when promptly diagnosed and treated. However, fewer than 5% of small nodules prove to be cancerous, and a key unanswered question is how to distinguish the malignant ones from those that are benign. Small nodules (especially those measuring up to 10 mm) are difficult to biopsy or resect, so the default option for evaluation is surveillance with serial CT scans to identify growth that is suggestive of cancer. However, the optimal frequency of surveillance has never been determined. Guidelines are based on low-quality evidence and provide conflicting recommendations. Thus, the proposed research will compare more intensive vs. less intensive protocols for surveillance among individuals who are found to have a small pulmonary nodule on chest CT.

Objectives: Our long-term goal is to improve decision-making and outcomes for patients with pulmonary nodules. This pragmatic clinical trial has the following aims:

1. Among individuals with small pulmonary nodules identified either incidentally or by screening, compare more vs. less intensive CT surveillance for the number of cancerous nodules that progress beyond stage T1a, the most curable stage of lung cancer.
2. Compare patient-reported outcomes of emotional distress, anxiety, general health status and satisfaction with the evaluation process.

Sub-aim: Compare provider-reported outcomes of knowledge, attitudes and beliefs about guidelines and practices for lung nodule evaluation, and provider satisfaction with the evaluation process.

3. Compare resource utilization and exposure to ionizing radiation.
4. Compare adherence to the recommended protocol for CT surveillance, and adherence to use of low radiation-dose techniques.

Methods: This study is a large, simple and pragmatic trial that compares two protocols for CT surveillance of potentially cancerous small lung nodules. The two protocols are both supported by guidelines from professional societies and consistent with current standards of care. Random assignment to intervention groups will occur at the level of the hospital or health system, and clinical champions at each site will work with clinical and operational leaders to implement the assigned protocol for surveillance during the study period. Eligible patients will be enrolled passively at one of 24 hospitals within 14 health care systems, using automated methods for identification, notification, and registration into the study. We plan to enroll approximately 35,200 unique patients over 28 months and follow them for two years to ascertain outcomes. Data will be collected from electronic health records (EHRs), local and statewide cancer registries, and surveys of patient-reported and provider-reported outcomes. All data will be transferred securely to the Data Coordinating Center for analyses using hierarchical models to account for clustering effects and to adjust for residual imbalances between groups. Sensitivity analyses and analyses to identify heterogeneity of treatment effect will be performed.

Patient Outcomes: Based on our extensive clinical experience, engagement with stakeholders, and qualitative research with patients to date, we have identified the following set of clinical and patient-centered outcomes of interest. Most of the outcomes are amenable to collection by reviewing EHRs.

1. Tumor (T) stage >T1a at diagnosis, according to the American Joint Committee on Cancer (AJCC) classification system, 7th edition, equivalently, stage >T1b according to the AJCC 8th edition;
2. TNM stage >T1aN0M0, according to the American Joint Committee on Cancer (AJCC) classification system, 7th edition, equivalently, stage >T1bN0M0 according to the AJCC 8th edition;
3. Timeliness of lung cancer diagnosis;
4. Lung cancer survival, obtained from membership files and state tumor registries;
5. Emotional distress during surveillance period, measured with the Impact of Event Scale-Revised (IES-R);
6. Anxiety, measured with the 6-item State-Trait Anxiety Inventory (STAI-6);
7. General health status;
8. Overall satisfaction with evaluation;
9. Health care resource utilization during the surveillance period;
10. Procedure-related complications, based on diagnosis codes documented in the EHR;
11. Adherence to recommended surveillance, for both patients and providers;
12. Effective radiation doses received.

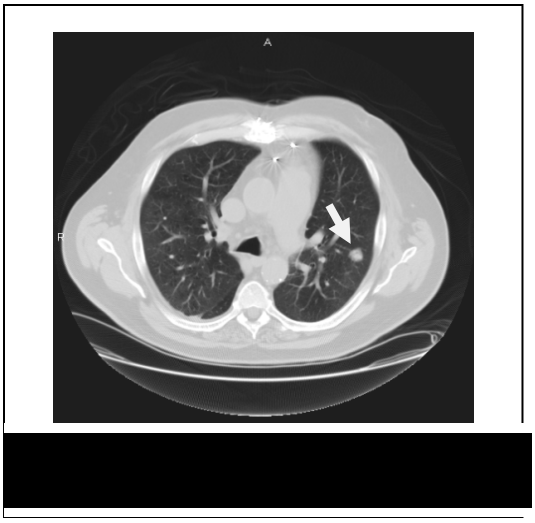
Patient and Stakeholder Engagement: We have assembled a team of researchers, patients, clinicians and stakeholders from health systems, advocacy groups, purchasers, and professional societies to help design and execute the study, and interpret and disseminate the results. Stakeholders have contributed to the framing of the research questions and the identification and prioritization of study outcomes. The researchers on the study team have had longstanding personal and/or professional relationships with many of the patients and

stakeholders, ensuring productive collaboration within the larger group. All members of the team will have an equal voice and all contributions will be valued.

Anticipated Impact: There is concern among clinicians, purchasers, and health systems that as lung cancer screening diffuses into practice, there will be an epidemic of surveillance imaging and downstream invasive testing—testing that is inconvenient, costly and potentially harmful. By comparing accepted protocols for CT surveillance in the context of routine clinical practice, our pragmatic trial will have a large and immediate impact on clinical care. By collaborating with stakeholders from health systems, professional societies, and advocacy groups, we will disseminate our findings widely, and facilitate implementation in diverse practice settings.

Background

A pulmonary nodule ("spot on the lung") is a spherical abnormality that measures up to 30 mm in diameter (**Figure 1**)[1]. Although 95% of pulmonary (lung) nodules are benign and harmless, a small number prove to be early lung cancer. The goal of this pragmatic clinical trial is to identify the optimum surveillance strategy that will maximize early diagnosis for individuals with cancerous nodules, while minimizing unnecessary surveillance of patients without cancer that can result in emotional stress, exposure to harmful ionizing radiation, and the discovery of additional incidental findings that may lead to unnecessary treatment. The trial will be conducted within diverse health care settings: the integrated delivery systems of the HMO Cancer Research Network, several large academic medical centers, and a Veterans Health Administration hospital, for results that are broadly applicable to the U.S. health care community. Continuous quality improvement will be applied to ensure consistent



standards across sites in enrollment, data collection, and performance of the surveillance CT scans. Sustained and authentic engagement of patients and other key stakeholders will continue through all project phases.

Chest CT is performed frequently in clinical practice. Over 7 million chest CT scans were performed in the U.S. in 2007,[2] and CT use has increased by approximately 10% per year. [3,4] Lung nodules are commonly identified on CT scans. In a study of nearly 450,000 CT scans of adult members of Kaiser Permanente Southern California (KPSC) from 2006– 2013, we found that 28% had a lung nodule[5]. In recent studies of lung cancer screening, nodules were discovered on 15% to 50% of chest CT scans. [6-13] The most influential trial of screening, the National Lung Screening Trial (NLST), found that 3 rounds of annual screening with low-radiation dose CT (LDCT) reduced lung cancer mortality by 20% among high-risk smokers and former smokers.[7] Based on these result, the U.S. Preventive Services Task Force and other groups recommend annual screening with LDCT in this population. [6,14-16] Since approximately 7 million Americans meet the Task Force's eligibility criteria for screening, [17] it is likely that even more lung nodules will be discovered once screening disseminates into practice. Most lung nodules have benign causes such as scarring or infection, but 5% to 10% are caused by cancer. Lung cancer is the most common cause of cancer death for both men and women worldwide. [18] Advanced lung cancer is almost always fatal, but cancerous nodules confined to the lung can be cured with surgery when promptly diagnosed and treated. Tumor size is a critical factor: 5-year overall survival is only 58% for people with tumors 3-5 cm in size, but 77% for patients with stage T1a disease (tumor size <20 mm). [19] Thus, many people could potentially benefit from identifying cancerous nodules as soon as possible. However, a key unanswered question is how best to distinguish cancerous from benign nodules to achieve prompt diagnosis of cancers but avoid surgery and other invasive procedures in people with harmless nodules. The cost of care for patients with nodules is substantial; in a study of veterans, per-patient

health care costs over 2 years exceeded $22,000 among individuals with benign nodules, and were greater than $50,000 among individuals with malignant nodules. [20]

Small nodules are not reliably characterized by imaging tests and are difficult or risky to biopsy. Hence, the preferred option for evaluation is usually surveillance with periodic CT scans. Under surveillance, most benign nodules remain stable in size, but detection of growth is presumptive evidence of cancer and should prompt more aggressive testing. The optimal frequency and duration of surveillance has not been determined.

For patients with small nodules that are detected incidentally, surveillance frequency and duration are guided by recommendations from the Fleischner Society and the American College of Chest Physicians (ACCP). [21-23] However, both sets of guidelines were informed by low-quality evidence. In the ACCP guidelines, 27 of 29 recommendations for lung nodule evaluation were considered “weak” and based on low-quality evidence.[23] Evaluation of nodules detected by screening is guided by more recent recommendations from the National Comprehensive Cancer Network (NCCN) and the Lung Imaging Reporting and Data System (Lung-RADS), developed by the American College of Radiology (ACR). Although the risk of cancer should be higher among populations that undergo screening, the Lung-RADS recommendations for surveillance are less intensive than the existing recommendations from the Fleischner Society and ACCP. Although no new evidence of direct relevance to the study from randomized trials or observational studies has been published, the Fleischner Society recommendations were recently revised and are less intensive than the existing recommendations. Indirect evidence from trials of lung cancer screening support the view that the risk of cancer is low in solid pulmonary nodules measuring ≤5 mm and subcentimeter nodules measuring <20 mm, leading some to believe that less intensive follow-up may be justified in these cases. However, just because the risk of cancer is low, it does not mean that less intensive follow-up will be as

effective as more intensive follow-up in identifying a cancerous nodule when it is still at a treatable stage.

Guideline effectiveness is also complicated by variable adherence. In the largest study to date, only 55% of lung nodule patients received follow-up concordant with ACCP guidelines,[24] in part because the guidelines are based on expert judgment and indirect evidence from uncontrolled studies of cancer prevalence from nodules of various sizes. Recommendations from these guidelines have never been validated in comparative effectiveness studies and optimal intervals for surveillance have yet to be determined, representing a major evidence gap. This study addresses that gap by comparing two surveillance protocols, both supported by guidelines and commonly used in current clinical practice. By rigorously comparing their benefits and harms in usual care settings, future decisions by patients and their doctors about the intensity of surveillance will be guided by high-quality evidence that is directly relevant to practice.

The frequency of surveillance has important implications for patients and outcomes. Excess surveillance is inconvenient, can result in emotional distress and anxiety, and might lead to unnecessary invasive testing. Each CT scan entails the risk of a new incidental finding; most are harmless, but each one requires additional surveillance. Greater than needed surveillance results in added exposure to ionizing radiation, a known carcinogen. At the population level, radiation exposure from CT scanning is estimated to cause as many as 1.5–2% of all incident cancers in the U.S.[25] These risks must be balanced by the benefit of finding cancerous nodules, which is greatest when nodules are small and localized to the chest. The optimal surveillance strategy would maximize the benefits of early diagnosis while minimizing the harms of excessive testing, which are considerable since most nodules are benign. The best surveillance strategy would also be simple, to maximize adherence.

Significance

ACCP guidelines for lung nodule evaluation explicitly call for additional research to "compare the benefits and harms of alternative management strategies among individuals stratified by cancer risk."[23] Medical professional societies have also identified lung nodule evaluation as a research priority area. In addition, when deliberating coverage policy for lung cancer screening, the Medicare Evidence Development and Coverage Advisory Committee (MEDCAC) expressed concerns about the lack of standardization in the evaluation of screen-detected nodules.[26] Thus, both professional societies and payers have noted that for lung cancer screening and surveillance of incidental nodules to be effective in practice, we need to know how to accurately, safely and efficiently distinguish malignant from benign nodules.

How Do the Specific Aims Relate to the Significance of the Proposed Study?

Frequent identification of incidental nodules is already widespread in current clinical practice. Providers, payers, and health systems are concerned that the potential benefits of LDCT screening will be mitigated by excessive surveillance imaging and downstream invasive testing that is inconvenient, potentially harmful and costly to patients. To address this concern, this large, pragmatic trial will compare more frequent vs. less frequent CT surveillance for patients with small pulmonary nodules ≤15 mm in diameter in diverse health care settings. We will measure outcomes selected by our stakeholders as important to patients, including progression of cancer to a less curable stage, timeliness of treatment, emotional distress and anxiety during surveillance, patient and provider satisfaction with the evaluation process, health care resource utilization, and adherence. Thus, completing our study aims will generate novel and generalizable knowledge about the comparative effectiveness of two strategies for nodule evaluation. This knowledge will benefit patients and clinicians who need better evidence about the potential benefits and harms of competing strategies for surveillance.

Potential for the Study to Improve Healthcare and Outcomes:

Lung cancer remains the number one cause of cancer death for both men and women in the United States,[27] killing more Americans each year than breast, prostate and colorectal cancer combined. Because most patients with lung cancer present with symptomatic, advanced disease, only 18% of lung cancer patients are still alive 5 years after diagnosis.[27] In contrast, lung cancer is often curable when it presents as an asymptomatic nodule, with 5-year survival rates of 55% to 80%.[19] However, up to 59% of patients who undergo resection for suspected cancer ultimately prove to have a benign nodule.[28-32]

A related concern is that an unknown number of patients with malignant nodules might miss opportunities for cure because diagnosis and treatment are delayed during the period of CT surveillance. Practice variation is another issue. In a prior study of evaluation practices for veterans with pulmonary nodules based on chart review, we found that almost 30% received less intensive care than recommended and 20% received more care than recommended, indicating substantial room for improvement.[24]

Lung cancer and lung nodules are understudied compared with other conditions that lead to comparable morbidity and mortality. This pragmatic, cluster-randomized trial of comparative effectiveness will be integrated into routine clinical practice in diverse settings so that the results will be broadly generalizable.

Patient-centeredness:

This study is patient-centered because it fills an unmet need for information that is critically important for informed shared decision-making by patients and clinicians. To further ensure patient centeredness, we engaged our patient and stakeholder team when developing this proposal to identify and prioritize patient-centered outcomes of greatest importance. Patients and stakeholders selected “cure of cancer” as the most important priority, and endorsed our

using the more feasible surrogate outcome of cancer progression beyond AJCC 7th edition stage T1a, the most curable stage. They also emphasized the importance of understanding patient preferences for more or less frequent surveillance. More specifically, while some stakeholders felt strongly that the most important imperative was to identify cancer early at all costs, others were concerned that more intensive surveillance would aggravate distress and anxiety associated with having a potentially cancerous nodule and "not knowing" its status. Listening to these views prompted us to expand the study design to include the survey of patient-reported outcomes during the surveillance period (Aim 2). Stakeholder partners representing clinicians, health systems, and payers had additional concerns about resource utilization, radiation exposure and adherence, providing the incentive for studying these outcomes in Aims 3 and 4.

