Supplemental Data

There were 472 participants in SNAP that underwent both CT and PET, and were reviewed by the Independent Research Readers Panel (IRRP). Of these, 344 had diagnosis by tissue or successful completion of follow-up. The follow-up period in this study was intended for determining in the absence of biopsy whether the person had a benign tumor. There was therefore no provision for determining that someone had malignant tumor in the absence of tissue biopsy. There were 128 participants for whom there was not an acceptable reference standard obtained. Fifty died or were lost to follow-up before meaningful clinical data could be obtained. There were 78 participants for whom an alternative reference standard could be established. Sensitivity analyses was performed to assess any influence the exclusion of these participants (who did not have a complete reference standard) might have on estimates of sensitivity and specificity.

A committee was assembled that included the study co-chairs, two site investigators, and a member of the IRRP. The committee was chaired by a site investigator who is a pulmonologist with extensive experience in the diagnosis and management of lung cancer. The other four committee members are nuclear medicine specialists with training in radiology. The members were blinded to the findings of the IRRP and site readers. However, they were provided with information concerning the CT characteristics of the nodule during follow-up and the patient's available

clinical history including surgery, chemotherapy treatment, radiation therapy, etc.

Supplemental Tables 2A and 2B present the result of these sensitivity analyses. Category A is the results from the confirmed sample, and is the basis for reporting the results in the main paper. There were 21 participants for whom diagnosis could be made based upon tissue from a distant site (Category B). This was considered as high quality evidence, but was not considered to be acceptable for reporting final results in this study. There were 57 participants for whom the reference standard could only be based on subsequent measurements of nodule size or clinical history. The committee members were asked to classify these as "probably" benign or malignant (Category C), or "possibly" benign or malignant (Category D). The criteria used to determine these classifications are detailed below:

PROBABLY MALIGNANT

-Initial significant growth of more than 2 mm or 10% (whichever value is larger) before treatment or death

PLUS

-Clinical diagnosis of lung cancer in hospital records or ICD-9 codes

PLUS

-Chemotherapy or radiation for lung cancer documented in records or ICD-9 codes

PROBABLY BENIGN

-Increase in size over 24 months of greater than 1 mm but not exceeding 2 mm or 10% (whichever value is larger)

OR

-No increase in size (change of 1 mm or less) or decrease in size after 18 months in subjects from whom images were not obtained at 24 months

POSSIBLY MALIGNANT

-Initial significant growth of more than 2 mm or 10% (whichever value is larger) before treatment or death

PLUS

- Clinical diagnosis of lung cancer in hospital records or ICD-9 codes WITHOUT chemotherapy or radiation for lung cancer documented in records or ICD-9 codes

OR

- Increase in size over any 12 month period of greater than 30% in diameter in the absence of either very rapid growth consistent with infectious or inflammatory process (i.e., doubling time: 30%

increase in diameter < 20 days) or subsequent increase in size without treatment for cancer

POSSIBLY BENIGN

-No increase in size (change of 1 mm or less) or decrease in size after 12 months in subjects from whom images were not obtained at 18 or 24 months

OR

-Very rapid growth consistent with infectious or inflammatory process (i.e., doubling time: 30% increase in diameter < 20 days) in subject who died or was lost to follow-up.

Supplemental Tables 2A and 2B demonstrate that the exclusion of participants lacking an acceptable reference standard was not likely to influence the estimates of sensitivity and specificity. For PET, there was nominal change in the specificity estimate with the inclusion of these additional participants. Sensitivity was reduced by 3 points with the inclusion of Category D participants. However, this change is not statistically significant, nor is it likely to be clinically significant. CT was not affected in any meaningful way with the inclusion of the additional participants.