Study Aims

Aim 1. Among individuals with small pulmonary nodules identified either incidentally or by screening, compare more versus less intensive surveillance for the number of cancerous nodules that progress beyond AJCC 7th edition stage T1a or AJCC 8th edition stage T1b. Approach: search EHRs and establish linkages with local and statewide cancer registries to ascertain the primary outcome and secondary outcomes of timeliness of lung cancer diagnosis and overall survival of patients in each study arm who prove to have a cancerous nodule.

Aim 2. Compare patient-reported outcomes of emotional distress, anxiety, general health status and satisfaction with the evaluation process. Compare provider satisfaction with the surveillance protocol and the evaluation process. Approach: survey patients via Internet or mail, using validated measures of emotional distress (IES-R), anxiety (STAI-6), and general health status. Patients will be surveyed at 1-2 and 13-24 months following nodule identification, and upon completion of surveillance to assess outcomes over time. Radiologists and ordering providers

(pulmonologists and thoracic surgeons) will be surveyed at the start of patient enrollment and at during the follow-up period. In addition, primary care physicians may be surveyed at the start of patient enrollment and during the follow-up period.

Aim 3. Compare resource utilization and effective radiation doses received. Approach: search EHRs to measure all resource utilization during the evaluation period that could be related to lung nodule evaluation, including total number of CT scans, positron emission tomography (PET) scans, invasive biopsy procedures, thoracic surgical procedures, outpatient visits, emergency department visits and hospitalizations. Search information contained in radiology transcripts or radiology information systems to estimate radiation exposure.

Aim 4. Compare adherence to the recommended protocols for CT surveillance, and adherence to use of low radiation dose techniques. Approach: review samples of dictated radiology reports and search EHRs to measure adherence and ascertain if the assigned surveillance protocol was recommended by the radiologist, ordered by the provider, and ultimately completed by the patient. We will also record adherence with low-radiation dose imaging techniques.

Patients and Settings

Patients will be identified, enrolled and studied at 14 health care delivery organizations (**Table 1**).

Table 1					
Health Care Organization	**Geographic Location**	**Presence of Endemic Mycosis**	**Will Enroll Patients with Screen-Detected Nodules**	**Type of System**	**Group Assign-ment (More vs. Less Intensive)**
Boston Medical Center	Northeast	No	No	Safety Net	More
Cleveland Clinic	Northeast	Yes	No	Referral	Less
Health Partners, MN	Midwest	Yes	Yes	Integrated	More
Kaiser Permanente Colorado	Mountain West	No	Yes	Integrated	Less
Kaiser Permanente Northwest	Northwest	No	Yes	Integrated	Less
Kaiser Permanente Southern California	Southwest	Yes	Yes	Integrated	Both*
Medical University of South Carolina	Southeast	Yes	Yes	University	More
National Jewish Health	Mountain West	No	No	Referral	More
Portland Veterans Affairs Med Center	Northwest	No	Yes	Integrated	Less
University of California Davis	West	Yes	Yes	University	Less
University of California Los Angeles	Southwest	Yes	No	University	More
University of California San Francisco	West	No	No	University	Less
University of Pennsylvania	Northeast	No	No	University	More
Vanderbilt University	Southeast	Yes	No	University	More

The organizations were chosen for geographic, socioeconomic, and racial/ethnic diversity of patients; and for diversity in the size and type of health care delivery system: fee for service, capitated HMO, integrated, and academic. We plan to enroll approximately 35,200 participants with small nodules at these sites during a 28-month period, providing sufficient statistical power for the study aims. We will include consecutive patients with small nodules, so the assembled cohort will reflect the diversity of patients seen at all participating institutions. The large number

and types of facilities and highly diverse populations will provide results that are generalizable to the broader U.S. population.

Inclusion Criteria: The target population includes adults with small lung nodules, who typically would be managed by CT surveillance in usual clinical practice. Thus, we will enroll all adult patients aged ≥35 years with at least one newly identified nodule measuring ≤15 mm in average cross-sectional diameter on chest CT. We increased the original age threshold from 21 years to 35 years, because lung nodules and (especially) lung cancer are extremely rare in individuals younger than age 35, and the new Fleischner Society guidelines are specifically designed for individuals who are 35 years-old or older. The intention to treat population will include patients with nodules of this size who do or do not undergo surveillance imaging. We anticipate most patients will be >50 years old, as both CT use and nodule detection are more common in older patients. Further, because CT scanning is used more frequently in persons with chronic illness, we expect a high proportion of patients to have comorbidities. To accommodate the implementation of lung cancer screening in many settings, we will include individuals with nodules detected either incidentally or by screening at 6 participating sites: HealthPartners, Kaiser Permanente (Colorado, Northwest, Southern California), Medical University of South Carolina, and UC Davis.

The ≤15 mm threshold for inclusion represents a change from the original study proposal, in which we initially adopted a size threshold of ≤10 mm. The change in the size threshold was suggested by several clinical stakeholders and subsequently vetted and approved by the Steering Committee (site PIs), the Stakeholder Advisory Group, and the Executive Committee. To provide justification for this modification, we address issues related to clinical relevance, generalizability, feasibility and validity.

Relevance to clinical practice was a potential concern, because the frequency of cancer increases with increasing nodule size, and some patients with nodules measuring 11 mm to 15 mm in size are not managed initially by CT surveillance, but rather by undergoing a more intensive evaluation with positron emission tomography (PET), invasive biopsy or surgery. However, most patients with nodules of this size (and their providers) still opt for surveillance as the preferred approach, especially in cases in which the patient does not have strong risk factors for lung cancer or the risk of procedure-related complications is judged to be high. Similarly, many patients and providers are risk averse and prefer to avoid biopsy or surgery unless the suspicion of cancer is particularly high, and surveillance can help to establish this by demonstrating growth over time. In fact, in recent studies performed in VA and community settings, between 46% and 78% of patients with even larger nodules measuring up to 30 mm in size were initially managed by surveillance.[33] [24] Thus, we anticipate that the majority of patients with nodules measuring between 11 mm to 15 mm will opt for surveillance, so the research question is quite relevant to this group. In addition, the change in the size threshold for inclusion will not require a modification of either surveillance protocol, since both protocols already specify that surveillance is one option among several for patients with nodules in the 8 mm to 15 mm size category.

By including a wider spectrum of patients with the target condition, the generalizability of the study findings will be greatly enhanced. Unpublished data from a sample of almost 70,000 members of Kaiser Permanente Southern California (KPSC) with incidentally detected pulmonary nodules indicate that nodules ≤10 mm make up 53% of the total population of patients with pulmonary nodules; by increasing the size threshold to ≤15 mm, we will include an additional 16% of the total nodule population. Even if as many as one third of the study patients with nodules measuring between 11 mm and 15 mm are not managed initially by CT

surveillance, and instead undergo an initial invasive biopsy or surgical procedure, this would still increase the size of the target population considerably and improve generalizability.

By increasing the size of the target population, we will also improve feasibility, both by casting a wider net that helps us to meet enrollment targets more efficiently, and by increasing the expected prevalence of cancer in the enrolled sample. Based on KPSC data, we estimate that 2% of patients with nodules measuring ≤10 mm will have a cancerous nodule, whereas 3% of patients with nodules measuring ≤15 mm will have a cancerous nodule (including almost 7% of patients with nodules measuring between 11 mm and 15 mm). Because the study is powered to demonstrate equivalence in stage distribution for patients with cancerous nodules, the increase in the expected frequency of cancer will enable us to meet our goals for sample size much more efficiently. Specifically, we anticipate that approximately 1,056 enrolled patients will have a cancerous nodule, which is greater than our initial estimate of approximately 900 patients. Of note, the original required sample size to achieve statistical power of 90% will not change, and the study will have slightly greater power than originally anticipated to demonstrate equivalence between the two comparison arms.

Finally, the proposed change will not introduce bias and therefore does not threaten validity, because both the more frequent and less frequent surveillance groups will include patients with nodules measuring ≤15 mm.

Exclusion Criteria: Recommendations for surveillance will not apply to patients with nodules outside of the target population, for whom the strategies that we are studying may not be appropriate: patients who are <35 years old, patients with nodules measuring >15 mm in size, whose risk of lung cancer (15% to 75%) is much higher than our target population, and patients with an active pulmonary or extrapulmonary cancer within the past 5 years, who are typically evaluated more aggressively because the risk of metastatic cancer is high. In addition, pregnant

women will be excluded because exposure to ionizing radiation should be minimized in pregnancy and it would not be ethically appropriate to assign a pregnant woman to a more frequent surveillance protocol. In addition, we will exclude non-members (at participating integrated health care systems), patients who previously requested not to be contacted for research purposes, and patients who are unreachable by email or mail.

We will rely on study radiologists to exercise their judgment about whether to exclude patients with associated pulmonary abnormalities such as pleural fluid (effusions), collapsed lung (atelectasis) or hilar or mediastinal lymph node enlargement, each of which increases the risk of an underlying lung cancer, patients with multiple pulmonary nodules, who sometimes require an evaluation approach different than that for lung cancer, and patients with benign appearing nodules (e.g. nodules that have a benign pattern of calcification or a characteristic benign appearance, such as an intrapulmonary lymph node or an arterio-venous malformation). In these cases, interpreting radiologists at all sites will make a determination of eligibility based on their clinical judgment.

Anticipated Characteristics of the Study Population: In a preliminary study at one of our sites (KPSC), we reviewed chest CT results for approximately 240,000 health plan members who underwent imaging between 2006 and December 2012, and identified 68,998 patients with a pulmonary nodule ≤30 mm in size. Virtually all nodules during this time period were detected incidentally, because these data predated implementation of lung cancer screening with LDCT. The mean age was 63.6±14.3 years; 55% were women; 48% were never smokers, 40% were former smokers and 11% were current smokers. The population was racially and ethnically diverse: 18% were Hispanic, 12% were non-Hispanic black, and 9% were Asian/Pacific Islander. Co-morbidities were frequent, including hypertension (58%), diabetes (22%), obstructive lung disease (14%), asthma (13%), coronary disease and/or heart failure (25%), and anxiety or depression (25%). Most (70%) nodules measured ≤15 mm. After excluding patients

with a prior history of cancer, 3% of patients with small nodules of this size were subsequently diagnosed with lung cancer within the next 2 years. We believe this sample is representative of patients with nodules identified in real-world settings and whom we will enroll and follow in the proposed trial. We anticipate having even greater patient diversity during the study because we will enroll participants from multiple geographic regions and from urban centers and safety net hospitals such as Boston Medical Center.

Research Design and Methodology

This study is an unblinded, prospective, pragmatic, cluster-randomized, comparative effectiveness trial of more intensive versus less intensive CT surveillance of patients found to have small pulmonary nodules. The study design includes elements of a pragmatic trial,[34,35] a large simple trial,[36] and a comparative effectiveness trial.[37] As a pragmatic trial, it will determine how the interventions work in usual care settings rather than the idealized settings and conditions of a conventional explanatory trial. As a large simple trial, it will enroll many participants from multiple sites, integrate most study procedures into existing clinical workflow, and perform few assessments or procedures exclusively for research purposes, relying heavily on EHRs and/or registries for recruitment of participants and data collection. As a comparative effectiveness trial, the goal is to inform health care decision-making by comparing at least two alternatives in order to determine which one works best, for whom, and under what circumstances.

We used the PRECIS criteria to characterize the design of our trial along the pragmatic-explanatory continuum.[38] We scored our trial design on a scale from 1-5, with 1 being most explanatory and 5 being most pragmatic:

- Eligibility criteria: participants will be enrolled from multiple sources; no exclusions based on likelihood of response to the intervention; all patients with condition of interest will be

enrolled (5);

- Flexibility of the experimental and control intervention(s): comparators represent recommendations for surveillance, but practitioners can deviate from recommendations based on clinical judgment; standardized methods encouraged for documentation and notification of results and recommendations, but providers may tailor to fit local conditions (4);

- Practitioner expertise: both interventions delivered by all practitioners who care for individuals with lung nodules at the study sites, including primary care providers, pulmonologists, oncologists, radiologists and surgeons (5);

- Follow-up intensity: no study visits outside of usual clinical care; no special procedures or assessments other than collection of patient-reported outcomes via surveys (4);

- Choice of the primary outcome: progression of cancer beyond stage T1a assessed by review of EHRs and linkage with cancer registries (for those with cancer); surrogate outcome (stage) is strongly linked to patient-important outcome (survival); no special adjudication of outcomes or training of outcome assessors (4);

- Participant adherence with the prescribed intervention: adherence measured purely for descriptive purposes at the conclusion of the trial; measurement not used to influence subsequent adherence (4);

- Practitioner adherence to study protocol: radiologists' adherence with protocols for recommended surveillance measured and fed back only early in the trial; strategies for improvement applied when adherence is poor (3);

- Analysis of the primary outcome: analysis by intention to treat; no post hoc exclusions;

subgroup analyses specified in advance and selected based on information needed for clinical decision-making, not merely to find large effects (5).

We chose cluster-randomization because the interventions will be applied at the level of the hospital and radiology department, and having different patients at the same facility receive different recommendations for surveillance would not be practical, efficient or effective. The study will be unblinded because both patients and providers need to be aware of the protocol to which they are assigned. We engaged clinical and non-clinical stakeholders in an intensive and iterative process of feedback and revision to design the initial surveillance protocols to be compared and to ensure equipoise between the two trial arms. In addition, the modified protocols were vetted extensively and approved unanimously by our 15-member Steering Committee (composed of principal investigators from all participating sites), and by our Stakeholder Advisory Group (composed of 5 patients, along with multiple representatives from advocacy groups, medical professional societies, and other health care stakeholders). Subsequently, the modification request was vetted and approved by our 3-member independent Data Monitoring and Safety Board.

Our choice of outcomes was also guided by several months of conversations with clinical, patient and non-patient stakeholders; our knowledge of the literature on clinical outcomes in lung nodule evaluation;[39,40] and recent studies performed by members of the research team on patient preferences related to diagnostic testing in general[41] and lung nodule evaluation in particular.[42-44] Aim 1 addresses the goal of an optimum surveillance strategy to diagnose cancer as early as possible to maximize the chances for cure; Aim 2 addresses the anxiety and distress that is experienced by patients with a nodule of uncertain cause, as well as the preferences and satisfaction of providers; Aim 3 focuses on avoiding complications and resource utilization, an important outcome for key stakeholders, including health systems and purchasers, as well as for patients, who often bear the inconvenience and co-payment costs of unnecessary testing. Last,

Aim 4 addresses concerns that more intensive surveillance may be associated with worse adherence by ordering physicians and patients.

All study procedures will be developed and overseen by the PIs. An independent Data and Safety Monitoring Board will monitor the study for unanticipated adverse events and safety, and perform interim analyses after 25% and 50% enrollment. The study will be registered at ClinicalTrials.gov. This proposal was developed in close collaboration with all stakeholders, and additional feedback will be sought (and the protocol will be modified as needed) during kick-off activities, including conference calls, webinars and an in-person meeting. All analyses will be planned in advanced and specified in the protocol, including sensitivity analyses and analyses to examine heterogeneity of treatment effect (HTE).

Cluster Randomization: Randomization to one of the guideline-based protocols will be performed at the level of the hospital or health system. All patients with small pulmonary nodules within a given hospital or health system will receive the same recommendations for surveillance.

Random assignment to one of the two intervention groups at the hospital or health system level will be performed using matching[45,46] and re-randomization[47] to ensure balance among several potential confounders: the volume of CT imaging performed, the setting type (integrated healthcare system vs. academic), KPSC vs. other institutions; racial and ethnic distribution; percentage of patients who are current, former, vs. never smokers; inclusion of patients with screening-detected nodules; frequency in using PET for patients with nodules 11-15mm; timing of notification letters to participants; and distribution of insurance type. A biostatistician at the Data Coordinating Center will generate the allocation scheme using statistical software developed in R by Robert Greevy at Vanderbilt[45,46]. All participating organizations have agreed to accept random assignment to one of the two groups. While many of the participating

institutions operate more than one hospital, all but one will undergo randomization at the health system level, because their information systems for managing radiology results are centralized, so assigning individual hospitals to different groups would be difficult. The one exception is KPSC, which operates hospitals in 11 distinct medical service areas (MSA); clinical and administrative leaders at this site have agreed to randomization at the level of the MSA. Site PIs, clinical champions, and their multidisciplinary collaborators (including radiologists, primary care physicians, pulmonologists, oncologists and surgeons) will work with key administrative leaders and other clinicians to implement the assigned imaging protocols during the study period. Although randomization is at the hospital level, the unit of analysis will be the individual patient, and we will compare patients randomized to more vs. less frequent surveillance, after adjusting for potential clustering within sites.

Data Collection and Patient Enrollment: Participants will be enrolled passively and added to the study cohort by the interpreting radiologist who identifies a pulmonary nodule. This will facilitate enrollment of all eligible patients. All adult patients undergoing chest CT who are identified as having a nodule ≤15 mm, not previously identified, are potentially eligible. The radiologist will interpret the CT scan based on radiographic and clinical information as per usual clinical practice and determine if the patient meets study criteria. If so, the radiologist will include pre-specified language provided by our trial into the standard dictated radiologist's report. This language will trigger a series of automated steps, to be executed in a site-specific manner according to local practices and workflow: a) inserting standardized text with the recommended protocol for surveillance in the dictated report; b) sending an IRB-approved notification letter to the patient; c) sending a notification to the provider who ordered the CT scan; and d) sending patient data to a local hospital-based study repository. Participants will not be compensated for use of their data or for completing survey questions in this study, although we will hold a raffle to encourage participation.

Each hospital will be provided with appropriately formatted templates for radiologists (often called “canned reports”) to insert into the dictated radiology transcripts. Templates will summarize findings (size and imaging characteristics of the nodule), and recommendations for surveillance based on the assigned protocol. Templates will not need to be used exactly as written and can be tailored by radiologists for each patient.

Currently, across settings, the text of radiology reports is stored within the radiology information system (RIS) and then forwarded to the EHR. Data transferred from the RIS to the EHR pass through an interface software engine that will be tailored to recognize the inclusion of the templates (i.e. particular text strings) or hash tags to identify eligible patients. This information will be used as an indicator to activate downstream systems that will notify both patient and provider. The IRB-approved letter to the patient will include provide a high level overview about the trial, let them know they have been tentatively enrolled in the trial, and offer an opportunity to opt-out of the trial (for data collection purposes). The letter will also provide contact information so patients can call the local project coordinator or site PI with questions. The letter (or electronic notification) to the provider (and/or copy of the dictated report) will be sent to the referring physician, explaining the finding, providing the recommendations for follow-up, and reminding them about the trial. Participating sites will also provide educational materials about the study to their treating physicians using various communication media (email, Webex, and/or face-to-face meetings) as appropriate for their site.

Subsequently, identifying information (e.g. patient name, medical record number, accession number) from the report and RIS will be sent to a secure, local data repository. Identifying information will be used to ascertain outcomes through linkage with internal hospital databases and external tumor registries, and to enable centralized distribution and processing of patient surveys. In addition, patient-level data will be exported to the central study repository and used to assess progress with enrollment and for study quality improvement purposes. Finally, a de-

identified sample of all chest CT radiology reports will be sent to the hospital-based study repository to ensure that all eligible patients (those with nodules) are enrolled. These data will be used for quality improvement activities to ensure enrollment of all eligible patients.

The study design calls for patient and providers to receive notification regarding the enrollment into the trial, although the method of communication will vary across sites. Study coordinators will review charts of a sample of patients to assess if notification is occurring, and if it is not above 95%, they will review the processes (electronic, manual) to improve these figures.

Each collaborating site has the capacity to introduce these procedures and several have already implemented similar systems. Our plan is to adapt existing systems for this pragmatic clinical trial. As an example, at UCSF clinicians are alerted to critically important radiology findings when the interpreting radiologist uses the words “critical alert” in the dictation. This prompts an interface engine to notify the referring clinician about a critical imaging finding through phone calls, emails, and/or electronic messages in the report. Similarly, a large and successful clinical-research collaboration across all University of California campuses, ATHENA, has successfully adopted a similar strategy of using HL7 text to identify patients who should be enrolled in various clinical programs around breast cancer screening and risk assessment. Likewise, at KPSC a significant findings tool was developed within the PACS reporting environment and is currently being used by radiologists throughout the health system to identify patients with lung nodules as part of an operational “safety net” to prevent lung nodule patients from being lost to follow-up. While different hospitals have different software for dictating reports, all allow inclusion of templates, and most systems allow the automatic execution of instructions or have developed unique strategies for doing this. We will work with participating hospitals to implement systems that meet their needs and fit best in their unique clinical environments.

Interventions to be compared (Tables 2 and 3): Based on multiple rounds of consultation with clinicians and stakeholders, including four group webinars and much discussion within and between sites, we selected and refined two surveillance protocols to compare.

The original protocol for more intensive imaging was based on published guidelines from the Fleischner Society (2005, 2013) and the National Comprehensive Cancer Network (NCCN), while most of the recommendations for less intensive imaging were derived from the American College of Radiology's Lung CT Screening Reporting and Data System (Lung-RADS). However, Fleischner Society and NCCN recommendations were recently revised, providing an ideal opportunity to compare the (more intensive) existing recommendations with the (less intensive) newly revised recommendations, a comparison that is most relevant to current clinical practice. The surveillance protocols are summarized in Tables 2 and 3. Additional details are provided in Appendix Tables C1-C3.

Table 2: Group A, More Intensive Surveillance

Nodule characteristics	Nodule size (mm)	Surveillance times (months from index CT)	Optional	For screening-detected nodules (if different)	Source
Solid	≤4	12		12, 24	Fleischner, 2005
	>4 to ≤6	6, 18			
	>6 to ≤8	3, 9, 24		3, 9, 21	
	>8	3, 9, 24	PET/CT, biopsy	3, 9, 21	
Solitary Part-Solid	Any size, solid component ≤5	3, 15, 27, 39			Fleischner, 2013
	Any size, solid component ≥5	3 (If persists, biopsy or resect)			
Solitary Non-solid (GGN)	<5	None		12, 24	Fleischner, 2013
	>5	3, 15, 27, 39			
Multiple Part- or Non-solid	≤5 GGN	24, 48			Fleischner, 2013
	>5 GGN	3, 15, 27, 39			
	"Dominant"	3 (If persists, biopsy or resect)			

Table 3: Group B, Less Intensive Surveillance					
Nodule characteristics	**Nodule size (mm)**	**Surveillance times (months from index CT)**	**Optional**	**For screening-detected nodules (if different)**	**Source**
Solid	<6	None	12	12, 24	Fleischner (2017)
	≥6 to ≤8	12, 24			
	>8	3, 15, 27	PET/CT, biopsy		
Part-solid	<6	None		12, 24	
	≥6 with solid component <6	6, 18, 30			
	≥6 with solid component ≥6	6 (if persists, biopsy or resect)			
Non-solid	<6	None		12, 24	
	≥6	12, 36		12, 24	
Multiple	Solid, largest <6	None	12	12, 24	
	Solid, largest ≥6	6, 18			
	Part- or Non-solid	6, 24	6 (If persists, biopsy or resect)		

The comparators selected will result in approximately 30% less imaging in the less intensive Group B. By using these guidelines as the foundation for the surveillance protocols to be compared, we will preserve the trial's equipoise, integrity and relevance now that the revised Fleischner Society recommendations have been released[48].

All CT scans performed for pulmonary nodule surveillance should be done using low radiation-dose technique, however, this vitally important aspect of imaging quality and safety is often overlooked during CT surveillance of lung nodules. Guidance on low radiation-dose techniques will be provided to each site and the study protocol will strongly recommend the use of these or equivalent low radiation-dose protocols for all follow up CT scans. During the first 3 months of the trial, site project coordinators will review the CT examinations and record radiation dose metrics (CT Volumetric Dose Index [CTDI vol] and Dose Length Product [DLP]) on a large sample of chest CTs performed at their site (10-50% depending on population; 100% for California hospitals as these data are included in the medical record by law) to determine if the

CTs were performed using appropriate techniques. Staff will be trained in radiation dose extraction techniques. Radiation dose metrics will be calculated monthly by site, provided to the Data Coordinating Center, and reviewed during regular meetings. Initiatives to ensure compliance with low-dose examinations will include site visits and radiology-specific lectures.

Table 4: Definition and Source Information for Outcomes, by Specific Aim				
Aim	**At Risk**	**Outcome**	**Definition**	**Source**
1	Participants with cancerous nodules	AJCC 7th edition Stage >T1a or AJCC 8th edition stage >T1b	Tumor size >20 mm at time of resection or radiotherapy.	Cancer Registry, EHR
		Time to diagnosis and treatment	Time to diagnosis and time to treatment with surgery, radiotherapy or chemotherapy, measured from date of index CT scan to date of first treatment.	
		Survival	Measured from date of index CT scan to death or censoring.	
2	All patients with nodules and access to email	Nodule-related distress	Measured with validated IES-R. Assessments performed 1-2 months after index CT scan, at 13-24 months, and at end of follow-up.	Self-administered web survey
		Anxiety	Measured with validated STAI-6. Assessments performed 1-2 months after index CT scan, at 13-24 months, and at end of follow-up.	
		General health status	Assessments performed 1-2 months after index CT scan, at 13-24 months, and at end of follow-up.	
		Satisfaction with evaluation	Measured with novel items, Likert-type scale. Assessment performed at 13-24 months and at the end of follow-up	
	All participating radiologists, ordering providers (pulmonologists, thoracic surgeons, and PCPs)	Satisfaction with surveillance protocol and notification systems	Measured with novel items, Likert-type scale during the follow-up period	Self-administered survey
3	All patients with nodules	Nodule-related resource utilization and total radiation exposure	Includes all CT scans; PET scans; other imaging tests; invasive biopsy procedures (bronchoscopic and percutaneous); thoracic surgical procedures; all outpatient visits, ED visits and hospitalizations during the surveillance period	EHR
4	All patients with nodules, random 10% sample for greater detail	Adherence with assigned surveillance protocol	EHR reviewed to determine whether surveillance imaging was completed per protocol; detailed review of radiology transcripts and orders to determine whether assigned protocol was recommended by radiologist and ordered by provider	EHR, radiology transcripts

Outcomes: The study will use readily available information from EHRs and local and statewide tumor registries to efficiently ascertain outcomes of importance to patients. We will supplement these clinical outcomes with surveys of patient-reported outcomes. The full panel of outcomes was selected based on our collective clinical expertise, and refined extensively based on iterative rounds of feedback from both clinical and non-clinical stakeholders (**Table 4**).

Aim 1 Outcomes, Lung Cancer Stage at Diagnosis: The primary study outcome is tumor progression beyond AJCC 7th edition stage T1a or, equivalently, AJCC 8th edition stage T1b (tumor size ≤20 mm). This size threshold was identified as the best cut point for discriminating prognosis by the most recent staging project of the International Association for the Study of Lung Cancer.[19] Unpublished data from KPSC also supports the prognostic significance of this cut point. We will use two approaches to determine the incidence and characteristics of lung cancer in the study cohort including size, stage and histology. First, we will create linkages to the NCI Surveillance, Epidemiology, and End Results (SEER) and state tumor registries. In addition, we will use data on enrolled patients from each hospital's local tumor registry. The former approach identifies patients who subsequently receive a lung cancer diagnosis at another institution, but will be complete only for patients enrolled during the first 12 months of enrollment. The later approach allows more rapid ascertainment of cancers among patients identified during the last 6 months of enrollment, and will be essentially complete for patients who receive care and retain their membership in one of the integrated delivery systems.

In response to anticipated delays in case abstraction and processing by both local and state tumor registries, we will also obtain information about lung cancers diagnosed during months not covered by state or local registry data based on ICD-10 codes, verified via chart review, and obtain information about tumor size and stage based on detailed chart review. We anticipate that the majority of lung cancer outcomes will be captured by either state-level or local-level cancer registries, both of which rely on standardized case abstraction performed by trained

registrars at local sites, so concordance between these sources is expected to be very high. Cases diagnosed during months not covered by local and state tumor registries will be abstracted from electronic health records according to North American Association of Central Cancer Registries (NAACCR) standards, using NACCR methods for coding tumor size and stage. To ensure consistency, a 10% sample of abstracted cases will be reviewed for accuracy and completeness by a 3-person adjudication committee composed of the principal investigator and 2 other lung cancer experts. In addition, we will collect at least 6 months of overlapping data from both state cancer registries and electronic health records to assess concordance.

Concordance of data between state and local tumor registries is expected to be very high. EHR data on lung cancer outcomes will be collected by an experienced clinician at each site using NAACCR standards, which will minimize discordance. Co-PIs Miglioretti and Smith-Bindman have extensive experience with the NCI-funded Risk of Pediatric and Adolescent Cancer Associated with Medical Imaging (RIC) study and the Breast Cancer Surveillance Consortium using data from state and local cancer registries and electronic health records and have found high concordance between sources. A recent analysis by Drs. Miglioretti and Smith-Bindman showed that ICD-10 codes had a sensitivity of 99% (95% CI 98% to 100%) for identifying cases of pediatric leukemia relative to the reference standard of local tumor registry and/or medical record review (Weinmann S, et al. Positive predictive value and sensitivity of ICD-9-CM codes for identifying pediatric leukemia. Pediatric Blood & Cancer 2021, e29283).

While the level of effort to request, receive and link data from the three sources is somewhat greater than initially anticipated, we have always planned to link cancer outcomes data from the state registries with baseline and follow-up data from electronic health records. We believe that the 9 months allotted between May 1, 2022 and January 31, 2023 will be sufficient for the DCC to perform the necessary linkages, data cleaning and analysis.

Timeliness of lung cancer care is an important dimension of quality.[49] While more timely care is not linked conclusively with better survival, timeliness is clearly important to patients and their caregivers. We will use information contained in EHRs and tumor registries to identify the date of the index CT scan, the date of diagnosis, and the date of first treatment (surgery, radiotherapy or chemotherapy). In addition to stage at diagnosis and times to diagnosis and treatment, we will measure overall survival from the date of the index CT scan, recognizing that we will have limited follow-up time for lung cancer deaths to accrue.

Aim 2, Patient- and Provider-Reported Outcomes (PC-3): In Aim 2, we will collect patient-reported outcomes of relevance to lung nodule evaluation. In particular, we will examine emotional distress and anxiety related to having a potentially cancerous nodule of uncertain cause. To measure distress and anxiety, we will use two validated, self-administered instruments that have been applied in prior studies of lung cancer screening, the IES-R and the STAI-6. The IES-R measures emotional distress resulting from an identified stressor (e.g. identification of a lung nodule) and includes the 3 domains of intrusive thoughts, avoidance and hyper-arousal.[50];[51] It is sensitive to change in patients who have positive lung cancer screening test results.[52,53] The STAI-6 is a brief, 6-item scale that was derived from the longer State-Trait Anxiety Inventory.[54,55] The STAI-6 has demonstrated feasibility and responsiveness to change in an uncontrolled trial of lung cancer screening.[56] Both instruments are available in English and Spanish-language versions (**Table 5**). In addition to these instruments, we will administer a one-item measure of general health status and novel items about overall satisfaction with the surveillance strategy and evaluation process.

Table 5: Patient-Reported Outcome Measures

Scale	Number of Domains	Items per Domain	Domain Names
Impact of Event Scale-Revised (IES-R)	3	Intrusion (8) Avoidance (8) Hyperarousal (6)	-Intrusion -Avoidance -Hyperarousal

Six-Item State Anxiety Scale (STAI-6)	1	6	-State anxiety
General health status	1	1	General health status

Each site will contact participants by regular mail or email and participants will be invited to complete a Web-based survey. Participants taking Internet surveys will provide informed consent and complete the REDCap-enabled surveys in the language of their choice, English or Spanish. Participants will be surveyed at approximately 1-2 and 13-24 months after enrollment, and at the end of surveillance.

To complement the patient perspective, we will also design a brief survey to capture provider satisfaction with the assigned protocol for surveillance and their perception about barriers to its implementation. Development of the brief physician survey will be led by Dr. Kaplan at UCSF, based on her ongoing qualitative pilot study of physician attitudes and beliefs about cancer screening.[57,58] The survey will primarily focus on radiologists and ordering providers (pulmonologists and thoracic surgeons) at each study site, because they are most directly involved with lung nodule evaluation, although we will explore the possibility of surveying primary care providers. Surveys will be distributed to radiologists and ordering providers at the start of the trial and near the end of the study enrollment period (approximately mid-way through the trial).

Aim 3 Outcomes, Nodule-Related Resource Utilization and Radiation Exposure: We will compare health care resource utilization between groups by conducting searches of structured data in EHRs for relevant ICD-9 procedure codes that appear during the surveillance period (from date of the index CT scan to the date of first cancer treatment or 2 years of follow-up, whichever comes first). Most study sites use the Epic EHR, which allows for efficient searching of utilization tables; other EHR systems have similar capabilities. We will capture all relevant imaging tests (chest CT, PET,), invasive biopsy procedures (bronchoscopy, transthoracic

needle biopsy), thoracic surgical procedures, outpatient visits, emergency department visits and hospitalizations, as we have done previously for the NCI and VA-funded CanCORS study.[59] We will also record procedure-related complications by searching for diagnostic codes for pneumothorax, respiratory failure and major bleeding (relying in part on operational definitions created by the Agency for Healthcare Research & Quality Patient Safety Indicators).[60]

We will measure effective radiation doses received by patients under the two different protocols. For the large subset of participants who received care in California, detailed information about CT technique must be documented in the radiology transcript by law. We estimate that approximately 22,000 individuals with nodules will be identified at one of the California sites during the 28-month enrollment period and we will randomly sample this group to obtain detailed information about diagnostic radiation exposure. We will also collect additional information about imaging protocols from radiologists at the sites outside of California who participate in the radiology Collaborative that we will set up as part of our quality assurance/quality improvement infrastructure.

We will collect data for three CT radiation dose metrics: the volume computed tomography dose index (CTDIvol), the dose-length product (DLP), and the effective dose. CTDIvol represents the average per-slice radiation dose, measured in milli-Gray (mGy), and reflects whether conventional or low dose settings are used for CT examinations. DLP is defined as the product of the CTDIvol and the scan length, with units of mGy-cm, and reflects the total radiation output during the entire CT scan. With longer imaging times, the DLP will increase, or if a particular area is imaged multiple times, such as happens in multiphase studies, the dose length product from each scan is added together. CT scans done for the surveillance of nodules should use a shorter rather than a longer scan length. Both CTDIvol and DLP are reported directly on the CT examination (typically stored as an image alongside the images from the CT scan) and can thus be extracted for any subjects through review of the images. Further, in California, by state law

that went into effect in 2013, CTDIvol and DLP must be reported in the radiology report and will be extracted from the electronic medical records. Because each site has developed their own technique for reporting the radiation dose metrics in the medical record, the technique for extracting these data will be locally determined. Effective dose, measured in milli-Sieverts (mSv), combines information about the radiation imparted to the patient, the scan location reflecting the organs irradiated, and the deleterious effect, primarily of developing cancer in the future. Effective dose increases with the dose length product and is a function of the radio-sensitivity of the tissues irradiated. Effective dose will be calculated from the DLP for each examination using existing established conversion formulas, and summed over the study period for each enrollee.

Aim 4 Outcomes, Patient and Provider Adherence: We will measure adherence to the assigned surveillance protocol at the level of the interpreting radiologist, the ordering provider, and the individual patient. First, we will measure adherence with each recommended surveillance test: was the test recommended by the radiologist, ordered by the provider, and completed by the patient? Adherence at the patient and provider levels can be extracted from structured data in the EHR. To measure adherence at the level of the interpreting radiologist, we will review a sample of radiology transcripts to determine whether the recommendation for surveillance were present and per protocol. These analyses will provide granular information about adherence on a test-by-test basis and pinpoint the level of non-adherence. In addition, we will measure overall adherence, defined as completion of all recommended tests for surveillance at the individual patient level.

Data Management

The Data Coordinating Center (DCC) at UC Davis will be the central location for data management. The DCC will collect data captured locally at each participating site. This collected

data will reside in a secure database housed on University servers located in a secure data center. The physical data center is located in a secure building. The physical servers are stored in a secure server farm with adequate cooling, power, backup power, fire suppression and physical building security. Only authorized personnel with specific card key access may enter the server room. The servers used by the DCC are housed next to the servers used to support the Hospital's Electronic Medical Record data and receive the same level of service and security.

Real time data protection offers the ability to synchronize data to another physically separate server located in another geographical location. This service provides a layer of protection should the primary server fail due to physical hazard or malfunction, or software or hardware failure. The DCC fail-safe servers are located in Davis, CA in the UC Davis campus data center. Off-site tape storage provides for one set of full data backup tapes to be stored off site with Iron Mountain.

Data will be collected using REDCap (Research Electronic Data Capture). REDCap was developed at Vanderbilt University, with collaboration from a consortium of institutional partners. REDCap is a software toolset and workflow methodology for electronic collection and management of research and clinical trial data. UC Davis hosts an instance of REDCap on servers secured in compliance with UC Davis Health System requirements and protocols.

REDCap data collection projects rely on a thorough study-specific data dictionary defined in an iterative self-documenting process by all members of the research team with planning assistance from the Biomedical Informatics Program. This iterative development and testing process results in a well-planned, well documented data collection strategy for individual studies. The REDCap system provides secure, web-based applications that are flexible enough to be used for a variety of types of research, provide an intuitive interface for users to enter data

and have real time validation rules (with automated data type and range checks) at the time of entry. These systems offer easy data manipulation with audit trails for reporting, monitoring and querying patient records, and an automated export mechanism to common statistical packages (SPSS, SAS, Stata, R/S-Plus). REDCap was developed specifically around HIPAA-Security guidelines. REDCap has been disseminated for use locally at other institutions and currently supports 240+ academic/non-profit consortium partners on six continents and over 26,000 research end-users (www.project-redcap.org).

The DCC will track and monitor data collection through REDCap and each participating institution will have access through the DCC to a REDCap project in order to maintain a secure REDCap database of enrolled patients for their institution. Patient identifiers (name, medical record number, social security number, address, contact information) linked to a unique study ID number will be kept in a file separate from other study data. Demographic information (age, sex, race/ethnicity) and selected clinical variables from the RIS (date and indication of imaging, acquisition protocol) will be linked with information from the EHR to collect relevant clinical data such as smoking history, information about health care utilization (physician visits of different types, imaging), and diagnosis of cancer from medical records, hospital pathology databases and local tumor registries. The database format will be created centrally by the DCC and provided to each site with a detailed data dictionary and a manual of study procedures to ensure consistent standards and definitions across sites. While the variables will be defined using a single data dictionary, the strategies for populating the database will be developed individually at each site, depending on available information systems and systems for extracting data to populate the database. Many of the sites use the Epic EHR, so a standard strategy will be developed for these sites. Other sites have homegrown informatics systems or research systems that have been developed for other projects that we can use. Institutional databases

can be used to coordinate enrollment and follow up activities, including providing physician and patient follow up, and for quality efforts to ensure standardization of all aspects of the trial.

Data from each enrollment site will be provided at least quarterly to the DCC through the provided REDCap project for cleaning, verification, quality assurance and analysis that will occur under the direction of Dr. Miglioretti. The preferred approach is to (securely) send data with identifiers, to enable linkages with cancer registries from states with a participating site, and to facilitate survey distribution and processing. Alternative plans to share limited data sets or de-identified data will be negotiated with individual sites as necessary. The DCC will oversee the linkages of the compiled study data with the appropriate SEER and state cancer registries to assess lung cancer outcomes, including size, stage and histology of all diagnosed cancers. Linkage procedures vary by registry, but typically each site will send their patient identifier file linked to the study ID number to the DCC, where matching to the cancer registry file will be performed by the DCC's data manager. Once matching has been completed, personal identifiers will be deleted to create a cancer file with the requested lung cancer outcomes data, linked to the study ID. The registries will send this dataset to the coordinating center to link with the study database via study ID for analysis.

We will implement systems to encourage adherence with standards for enrollment, follow up, and performance of CT scans obtained for nodule surveillance. These systems and quality initiatives will be provided equally to all sites. As part of the quality initiatives, the study PIs may visit each site to meet with study personnel and deliver Grand Rounds presentations within radiology, primary care and pulmonary medicine, focused on raising awareness of the trial and study requirements. We anticipate that the intensity of quality assurance efforts will be highest during the first 3-6 months of enrollment, when all sites will be fine tuning procedures for enrollment, performance of CT, communication with patients and providers, and adherence to study recommendations.

Statistical Considerations and Data Analysis

Objectives

The primary study objectives include, among individuals with small pulmonary nodules identified either incidentally or by screening, comparing more vs. less intensive CT surveillance for the progression of cancer beyond AJCC 7th edition stage T1a disease (or T1b in the AJCC 8th edition) (Aim 1), nodule-related emotional distress (Aim 2), health care resource utilization and exposure to diagnostic radiation (Aim 3), and overall adherence with the recommended surveillance protocol (Aim 4).

The secondary objectives include comparing more vs. less intensive CT surveillance for the timeliness of lung cancer diagnosis and overall survival of patients in each study arm who prove to have a cancerous nodule (Aim 1), patient general health status and satisfaction with the evaluation process and the provider satisfaction with the surveillance protocol and the evaluation process (Aim 2), and adherence to use of low radiation-dose techniques (Aim 4). In general, we hypothesize that more intensive follow-up will result in more CT imaging (by design), greater radiation exposure, more downstream testing, greater anxiety and emotional distress, slightly shorter times to cancer diagnosis, and worse adherence to follow-up, and no effect on cancer stage distribution or overall survival.

Primary endpoints:

Aim 1: Proportion of lung cancers that are greater than 20mm

- Measured as progression beyond AJCC 7th edition T stage of T1a or, equivalently, AJCC 8th edition T stage of T1b

Aim 2: Nodule-related anxiety and emotional distress at 1-2 months, 13-24 months, and ≥25 months after nodule identification

- Measured using the IES-R and STAI-6

Aim 3:

1. Health care resource utilization within 26 months following nodule identification
    - Number of visits within 26 months of nodule identification, including the following exams (separate analyses by type of exam):
        - Chest CT scans
        - Chest x-rays
        - PET or PET/CT scans
        - Biopsies (look separately and together): Bronchoscopic biopsy, percutaneous biopsy, surgical biopsy/resection, extrapulmonary biopsy.
    - We will not include any exams after the date of cancer diagnosis, but will include biopsy that led to the cancer diagnosis
2. Serious complications within 7 days and within 30 days of a chest CT, chest x-ray, PET or PET/CT scan, or biopsy (bronchoscopic biopsy, percutaneous biopsy, surgical biopsy/resection, extrapulmonary biopsy) performed within 26 months of baseline CT:
    - Pneumothorax
    - Pneumothorax requiring tube thoracostomy
    - Bleeding (pulmonary hemorrhage)
    - Acute respiratory failure
    - Acute renal failure
    - Allergic reaction to iodinated contrast material requiring hospitalization
    - Acute myocardial infarction
3. Cumulative radiation dose measured by DLP and effective dose

- Including radiation exposure from surveillance and diagnostic imaging following baseline CT scan through 2 years of follow-up (including 24 month exams).

Aim 4: Adherence with the recommended surveillance protocol

1. Radiologist adherence to assigned protocol
2. Physician adherence to ordering test
3. Patient adherence
    a. Adherence to first recommended surveillance exam
    b. Adherence to full protocol

Secondary endpoints:

Aim 1:

1. Proportion of lung cancers that are greater than 20mm, N0, MO
    - Measured as progression beyond AJCC 7th edition TNM stage of T1aN0M0 or AJCC 8th edition stage of T1bN0M)
2. Timeliness of lung cancer diagnosis
    - Measured as time from date of index CT scan to confirmed lung cancer diagnosis, defined as histologic confirmation via pathology, among individuals diagnosed with lung cancer.
3. Timeliness of lung cancer treatment
    - Measured as time from date of index CT scan to data of first treatment including surgery, radiotherapy or chemotherapy, among individuals diagnosed with lung cancer. If a patient receives no treatment, we will censor at end follow-up.
4. Overall survival of patients who prove to have a cancerous nodule,
    - Measured from date of index CT scan to death or censoring among individuals diagnosed with lung cancer.

Aim 2:

1. Patient general health status measured at 1-2 months, 13-24 months, and ≥25 months after nodule identification
2. Patient satisfaction with the evaluation process measured at 13-24 months and ≥25 months after nodule identification
3. Ordering specialist, primary care, and radiologist knowledge attitudes and beliefs about existing guidelines for pulmonary nodule evaluation measured at baseline
4. Provider satisfaction with the surveillance protocol and the evaluation process measured during the follow-up period

Aim 4: Adherence to use of low radiation-dose techniques

Hypotheses:

We specify the following hypotheses by aim:

H1a: Among participants with nodules that prove to be cancerous, less intensive surveillance will be non-inferior to more intensive surveillance for: (a) the frequency of cancer progression beyond AJCC 7th edition stage T1a disease (or T1b for AJCC 8th edition) and (b) the frequency of cancer progression beyond AJCC 7th edition TNM stage T1aN0M0 (or T1bN0M0 for AJCC 8th edition). In addition, among participants with *lung nodules*, less intensive surveillance will be non-inferior to more intensive surveillance for: (a) the frequency of cancer progression beyond AJCC 7th edition stage T1a disease (or T1b for AJCC 8th edition); and (b) the frequency of cancer progression beyond AJCC 7th edition TNM stage T1aN0M0 (or T1bN0M0 for AJCC 8th edition).. Furthermore, in exploratory analyses for HTE, we hypothesize that individuals who are most likely to benefit from more intensive surveillance will include older people, people with larger nodules (>8 mm), patients with solid or part-solid nodules (compared with non-solid

nodules), and individuals with nodules detected by screening (compared with individuals who have nodules that are incidentally detected).

H1b: Among participants with nodules that prove to be cancerous, more intensive surveillance will result in more timely diagnosis than less intensive surveillance.

H1c: Among participants with nodules that prove to be cancerous, less intensive surveillance will be non-inferior to more intensive surveillance for overall survival. Among participants with lung nodules, less intensive surveillance will be non-inferior to more intensive surveillance for overall survival.

H2a: Among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with greater levels of emotional distress and state anxiety at 13-24 months, but levels of distress and anxiety similar to less intensive surveillance at 1-2 months and ≥25 months.

H2b: Among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with scores similar to less intensive surveillance for general health status at 1-2 months, 13-24 months and ≥25 months.

H2c: Among participating radiologists and pulmonologists, satisfaction with the surveillance protocol will be greater among those who practice at institutions assigned to the less intensive protocol.

H3a: Among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with greater nodule-related resource utilization, including more frequent use of CT scanning, PET scanning and invasive biopsy.

H3b: Among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with a greater number of serious complications related to invasive procedures, including pneumothorax, pneumothorax requiring tube thoracostomy, bleeding, acute respiratory failure, acute renal failure and acute myocardial infarction.

H3c: Among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with greater total exposure to diagnostic radiation than less intensive surveillance.

H4: Among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with worse adherence to the surveillance protocol than less intensive surveillance, when measured at the level of the interpreting radiologist, the ordering provider and the patient.

Analysis Plan

In most cases, the unit of analysis will be the individual patient. All primary analyses will be by intention to treat, although we will also perform sensitivity analyses by surveillance imaging actually received. The intention to treat population will include all patients with qualifying nodules, even ones who do not undergo surveillance as the initial step in their evaluation (i.e., patients that proceed directly to tissue diagnosis). Sensitivity analyses will evaluate outcomes by strategy recommended and by surveillance imaging received.

Descriptive statistics will be generated for continuous variables (e.g., mean, standard deviation, median, range) and categorical variables (frequency) at baseline and follow-ups when appropriate. Bivariate analyses will include comparisons of means or medians (as appropriate) for continuously measured variables, cross tabulations of categorical variables, and Kaplan-Meier plots for time to event outcomes. Count outcomes, such as the number of procedure-related complications, will be treated as Poisson random variables. Analysis of time to

diagnosis, treatment, and survival will use Cox regression, after checking the proportionality assumption by creating log-minus-log plots.

H1a - Primary endpoint –lung cancer stage >T1a at diagnosis:

To examine whether among participants with cancerous nodules, less intensive surveillance will be non-inferior to more intensive surveillance for the frequency of cancer progression beyond AJCC 7th edition stage T1a (or T1b for 8th edition) disease, we will use hierarchical logistic regression that includes random site effects to account for clustering of patients within sites and to capture institutional characteristics that are not explicitly modeled. Hierarchical logistic regression models will be fitted with and without adjusting for potential confounders including age, gender, ethnicity/race, smoking, body mass index (BMI), baseline nodule size, and any facility-level factors balanced during the randomization.

We will conduct a main analysis and several sensitivity analyses.

1. The main intent to treat analysis will include all enrolled participants in both study arms
2. A modified intent to treat sensitivity analysis will include:
    a. For integrated healthcare systems, all active members
    b. For other systems, patients who have at least one follow-up visit
3. On treatment sensitivity analyses will include those who met the definition of adherence to first recommended test and to overall surveillance protocol
4. We will conduct another sensitivity analysis using all nodules, instead of cancerous nodules, as the denominator to evaluate rates instead of proportions, given some concern about the potential for overdiagnosis in the more intensive arm. However, we are concerned about evaluating rates among all nodules due to the potential for differential enrollment of very small, likely insignificant nodules across sites. Inclusion of these nodules in overall rates will inflate the denominator without an increase in

numerator, resulting in smaller rates. Thus, we will evaluate variation in lung cancer rates across site within arms. Significant heterogeneity among the sites will provide evidence that the rates may be unreliable due to differential enrollment.

5. We will perform additional sensitivity analyses to test the robustness of findings to the source of data. The primary analysis will include data from all three sources (local and state tumor registries and electronic health records). The first sensitivity analysis will be limited to lung cancer outcomes from both state and local tumor registries, truncating follow-up time as necessary to ensure non-differential follow-up across sites (likely at a minimum of 17-18 months). Another sensitivity analysis will be limited to lung cancer outcomes from only state registries, truncating follow-up as per above.
6. A separate analysis will examine patients who adhere to the recommendations for evaluation compared to those who do not, within each of the two study arms and for both arms combined, for the primary outcome of tumor (T) stage >AJCC 7th edition stage T1A (or AJCC 8th edition stage T1b).

H1a - Secondary endpoint:

TNM stage: We will conduct a secondary analysis to compare the presence of AJCC-7 TNM stage >T1aN0M0 or AJCC-8 TNM stage >T1bN0M0 between study arms. This is a secondary analysis because it is not possible to ascertain TNM stage at the time of nodule identification, so we will not be able to compare rates of progression for this outcome. Nevertheless, we will be able to make an unbiased comparison of the presence of TNM stage >T1aN0M0 at the time of diagnosis. Because it is possible that a small percentage of enrolled patients will have minimal or occult N1-3 and/or M1 disease at the time of enrollment, the baseline risk using this outcome definition will be greater than what we assumed for the primary outcome of tumor (T) stage >T1a, and therefore we will have less power to show non-inferiority in this sensitivity analysis.

H1b - Secondary endpoints:

Timeliness of diagnosis: To investigate whether among participants with cancerous nodules, less intensive surveillance will be non-inferior to more intensive surveillance for timeliness of diagnosis, we will restrict our analysis to participants with cancerous nodules and use Cox regression models where time from initial nodule identification to diagnosis is the outcome variable.

Timeliness of treatment: To investigate whether among participants with cancerous nodules, less intensive surveillance will be non-inferior to more intensive surveillance for timeliness of treatment, we will restrict our analysis to participants with cancerous nodules and use Cox regression models where time from initial nodule identification to treatment is the outcome variable. Patients who die or do not receive treatment will be censored at end of follow-up. We will consider using competing risk models, as necessary.

H1c: Secondary endpoint

Survival: To investigate whether among participants with cancerous nodules, less intensive surveillance will be non-inferior to more intensive surveillance for overall survival, we will restrict our analysis to participants with cancerous nodules and use Cox regression models where time from initial nodule identification to death as the outcome variable. Patients will be censored at end of follow-up.

For all time-to-event analyses, clustering by site will be handled by specifying subjects nested within each cluster and using robust sandwich estimators for statistical inferences.[61] The proportionality assumption will be checked by creating log-minus-log plots. To explore whether individuals who are most likely to benefit from more intensive surveillance, subgroup analyses will examine the effect of age (using the median and/or quartile values as cut points), nodule size (>8 mm), nodule attenuation (solid or part-solid nodules compared with non-solid nodules), and mode of detection (detected by screening compared with incidentally detected).

Analytically, we will include interaction terms between each of these factors with treatment in the models, respectively. For significant interactions, we will also conduct subgroup analysis to explore the treatment effects in corresponding subgroups. Since the study is not powered to test interaction terms, these analyses will be exploratory. For H1b, to study whether among participants with cancerous nodules, more intensive surveillance will result in more timely diagnosis and treatment than less intensive surveillance, we will adopt similar strategy used for comparing overall survival described above.

H2a:

Patient surveys: To examine whether among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with greater levels of emotional distress and state anxiety at 13-24 months, but levels of distress and anxiety similar to less intensive surveillance at 1-2 months and ≥25 months, comparisons of the levels of emotional distress and state anxiety (measured according to the scoring systems detailed in Table 4) will be assessed at 1-2, 13-24, and ≥25 months from the first CT scan, respectively, again using a hierarchical model to capture site-to-site variability. At each time point, we will compare the levels between surveillance groups via a mixed effects model, where site will be considered as random effects to account for within-cluster correlations. Formal tests of both between-site and between-subject heterogeneity will be used to rule out sources of clustering that are not warranted. Structural Equation Models will be used to compare levels of distress and anxiety at 13 months, using second-order factors when valid subscales exist, to assess whether accounting for measurement error in the survey items changes our conclusions. Finally, we will use Latent Class Regression (LCR) to identify subgroups of the study population whose change in levels of emotional distress and state anxiety may differ, since LCR allows for modeling of time trends along with adjustment demographic and clinical characteristics within an arbitrary number of latent classes.[62] For example, there could be the following three classes: 1) patients whose

levels are elevated at 1-2 months after being notified of a nodule, but then drop off and stay low throughout follow-up; 2) patients whose levels are elevated at 1-2 months and stay high throughout follow-up; 3) and patients whose levels are normal at 1-2 months following surveillance, and then increase during follow-up. The differences in changes of levels of distress and anxiety between the latent classes over time would then be assessed by including a piecewise linear function of time in the LCR model, as described above.

For H2b, to investigate whether among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with scores similar to less intensive surveillance for general health status at 1-2 months, 13-24 months, and ≥25 months, a mixed effects model, including all subjects who complete at least two sets of surveys and the time variable as a categorical variable, and will allow the parameters corresponding to time trends to be random effects, so as to allow for between-subject heterogeneity in those trends. Formal tests of both between-site and between-subject heterogeneity will be used to rule out sources of clustering that are not warranted. The mixed effects models will be adjusted for potential confounders including age, gender, ethnicity/race, smoking, BMI, nodule size, and any facility-level factors balanced during the randomization. Interaction between the treatment assignment and time will also be considered in the models and subgroup analysis will be conducted at separate time points if the interaction term is significant. For H2c**,** to examine whether among participating radiologists and pulmonologists, satisfaction with the surveillance protocol will be greater among those who practice at institutions assigned to the less intensive protocol, similar analysis strategy for H2b will be adopted.

Similar methods will be used to evaluate ordering specialist, primary care, and radiologist knowledge attitudes and beliefs about existing guidelines for pulmonary nodule evaluation and provider satisfaction with the surveillance protocol and the evaluation process measured at approximately mid-way during follow-up.

Aim 3: Health care resource utilization and exposure to diagnostic radiation

Descriptive statistics will be generated for radiation exposure and frequency of follow-up CT scans and other nodule-related resource utilization will be obtained by site and study arm. For each type of nodule-related resource utilization, we will also create a binary variable to indicate whether a patient has any exam of that type (1 yes and 0 no) and obtain the corresponding frequency. For H3a, to examine whether among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with a greater number of specific follow-up tests and types of follow-up tests (e.g., imaging, non-surgical biopsy, surgery), we will use mixed-effects Poisson regression model including random site effects to account for clustering of patients within sites. We will also model the frequency at which at least 1 specific follow-up test or type of follow-up test is performed using hierarchical logistic regression that includes random site effects. All regression models mentioned above will be fitted with and without potential confounders including age, gender, ethnicity/race, smoking, BMI, nodule size, and any facility-level factors balanced during the randomization. For H3b, to examine whether among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with a greater number of serious complications related to invasive procedures, similar strategy used for H3a will be adopted. We will model the number of serious procedure-related complications using mixed-effects Poisson regression models and the frequency of patients with any serious complications using hierarchical logistic regression models. For H3c, to investigate whether all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with greater total exposure to diagnostic radiation than less intensive surveillance, we will use mixed effects models and log-transform the radiation exposure to make normal assumption more appropriate.

For H4, binary variables will be generated to indicate adherence to the recommended protocols measured at the level of the ordering provider and the patient. Cross tabulations of the

adherence indicators and the surveillance arm will be obtained. To examine whether among all participants with a lung nodule ≤15 mm, more intensive surveillance will be associated with worse adherence to the surveillance protocol than less intensive surveillance, we will use hierarchical logistic regression that includes random site effects, with and without adjusting for potential confounders including age, gender, ethnicity/race, smoking, BMI, nodule size, and any facility-level factors balanced during the randomization.

Sample size and power calculations by aim

Preliminary data from a retrospective study of patients who underwent one or more chest CT scans at KPSC (2006-2012) showed that 76% of chest CT scans are performed in unique patients, and that approximately 30% of these individuals will have a prior history of cancer. Of the remainder, approximately 32% of unique patients (with no prior cancer diagnosis) who undergo chest CT will have a pulmonary nodule ≤30 mm, and 72% of all nodules measure ≤15 mm. Thus, we expect that approximately 16% of all unique patients who undergo chest CT will have a qualifying pulmonary nodule measuring ≤15 mm and no prior history of cancer. Data provided by participating sites indicates that more than 376,000 chest CT scans will be performed during the planned 20-month enrollment period in over 279,500 unique individuals (**Table 1**). This translates to an estimated target population size of approximately 45,000 patients with qualifying nodules. Allowing for site-level variation in the proportion of patients with a prior history of cancer and patients with nodules detected on prior chest CT scans, we will passively enroll a minimum of 35,200 individuals with a nodule ≤15 mm at one of the 14 health systems (24 individual study sites for cluster randomization) during the initial 28-month recruitment period. To allow time to complete the evaluation and accrue outcomes, patients will be followed for at least 2 years during the funding period. Among these individuals with small pulmonary nodules, we assume (based on data from KPSC) that 3.0% will be diagnosed with lung cancer, resulting in a sample size of approximately 1,056 patients with cancerous nodules.

We will continue recruitment through the end of the study to collect additional data for HTE analyses with greater statistical power for Aim 1. We assume 10% loss to follow-up before final diagnosis is established based on KPSC membership data. This yields an estimated sample size for analyses of 960 individuals with cancerous nodules for Aim 1, and 32,000 individuals with small nodules for Aims 2-4.

Aim 1: For the primary outcome, the baseline risk of progression has not been reported. Conservatively, we assume that 5% of patients will progress from a tumor ≤15 mm to a tumor >20 mm during the surveillance period. Assuming a two-sided alpha of 0.05, and an intraclass correlation coefficient (ICC) of 0.01, a sample size of 888 unique patients with cancerous lung nodules (444 per arm) is required to achieve 83% power to detect an absolute increase in tumor progression of 5 percentage points. Thus, our anticipated sample size of 960 unique patients with cancerous lung nodules provides adequate power to show non-inferiority of the less intensive surveillance strategy for our primary outcome of lung cancer stage >T1a at diagnosis. The most uncertainty in our assumptions is for the ICC. We will have at least 80% power to detect a 5% increase in our primary outcome for ICC values of 0.012 or less. At the baseline ICC value of 0.01, a total sample size of 32,560 enrolled (including 888 patients with cancerous nodules) will have 0.83 power; for an ICC value of 0.012, a total sample size of 35,200 enrolled (960 patients with cancer) will also have 0.83 power.

Aim 2: Prior studies of anxiety and distress in lung cancer screening have shown that the identification of a lung nodule leads to anxiety and emotional distress that subsequently dissipates over 1-2 years.[56 63 52] For our analyses of anxiety and emotional distress, we will have >90% power to detect an effect size of 0.22 standard deviation, assuming an ICC of 0.01, even if the response rate is as low as 5%, resulting in an estimated sample size of approximately 792 respondents in each arm.

Aim 3: The primary outcome for this aim is the total number of non-invasive and invasive procedures performed from the date of the index CT scan to the end of surveillance. We assume that patients who undergo more intensive surveillance will have 3.5 procedures on average. A sample size of 32,000 individuals with small nodules provides 90% power to detect a difference of 0.14 standard deviations for this outcome, assuming a two-sided alpha of 0.05 and an ICC of 0.01.

We will also compare the frequency of invasive testing among individuals with benign nodules. Based on data from the NLST, we assume that 5% of patients with benign nodules will undergo at least one invasive test. With an estimated sample size of 15,520 individuals with benign nodules in each group, and assuming a 2-sided alpha level of 0.05 and an ICC of 0.01, we will have 90% power to detect a difference between groups in the use of invasive testing of 3.4 percentage points.

Aim 4: For the analysis of radiologist adherence to inclusion of the recommendation for surveillance per protocol in the radiology report, we assume that radiologists will adhere to the assigned protocol for at least 85% of participants in the more intensive group. We estimated the ICC to be 0.19 based on preliminary data on radiologist variability in number of enrolled patients at the KPSC facilities. Assuming a 2-sided alpha of 0.05, we need to review at least 130 reports for enrolled patients at each participating facility to achieve 80% power to detect an improvement in adherence of 5.0 percentage points.

For the analysis of ordering physician adherence with ordering the surveillance test according to the assigned protocol, we assume that 80% of physicians in the more intensive group will be adherent. Assuming a 2-sided alpha of 0.05, our sample size of 16,000 individuals with small nodules in each group will provide >90% power to detect an improvement in adherence of 5.3 percentage points, assuming an ICC of 0.01-0.19.

For the analysis of patient overall adherence with surveillance, we assume that 27% of participants in the more intensive group will be non-adherent, based on the results of our study of guideline-concordant follow-up in veterans with pulmonary nodules.[24] Assuming a 2-sided alpha of 0.05, our sample size of 16,000 individuals with small nodules in each group will provide 90% power to detect an improvement in non-adherence of 5.9 percentage points, assuming an ICC of 0.01.

Subgroup Analyses/Heterogeneity of Treatment Effect: During the 7.5-year study period, we will have sufficient power to achieve our primary aims based on the planned 28 months of enrollment. We limited the period of enrollment to enable follow-up of all enrolled patients for at least 2 years, to reliably exclude a diagnosis of cancer when absent, and to assess cancer outcomes when present. Analyses will include whether outcomes vary by CT indication (lung cancer screening or unrelated to lung cancer screening), lung cancer risk factors (primarily smoking history), nodule density (solid, part-solid or non-solid), health care setting, demographic characteristics, and geographic region. Each of these subgroup analyses will be guided by the clinically plausible, pre-specified hypothesis that more intensive surveillance is more effective in subgroups that are at higher risk for more aggressive forms of lung cancer, such as older individuals; smokers and former smokers; patients with screening-detected nodules, solid nodules, or nodules of larger size; and patients from regions where a large portion of nodules are caused by endemic fungal infections.

To examine differences in “treatment” effect, we will include relevant interaction terms in all regression models, and report magnitudes of effects and standard errors for each subgroup level when interactions are statistically significant.

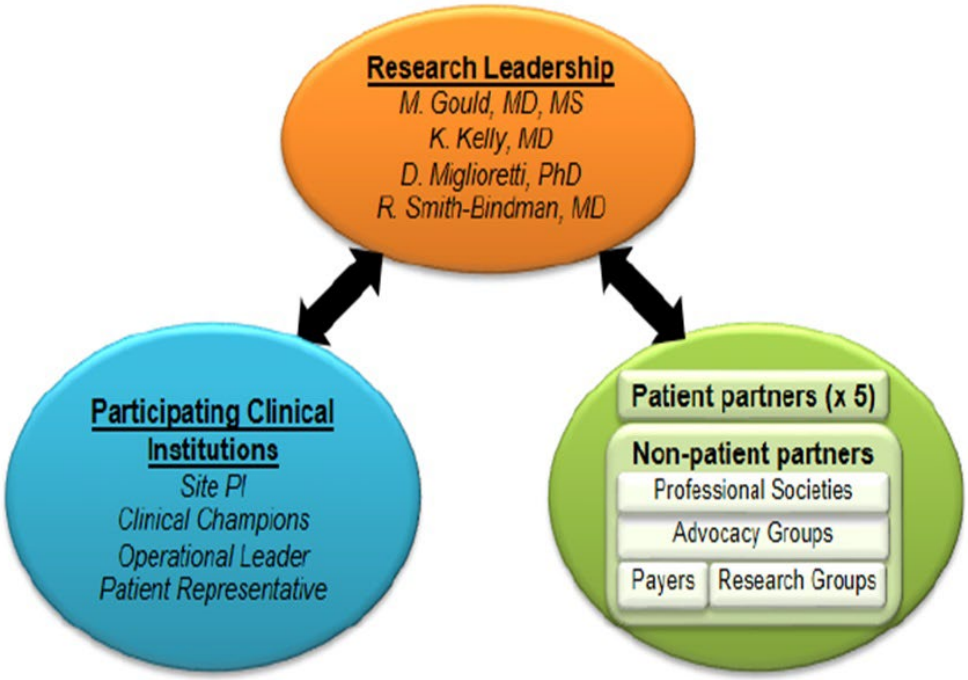
Missing Data: Two types of missing data might occur: (1) core study variables (such as potential confounders) may be missing for some patients because of incomplete capture in the

EHR, and (2) patients might disenroll from a health system and/or leave the capture area. We will minimize missingness of core study variables through intensive quality assurance. The Data Coordinating Center will review monthly reports for potential data issues and follow-up with site coordinators and programmers to abstract missing data, correct data issues, and document in the study database when missing data cannot be found.

Despite these efforts, we expect to have some data missing at random because of information not being documented in the EHR, as opposed to informatively missing, where the probability of being missing depends on the missing value. Thus, we will use multiple imputation methods to account for missing data.[64,65] In addition, we will report and compare the characteristics of patients with and without missing data. We expect a maximum of 10% of patients to be lost to follow-up within 2 years of the index CT scan. To minimize missingness of cancer outcomes on these patients, we will link with cancer registries from states with a participating site. Sites will document the date when patients disenroll or die and the last patient encounter in the health system. We will censor these patients in time-to-event analyses. For analyses restricted to patients with a lung cancer diagnosis, we will perform sensitivity analysis to determine the potential effects of incomplete capture of lung cancers in patients who disenroll. For example, we will assume some percentage of these patients were diagnosed with lung cancer and vary the percentage with stage >T1a tumors.

Patient and Stakeholder Engagement

This collaboration between scientists, patients and multiple stakeholders will be guided by the PCORI engagement principles: reciprocity, co-learning, partnership, trust, transparency and honesty. The PIs have identified key stakeholders who have agreed to serve as study partners and assist us with study design and execution and interpretation and dissemination of results. The preliminary governing structure that we developed is illustrated in the Figure.



This structure will facilitate interactions between each of the following groups: Executive Committee (EC): The EC will include the 4 study PIs, 1 patient stakeholder, 1 non-patient stakeholder, and 1 site PI. The EC will provide an efficient forum for discussion of all study-related issues and a mechanism for rapid decision-making when necessary. The EC will meet during monthly conference calls.

Steering Committee (SC): the SC will include the 4 study PIs, all site PIs, 1 patient stakeholder, and 1 non-patient stakeholder. The CAT will provide a forum for sharing best practices across sites and advise about all matters related to implementing study procedures at the local sites. The SC will be the primary decision-making body for all matters related to the conduct of the study at the individual local sites. The SC will meet during monthly conference calls in year 1 and quarterly thereafter.

Stakeholder Advisory Group (SAG): The SAG will include the 4 study PIs, all patient and non-patient stakeholders. The SAG will meet quarterly by conference call and advise about all matters related to the research. The SAG will have primary responsibility for all decisions regarding research methodology, publications, communications and dissemination of study results.

Local Study Teams (LST): The local team at the individual study sites will include the site PI, one or more clinical champions, a local administrative leader, and a local patient representative. The LST will assist with implementing study procedures and the assigned protocols for surveillance, and will address unique issues that may arise at an individual site.

Ethical Issues Including Informed Consent

All participating sites will comply with the Department of Health and Human Services (HHS) regulations for the protection of human subjects, 45 CFR Part 46, which provide a systematic means, based on established ethical principles, to safeguard the rights and welfare of individuals who participate as subjects in research activities supported or conducted by the HHS. All sites must have on file with OHRP an acceptable Assurance of Compliance.

Risks to Human Subjects

Human Subjects Involvement, Characteristics, and Design.

This study is a large pragmatic trial with cluster randomization by hospital or health system comparing more (Arm A) versus less (Arm B) intensive chest CT surveillance for identifying lung cancers in subjects with small indeterminate pulmonary nodules. All subjects enrolled in this study have one or more small (≤15 mm) pulmonary nodule(s) identified on a chest computerized tomography (CT) scan that is determined to be of unclear significance. The most life threatening etiology for an indeterminate nodule is lung cancer but the vast majority of nodules are benign. Small nodules are difficult to biopsy or resect, so surveillance with serial CT scans is the default option for distinguishing cancerous nodules from those that are benign. Growth of a nodule on serial CT scans is suggestive of cancer and warrants an additional evaluation, while a nodule that does not grow over 2 years is usually presumed benign.

We are comparing two well-established protocols for surveillance of these nodules – both consistent with current clinical practices. Our goal is to identify the optimal strategy for CT surveillance of these nodules, and identify tradeoffs between the more and less intensive approaches. The primary outcome is the percent of lung cancer patients with early stage (T1a) lung cancer. Additional outcomes include 1) emotional distress, anxiety, general health

status and overall satisfaction with the evaluation process; 2) physician satisfaction with the surveillance strategy and surveillance process; 3) cumulative radiation exposure; 4) healthcare utilization related to nodule evaluation; and 5) adherence to the assigned protocol for CT surveillance.

All subjects 35 years or older who have at least one newly identified pulmonary nodule measuring ≤15 mm detected on a chest CT conducted for any reason will be passively enrolled into the study. Exclusion criteria include subjects with a known extrathoracic cancer within the past 5 years. For sufficient power for establishing non-inferiority of the less intensive strategy, we estimate that we will need to identify 279,500 subjects who underwent chest CT to identify 35,200 subjects with small lung nodules of interest. Subjects with nodules will be identified by practicing radiologists, who will trigger an alert in the Radiology Information System (RIS) to embed a template with the recommended surveillance protocol into the dictated radiology report. In addition, sites will use manual and/or automated methods to search radiology reports and identify subjects with small nodules who may not have been enrolled, thereby allowing us to measure adherence at the level of the interpreting radiologist. Additional alerts flowing from the RIS to the Electronic Medical Record (EMR) will generate letters or emails to the ordering physician and patient notifying them of the results, explaining the results and the recommendation for follow-up, describing the study goals and duration, and specifying the types of data to be collected. Contact information for the project coordinator and Principal Investigator (PI) will be provided. An alert will also be created to send the radiology report to the local site data repository that will then be used by protocol coordinators to extract study outcome information through linkage with the internal health systems database and cancer registries. Patients will complete validated measures of anxiety, emotional distress and general health status via the Internet. Patient– and physician–level data will be exported to the central data repository at Data Coordinating Center.

Prospective subjects will be enrolled passively from all participating institutions and their network sites. The institutional PI will be responsible for overseeing the study to its completion and that all local, state and federal guidelines for the conduct of human subject research are followed. Dr. Gould and the research leadership team are ultimately responsible for the successful execution and completion of this study. Real time study monitoring will be performed to ensure accrual goals are met, patient safety is maintained, and data are collected in a timely manner and are of high quality. The research leadership committee will conduct training for the local study coordinator and site PI. The research leadership team will be available to address questions or concerns as they arise. Monthly communications will provide all sites with timely updates and information.

Sources of Materials

As specified in the protocol, the EHR will be accessed to search existing records for variables and outcomes of interest (e.g. demographics, smoking status, cancer history, radiology and pathology reports). Identifiers (name, gender, address, date of birth, and/or Social Security number) will be used for linking with local and state cancer registries to ascertain tumor size and stage in the patients with nodules that prove to be cancer. Patient-reported outcomes will be collected using validated instruments via the Internet or regular mail at 1-2 months, 13 months and 25 months after the date of the index CT scan. Other than completing these questionnaires, no special study visits, procedures or assessments will be performed. Selected physicians at each institution will be asked to participate in a survey to assess their satisfaction with the assigned protocol for surveillance. Data transfer agreements will be in place prior to all data transfers.

Subjects' privacy will be respected at all times, and all procedures will be HIPAA-compliant. All data will remain confidential and only people who are involved in the study will have access to private information. No information identifying individual study subjects will be released except as required by law. Subjects will be tracked in a local database system and only de-identified data will be sent to the central database at the Data Coordinating Center for analysis. Only certified staff will have access to the database, which will be securely password-protected in a REDCap database.

Potential Risks

The study poses minimal risk to the patient because CT surveillance will be performed as part of routine medical care, and both protocols for surveillance are based on guidelines from professional societies. Randomization will occur at the level of the hospital or health system, and all participating institutions have indicated that there is equipoise between the comparators. Thus, all aspects of care, regardless of the group assignment to more versus less active surveillance, fall within the current standards of clinical practice. All nodule patients who receive care at a given site will be followed using the same protocol for surveillance. To the extent that adherence with current guidelines is often poor in routine clinical practice, receipt of care at a participating site may even be beneficial if adherence is improved.

A rare but serious harm that could occur is loss of confidentiality of personal health information (PHI), but every effort will be made to protect PHI. Patient and physician questionnaire data will be de-identified prior to transfer to the Data Coordinating Center for Analysis.

Adequacy and Protection Against Risks

Recruitment and Informed Consent

All adult subjects undergoing a chest CT scan who have a small pulmonary nodule(s) will be enrolled in the study when an interpreting radiologist triggers an alert indicating the abnormal

finding of a lung nodule. This alert will set in motion an automated process that will notify the patient and the ordering provider about the CT findings and the study. The patients will thus be passively enrolled in the trial (i.e. we will ascertain outcomes of interest by reviewing electronic health records over the follow up period. Other than accessing their records, and inviting them to complete a survey describing their experiences, there is no planned change in the care they receive.) They will be notified that they have been included in the study, told what the study entails (i.e. following the course of their care) and they will be provided with information on how to opt out of the study and thus not have any of their information included in our analysis should they choose. Any subjects who choose to opt out will have all of their data removed from the study archives. Subjects will have the opportunity to opt out when they receive the invitation to complete surveys of patient-reported outcomes. We do not anticipate accrual issues. After a discussion with our patient partners we believe the majority of patients will agree to participate because of the “standard of care” nature of this study therefore a recruitment plan will not be devised. However, if underperforming sites are identified, a site - specific recruitment plan will be devised, implemented and monitored.

Each participating institution must obtain IRB approval to conduct this study. Sites will request an institutional waiver of consent. Obtaining individual patient consent is not possible, because there is no opportunity to contact the patient in between the time that the nodule has been identified by the interpreting radiologist and the time that the intervention is delivered (in the form of recommendations for surveillance that will be embedded in the formal radiology report). In addition, receipt of care at a participating site should not pose any additional risks to patients, since both options for surveillance are based on guidelines and consistent with current standards of care. Recent papers have argued that the greater risk to patients is not to perform the study of comparative effectiveness, because the risks associated with participating in the study might be very small.

All subjects (including physicians) will receive written communication describing the purpose of the study, study procedures, and the risks and benefits of participating in the study. If additional information about the study is needed the study coordinator and/or local PI can be contacted. No subject will be included in the study against his or her will. Subjects also have the right to withdraw from the study at any time.

Protections Against Risk

All investigators and research staff at participating sites will be required to maintain up-to-date training in human subject protection and good clinical practice through *Collaborative Institutional Training Initiative (CITI; https://www.citiprogram.org/default.asp)* and HIPAA training. To protect subjects from a breach in confidentiality of personal health information, all data will be stored on secure servers that are password protected. Access to the database will be granted to as few as possible with a need to know. Additional security precautions include encryption, digital certification, audit logs, and firewall protection. Unique study-specific identifiers will be assigned to support accurate linkage of data on the same individual across time frames and data files. Encrypted data will be transferred from participating sites to the UC Davis central database using a web-based secure file transfer site and servers. This method is commonly used.

State cancer registries are experienced with data linkage procedures. Sites will send identifiers to the UC Davis Data Coordinating Center, who in term will link these data with the cancer registries. The state cancer registries will do the linkages, determine who developed cancer, and provide relevant information to the data coordinating center, stripped of personal identifiers except the study identification number. All transfer of identifiers and cancer information will be through appropriate encrypted means following established data transfer protocols of each participating institution and cancer registry. The cancer registry will destroy files with personal health information after they are no longer deemed of use for the project.

The Data Safety Monitoring Board (DSMB) employed for this study will provide another layer of protection. The DSMB will meet twice each year by conference call to review unanticipated adverse events. When data are presented or published, subjects will not be individually identified.

Potential Benefits of the Proposed Research to Human Subjects and Others

Given patients will receive similar care whether or not they participate in the study, there will be no direct benefit of participation. There is no direct benefit for physicians participating in the surveys. The benefit for subjects and physicians is in the knowledge gained to improve evidenced based health care.

This study follows standard practices with minimal additional risks. Although there is no direct benefit to the subjects or physicians contributing data the new knowledge derived is anticipated to have a substantial impact on overall health care.

Importance of the Knowledge to be Gained

Lung cancer is the leading cause of cancer deaths worldwide for both men and women. One reason for the poor outcomes is the lack of a screening and early detection tool. As a consequence, patients are typically diagnosed with late stage disease for which there is no cure. Recently, low dose CT was shown to detect early stage lung cancer with a significant decrease in lung cancer mortality. However small pulmonary nodules detected on CT are problematic because they are infrequently cancerous, and the benefits of intensive testing to identify cancer must be balanced by considering the harms of unnecessary testing in patients with nodules that are benign. Several guidelines have been developed to make recommendations for the evaluation of small pulmonary nodules, but the evidence supporting them is limited. Moreover adherence to guidelines is highly variable. At the completion of this

study we will have empirical data to support an evidence-based guideline for the surveillance of small pulmonary nodules.

The adoption of a standard guideline for the evaluation of small pulmonary nodules that can immediately impact the lives of millions of at risk subjects justifies the minimal risk participants in this trial may experience. Subjects can be reassured that since the trial follows clinically accepted practice, was designed with patient input, and has a data and safety monitoring plan, every effort is being made to minimize risk.

Data and safety monitoring plan

An independent data and safety monitoring committee and plan will be created for this trial. In addition, each participating institution must abide by their institutional data and safety monitoring guidelines. The site PI is responsible for ensuring adherence to all local, state and federal guidelines.

Clinical Trials.gov Requirements

This trial has been registered with ClinicalTrials.gov.

Inclusion of Women and Minorities

1. Describe the planned distribution of subjects

The distribution of subjects by sex/gender, race and ethnicity is provided in the accompanied enrollment table (**Table 6**).

Table 6: Distribution of subjects by sex/gender, race and ethnicity			
Ethnic Category	Females	Males	Total
Hispanic or Latino	2,990	2,894	5,884
Not Hispanic or Latino	14,842	14,374	29,216
Ethnic Category: Total of All Subjects *	17,882	17,318	35,200
Racial Categories			
American Indian/Alaska Native	214	208	422
Asian	948	918	1,866
Native Hawaiian or Other Pacific Islander	36	34	70
Black or African American	2,360	2,286	4,646
White	14,323	13,872	28,195
Racial Categories: Total of All Subjects *	17,882	17,318	35,200
* The "Ethnic Category: Total of All Subjects" must be equal to the "Racial Categories: Total of All Subjects."			

The distribution is based on the general population in the United States using the 2013 Census Data. The sample size calculation is based on the estimated 35,200 subjects projected to have a small pulmonary nodule(s) on a chest CT.

2. Describe the subject selection criteria and rationale

Subjects selected to participate in this study is based on a chest CT finding of a small pulmonary nodule and not on the basis of sex/gender or racial/ethnic group. To the best of our knowledge there is no data to suggest that sex/gender or racial/ethnic group influences the incidence of small pulmonary nodules.

3. Provide a compelling rationale for exclusion of any sex/gender, racial, or ethnic group

There is no exclusion of subjects based on sex/gender, racial or ethnic group.

4. Describe proposed outreach programs for recruiting sex/gender, racial and ethnic group members as subjects.

Each participating institution will be responsible for developing an outreach program if a current program does not exist to ensure enrolled subjects match their catchment region. If gaps in sex/gender and/or racial/ethnic enrollment are identified an action plan will be designed, executed, monitored and revised as needed. Trial demographics will be a standing agenda item on the steering committee conference calls.

NIH-Defined Phase III Clinical Trial Requirements

There is no data to suggest that sex/gender or racial/ethnic group influences the incidence of small pulmonary nodules therefore we do not expect the surveillance intervention to detect a difference among these groups.

Inclusion of Children

The research topic to be studied is not relevant to children.

References

1. Ost D, Fein AM, Feinsilver SH. Clinical practice. The solitary pulmonary nodule. N Engl J Med 2003;348:2535-42.
2. Berrington de Gonzalez A, Mahesh M, Kim KP, et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. Arch Intern Med 2009;169:2071-7.
3. Smith-Bindman R. Is computed tomography safe? N Engl J Med 2010;363:1-4.
4. 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 2012;307:2400-9.
5. Gould MK. Lung Cancer Screening in Individuals with Chronic Obstructive Pulmonary Disease. Finding the Sweet Spot. Am J Respir Crit Care Med 2015;192:1027-8.
6. Bach PB, Mirkin JN, Oliver TK, et al. Benefits and harms of CT screening for lung cancer: a systematic review. Jama 2012;307:2418-29.
7. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395-409.
8. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999;354:99-105.
9. Menezes RJ, Roberts HC, Paul NS, et al. Lung cancer screening using low-dose computed tomography in at-risk individuals: the Toronto experience. Lung Cancer 2010;67:177-83.
10. Pastorino U, Bellomi M, Landoni C, et al. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. Lancet 2003;362:593-7.
11. Swensen SJ, Jett JR, Hartman TEea. CT screening for lung cancer: five-year prospective experience. Radiology 2005;235:259-65.
12. Veronesi G, Bellomi M, Scanagatta P, et al. Difficulties encountered managing nodules detected during a computed tomography lung cancer screening program. J Thorac Cardiovasc Surg 2008;136:611-7.
13. Wilson DO, Weissfeld JL, Fuhrman CR, et al. The Pittsburgh Lung Screening Study (PLuSS): outcomes within 3 years of a first computed tomography scan. Am J Respir Crit Care Med 2008;178:956-61.
14. Screening for lung cancer. 2013. (Accessed July 11, 2014, 2014, at http://www.uspreventiveservicestaskforce.org/uspstf/uspslung.htm.)
15. Wender R, Fontham ET, Barrera E, Jr., et al. American Cancer Society lung cancer screening guidelines. CA Cancer J Clin 2013.
16. Providing guidance on lung cancer screening to patients and physicians. American Lung Association, 2012. at http://www.lung.org/lung-disease/lung-cancer/lung-cancer-screening-guidelines/lung-cancer-screening.pdf.)
17. Sox HC. Better evidence about screening for lung cancer. N Engl J Med 2011;365:455-7.
18. Fact Sheets: Cancer. 2014. (Accessed July 11, 2014, at http://www.who.int/mediacentre/factsheets/fs297/en/.)
19. Rami-Porta R, Ball D, Crowley J, et al. The IASLC Lung Cancer Staging Project: proposals for the revision of the T descriptors in the forthcoming (seventh) edition of the TNM classification for lung cancer. J Thorac Oncol 2007;2:593-602.
20. Barnett PG, Ananth L, Gould MK. Cost and outcomes of patients with solitary pulmonary nodules managed with PET scans. Chest 2010;137:53-9.

21. 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;237:395-400.
22. 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;266:304-17.
23. 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-30S.
24. 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;174:871-80.
25. Brenner DJ, Hall EJ. Computed tomography--an increasing source of radiation exposure. N Engl J Med 2007;357:2277-84.
26. Lung cancer screening with low dose computed tomography. 2014. (Accessed July 11, 2014, at http://www.cms.gov/medicare-coverage-database/details/medcac-meeting-details.aspx?MEDCACId=68#agenda.)
27. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin 2014;64:9-29.
28. Bernard A. Resection of pulmonary nodules using video-assisted thoracic surgery. The Thorax Group. Ann Thorac Surg 1996;61:202-4; discussion 4-5.
29. Mack MJ, Hazelrigg SR, Landreneau RJ, Acuff TE. Thoracoscopy for the diagnosis of the indeterminate solitary pulmonary nodule. Ann Thorac Surg 1993;56:825-30; discussion 30-2.
30. Grogan EL, Jones DR, Kozower BD, Simmons WD, Daniel TM. Identification of small lung nodules: technique of radiotracer-guided thoracoscopic biopsy. Ann Thorac Surg 2008;85:S772-7.
31. Stiles BM, Altes TA, Jones DR, et al. Clinical experience with radiotracer-guided thoracoscopic biopsy of small, indeterminate lung nodules. Ann Thorac Surg 2006;82:1191-6; discussion 6-7.
32. Chen W, Chen L, Yang S, et al. A novel technique for localization of small pulmonary nodules. Chest 2007;131:1526-31.
33. Tanner NT, Aggarwal J, Gould MK, et al. Management of Pulmonary Nodules by Community Pulmonologists: A Multicenter Observational Study. Chest 2015;148:1405-14.
34. Patsopoulos NA. A pragmatic view on pragmatic trials. Dialogues in clinical neuroscience 2011;13:217-24.
35. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. Jama 2003;290:1624-32.
36. Care RoVS-DH. Large simple trials and knowledge generation in a learning health system. Washington, DC: Institute of Medicine; 2013.
37. Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine. Ann Intern Med 2009;151:203-5.
38. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne 2009;180:E47-57.
39. 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 2013;143:e93S-120S.
40. Gould MK, Sanders GD, Barnett PG, et al. Cost-effectiveness of alternative management strategies for patients with solitary pulmonary nodules. Ann Intern Med 2003;138:724-35.

41. Solberg L S-BR, et al. It is Time to Ask Patients What Outcomes are Important to Them. Submitted for publication 2014.
42. 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. Annals of the American Thoracic Society 2013;10:330-5.
43. 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 expectations : an international journal of public participation in health care and health policy 2012.
44. 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;143:672-7.
45. Greevy R, Lu B, Silber JH, Rosenbaum P. Optimal multivariate matching before randomization. Biostatistics 2004;5:263-75.
46. Greevy RA, Jr., Grijalva CG, Roumie CL, et al. Reweighted Mahalanobis distance matching for cluster-randomized trials with missing data. Pharmacoepidemiol Drug Saf 2012;21 Suppl 2:148-54.
47. Morgan KL, Rubin, D.L. Rerandomization to improve covariate balance in experiments. The Annals of Statistics 2012;40:1263–82.
48. 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 2017:161659.
49. America CoQoHCi. Crossing the Quality Chasm. Washington, DC: National Academies Press; 2001.
50. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. Psychosomatic medicine 1979;41:209-18.
51. Weiss DS. The Impact of Event Scale: Revised. In: Wilson JP, Tang CS, eds. Cross-cultural assessment of psychological trauma and PTSD. New York: Springer; 2007.
52. van den Bergh KA, Essink-Bot ML, Borsboom GJ, Scholten ET, van Klaveren RJ, de Koning HJ. Long-term effects of lung cancer computed tomography screening on health-related quality of life: the NELSON trial. Eur Respir J 2011;38:154-61.
53. van den Bergh KA, Essink-Bot ML, Borsboom GJ, et al. Short-term health-related quality of life consequences in a lung cancer CT screening trial (NELSON). Br J Cancer 2010;102:27-34.
54. Tluczek A, Henriques JB, Brown RL. Support for the reliability and validity of a six-item state anxiety scale derived from the State-Trait Anxiety Inventory. Journal of nursing measurement 2009;17:19-28.
55. Marteau TM, Bekker H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). The British journal of clinical psychology / the British Psychological Society 1992;31 (Pt 3):301-6.
56. Byrne MM, Weissfeld J, Roberts MS. Anxiety, fear of cancer, and perceived risk of cancer following lung cancer screening. Med Decis Making 2008;28:917-25.
57. Kaplan CP, Haas JS, Perez-Stable EJ, Des Jarlais G, Gregorich SE. Factors affecting breast cancer risk reduction practices among California physicians. Preventive medicine 2005;41:7-15.
58. Haas JS, Kaplan CP, Gregorich SE, Perez-Stable EJ, Des Jarlais G. Do physicians tailor their recommendations for breast cancer risk reduction based on patient's risk? Journal of general internal medicine 2004;19:302-9.
59. Gould MK, Wagner TH, Schultz EM, et al. Facility-level analysis of PET scanning for staging among US veterans with non-small cell lung cancer. Chest 2014;145:839-47.

60. McDonald KM, Romano PS, J. G, et al. Measures of patient safety based on hospital administrative data-the patient safety indicators. Rockville, MD: Agency for Healthcare Research and Quality; 2002.
61. Weig LJ, Lin DY, Weissfeld L. Regresion analysis of multivariate incomplete failure time data by using the marginal distributions. Journal of American Statistical Association 1989;84:1065-73.
62. Leisch F. FlexMix: a general framework for finite mixture models and latent class regression in R. Journal of Statistical Software 2004;11.
63. van den Bergh KA, Essink-Bot ML, Bunge EM, et al. Impact of computed tomography screening for lung cancer on participants in a randomized controlled trial (NELSON trial). Cancer 2008;113:396-404.
64. He Y, Zaslavsky AM, Landrum MB, Harrington DP, Catalano P. Multiple imputation in a large-scale complex survey: a practical guide. Statistical methods in medical research 2010;19:653-70.
65. Little RJ. Statistical Analysis with Missing Data. New York: Wiley; 1986.

Appendices

Appendix A: Patient Notification Form

PATIENT INFORMATION SHEET

Funder: Patient-Centered Outcomes Research Institute

Investigator: Michael Gould, MD, Senior Research Scientist

Southern California Permanente Medical Group

Department of Research & Evaluation

100 S. Los Robles, 2nd Floor

Pasadena, CA 91101

844-795-3878 (toll free)

We are writing to let you know about a voluntary research study we are doing at Kaiser Permanente Southern California. This study is about the follow-up of patients with a small lung nodule (spot on the lung). This study will help us understand how best to care for patients with small lung nodules.

- We are letting you know about the study because your doctor found a small lung nodule on your recent chest CT scan that may require follow-up.
- Most small lung nodules are harmless. About 2 to 3% are found to be lung cancer. Most cancerous nodules are potentially curable.
- This study plans to find out how often doctors should do follow-up chest CT scans on patients in order to:
    - Find any cancerous nodules as early as possible;
    - Limit patient exposure to chest scan radiation.

Do I have to be in this study?

Being in this study is your choice. You do not have to be in any research project offered by Kaiser Permanente. For information on how to request that your health information not be included in this study, please see the paragraph below titled “What if I don’t want to be in this study.” Your choice will not affect your eligibility for treatment at Kaiser Permanente Health Plan hospitals. Your choice will not affect the medical care you receive.

If after reading this letter, you still have questions or you do not understand what the research study is about, you can call Dr. Michael Gould at 844-795-3878 (toll free).

If your provider has not contacted you previously about the nodule and you would like more information, or if you have questions about follow-up care for patients with small lung nodules, please contact your primary care provider’s office.

The study is using an accurate computer program to identify patients with nodules. However, there is a small chance that the computer made a mistake and you have received this letter in error. If you are

uncertain about whether you actually have a nodule, please contact your primary care provider and review your chest CT scan results together.

What is this study about?

The study will compare two sets of guidelines for the follow-up of patients with small lung nodules.

Professional medical societies with expert knowledge about caring for patients with lung nodules developed both sets of guidelines. Both guidelines are in line with current standards of care. One guideline recommends earlier or more frequent scanning than the other one. More frequent scanning isn't always better for health. The study is being done because doctors do not know which guideline is better for patients.

Your hospital within Kaiser Permanente Southern California is taking part in this study. Your hospital is following one of the guidelines for follow-up.

- All patients (like you) who had a small lung nodule found on a recent chest scan have had recommendations for follow-up, based on your hospital's guideline, put into a report in your electronic medical record.
- Following the recommendations in the report is best for most patients but is not required.
- You can talk with your doctor about whether the recommendations are a good fit for you.
- Depending on what you want or your doctor's judgment, you and your doctor may choose not to follow the recommendations in the report.

If I participate in this study, what happens next?

- We will collect the following information from your electronic medical record:
  - General information about you and your health (for example, age, smoking history and health conditions)
  - Basic information about your lung nodule (for example, nodule size)
  - Information about the medical care that you receive (for example, the dates of your chest scans).
- We will invite you to take a survey about your feelings and experiences related to lung nodule follow-up, and your satisfaction with the care you receive.

How long will the study last?

We will collect information from your electronic health record for approximately two (2) years after your nodule was found.

What are the possible risks if I am in this study?

There are minimal risks to people enrolled in this study. The main risk is a loss of confidentiality because of sharing of health information, including some identifiable information such as your name and social security number, between members of the research team, including our research partners at the University of California Davis Data Coordinating Center, who will perform all data analysis.

How will you keep my information private?

To maintain your confidentiality, we will take the following steps:

- We will only share your health information with other members of the research team.
- Any research staff who has access to your health information will be extensively trained in how to keep this kind of information private.
- Our results will never be presented in any way where you may be identified.

What are the possible benefits if I am in this study?

There may not be any direct benefit to you as a result of taking part in the research study.

Will I be paid for being in this study?

You will not be paid for being in this study.

Do I have to pay to be in this study?

You do not have to pay anything to be in this research study.

What if I don't want to be in this study?

If you agree to have information about you and your lung nodule used in this research study, then you do not have to do anything.

- However, if you **do not** want any of your information used in this study, please call study staff at 844-795-3878 (toll free).
- If you choose not to be in this study, none of your information will be collected or shared with our research team, and we will not contact you further about this research study.

What if I have questions?

If you have any questions about this research study, please call Dr. Michael Gould at 844-795-3878 (toll free).

If you have any questions regarding your rights as a study subject you may contact: Armida Ayala, PhD, Director, Human Research Subjects Protection Office at 626-405-3665 or Armida.Ayala@kp.org

How do I take the survey?

We would like to hear your thoughts about lung nodule follow-up and your satisfaction with the care you receive. To take our survey, please go to this link: www.watchthespot.org/kpsc/

Appendix B: Patient survey excerpts: State-Trait Anxiety Inventory, Impact of Event Scale-Revised, general health status

Section 2 (Impact of Event Scale - Revised)										
Below is a list of difficulties people sometimes have after stressful life events. Please read each item, and then indicate how distressing each difficulty has been for you DURING THE PAST SEVEN DAYS with respect to the lung nodule identified on your recent chest CT scan. How much were you distressed or bothered by these difficulties?										
	1	Any reminder brought back feelings about it.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	2	I had trouble staying asleep.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	3	Other things kept making me think about it.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	4	I felt irritable and angry.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	5	I avoided letting myself get upset when I thought about it or was reminded of it.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	6	I thought about it when I didn't mean to.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	7	I felt as if it hadn't happened or wasn't real.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	8	I stayed away from reminders of it.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	9	Pictures about it popped into my mind.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	10	I was jumpy and easily startled.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	11	I tried not to think about it.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	12	I was aware that I still had a lot of feelings about it, but I didn't deal with them.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	13	My feelings about it were kind of numb.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	14	I found myself acting or feeling like I was back at that time.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	15	I had trouble falling asleep.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	16	I had waves of strong feelings about it.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	17	I tried to remove it from my memory.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	18	I had trouble concentrating.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	19	Reminders of it caused me to have physical reactions, such as sweating, trouble breathing, nausea, or a pounding heart.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	20	I had dreams about it.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	21	I felt watchful and on-guard.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
	22	I tried not to talk about it.	Not at all	A little bit	Moderately	Quite a bit	Extremely	Prefer not to answer		
Section 3 (State Trait Anxiety Inventory - 6)										
A number of statements which people have used to describe themselves are given below. Read each statement and then click the most appropriate box to indicate how you feel right now, at this moment. There are no right or wrong answers. Do not spend much time on any one statement but give the answer which seems to describe your present feelings best.										
	1	I feel calm	Not at all	Somewhat	Moderately	Very much	Prefer not to answer			
	2	I feel tense	Not at all	Somewhat	Moderately	Very much	Prefer not to answer			
	3	I feel upset	Not at all	Somewhat	Moderately	Very much	Prefer not to answer			
	4	I am relaxed	Not at all	Somewhat	Moderately	Very much	Prefer not to answer			
	5	I feel content	Not at all	Somewhat	Moderately	Very much	Prefer not to answer			
	6	I am worried	Not at all	Somewhat	Moderately	Very much	Prefer not to answer			
Section 4 (General health status)										
Please pick the one answer that is right for you.										
	1	In general, would you say your health is:	Excellent	Very good	Good	Fair	Poor	Prefer not to answer		

Appendix C: Surveillance Protocols

Table C1: Surveillance protocol for incidental nodules, Group A, more intensive surveillance	
Patient and Nodule characteristics	**Surveillance times (months from index CT)**
Solid, with Risk Factors	
≤4	12
>4 to ≤6	6, 18
>6 to ≤8	3, 9, 21 to 24
>8	3, 9, 21 to 24 or PET or tissue
Solid, without Risk Factors	
≤4	Optional (12)
>4 to ≤5	12
>5 to ≤8	6, 18
>8	3, 9, 21 to 24 or PET or tissue
Part-solid	
Solitary, solid component ≤5	3, 15, 27, 39
Solitary, solid component >5	3, tissue if persists
Multiple, solid component ≤5	3, 15, 27, 39
Multiple, solid component >5	3, tissue if persists and solid ≥5
Non-solid	
≤5, solitary	None
>5, solitary	3, 15, 27, 39
≤5, multiple	24, 48
>5, multiple	3, 15, 27, 39

Table C2: Surveillance protocol for incidental nodules, Group B, Less intensive surveillance	
Nodule characteristics	**Surveillance times (months from index CT)**
Solid	
<6, solitary or multiple	None (Optional at 12 months in patients with risk factors)
≥6 to ≤8, solitary	12, 24
≥6 to ≤8, multiple	6, 18
>8, solitary or multiple	3, 15, 27 or PET or tissue
Part-solid	
<6, solitary	None
<6, multiple	6, 24, 48
≥6, solid component <6	6, 18, 30, 42, 54, 66
≥6, solid component ≥6	6, tissue if persists
Non-solid	
<6, solitary	None
≥6, solitary	12, 36, 60
<6, multiple	6, 24, 48
≥6, multiple	6, 24, 48

Table C3: Surveillance protocols for screening-detected nodules			
Lung-RADS Category	**Nodule characteristics**	**Surveillance times (months from index CT)**	
		Group A	**Group B**
1	No nodule	12, 24…	12, 24…
	Calcified nodule		
	Fat-containing nodule		
2	Solid nodule <4	12, 24…	12, 24…
	Solid nodule <6	6, 18, 30…	
	New solid nodule <4		
	Part-solid nodule <6		
	Non-solid nodule <20		
	Non-solid nodule ≥20, unchanged or slow-growing		
	Category 3 or 4 nodule unchanged for >3 months		
3	Solid nodule ≥6 to <8	3, 15, 27…	6, 18, 30…
	New solid nodule, ≥4 to <6		
	Part-solid nodule ≥6, with solid component <6		
	Part-solid nodule, new <6		
	Non-solid nodule ≥20		
4A	Solid nodule ≥8 to <15	3, 9, 21… (PET may be used when solid component ≥8)	3, 15, 27… (PET may be used when solid component ≥8)
	Solid nodule, new ≥6 to <8, or growing <8		
	Part-solid nodule ≥6, with solid component ≥6 to <8		
	Part-solid nodule, new or growing, with solid component <4		
	Endobronchial nodule		
4B	Solid nodule ≥15	CT at 3 months, PET or tissue, depending on probability of cancer and comorbidities	
	Solid nodule, new or growing ≥8		
	Part-solid nodule, with solid component ≥8		
	Part-solid nodule, new or growing, with solid component ≥4		
4X	Category 3 or 4 nodules with additional features that suggest cancer	CT at 3 months, PET or tissue, depending on probability of cancer and comorbidities	

*Note: Patients with nodules described in boxes with gray shading are NOT eligible for Watch the Spot