Details – Methods of Image Acquisition and Processing

CT scans were performed according to the standards of the American College of Radiology in place in 1998 *(12).* The use of iodated intravenous contrast in scanning was not mandatory and was left to local practice. High-speed spiral CT was utilized, with a requirement that slices be no larger than 3 mm through the area in which the target nodule was located. Prior to initiation of the study, the performance and minimal resolution requirements (7 mm FWHM) of the PET cameras at the ten sites were validated by use of a standard phantom (CTI, Knoxville, TN, USA) and questionnaire. On the day of the PET study, participants were requested to fast for a minimum of 4 h. Prior to injection, a fasting blood sugar was taken, and participants with a blood sugar in excess of 200 were rescheduled. An injection of ^{18}F-FDG ranging from 10 to 20 mCi was given. Local investigators had the option to provide a dosage not to exceed 140 microcuries/kg of body weight if that was local practice. The PET study was initiated 45–60 min after injection with localization of the SPN based on chest radiograph. If there was prior knowledge of an adrenal mass, an attempt was made to include it in one of the anatomic positions. At a minimum, two attenuation-corrected images of the lung region were required to be taken to encompass the area of the target nodule and to include as much of the hilum, mediastinum as possible. In addition, a whole body nonattenuation corrected study covering the area of the body from the jaw to iliac crest was obtained. These transmission

scans were acquired in 7–10 min per bed position, while emission scans were 10 min per bed position for the lung region and a minimum of 5 min for regions not including the lesion. Maximum slice thickness for PET studies was 6mm.

Visual and semi-quantitative estimates of ^{18}F-FDG uptake were estimated and reported. Semi-quantitative determination of PET ^{18}F-FDG uptake in the SPN was calculated selecting the slice with greatest lesion activity. The lesion was circumscribed with the manufacturer's region-of-interest application, and average and maximum activity was measured for calculation of standardized uptake values (SUVs). PET images were reconstructed according to manufacturer's guidelines. A PET reader at each site interpreted the images and provided a report using a SNAP case report form. The on-site reading was not used to estimate accuracy of PET, but to assess intra-reader reliability and the influence of PET results on clinical decision-making.

Supplemental Table 1
Description of SNAP Participants

	All Participants (% or sd) (n=472)		Participants with Reference Standard (% or sd) (n=344)		Participants without Reference Standard (% or sd) (n=128)	
Age (Years)	66.2)	(10.9)	65.8	(10.7)	67.1	(11.6)
Gender (Male)	458	(97)	337	(98)	121	(94)
History of Tuberculosis	13	(3)	9	(3)	4	(3)
Smoking History						
Currently Smoking	213	(45)	155	(45)	58	(45)
Pack Years (current smokers only)	57.7	(32.5)	58.7	(33.0)	54.9	(31.3)
Previously smoked	230	(49)	167	(48)	63	(49)
Pack Years (prev smokers only)	59.2	(42.7)	60.3	(44.4)	56.0	(37.9)
Years since quitting smoking	13.9	(12.0)	14.4	(11.9)	12.7	(12.1)
Never smoked	29	(6)	22	(7)	7	(6)
CT Nodule Size (mm)	16.2	(6.9)	16.4	(6.7)	15.8	(7.5)
CT Nodule Location						
Upper left	103	(22)	76	(22)	27	(21)
Upper right	135	(28)	98	(28)	37	(29)
Middle left	32	(7)	23	(7)	9	(7)
Middle right	60	(13)	43	(13)	17	(13)
Lower left	61	(13)	47	(14)	14	(11)
Lower right	81	(17)	57	(16)	24	(19)

Supplemental Table 2A

Impact on Estimate of Diagnostic Accuracy of PET of Inclusion of Participants Without Complete Reference Standard

	Reader Interpretation						Diagnostic Statistic (%)	
Reference Standard	Definitely Benign	Probably Benign	Indeter-minate	Probably Malignant	Definitely Malignant	Total	Sensitivity	Specificity
A. Confirmed Sample								
Benign	54	78	3	16	9	160	91.3	82.5
Malignant	2	14	1	55	112	184		
B. Diagnosis Based Upon Remote Site*								
Benign	1	2	0	0	2	5		
Malignant	0	1	0	5	10	16		
Accuracy for group							93.8	60.0
Cumulative accuracy							91.5	81.8
C. Finding of “Probably” by Expert Panel								
Benign	9	4	0	1	0	14		
Malignant	0	0	1	5	5	11		
Accuracy for group							100.0	92.8
Cumulative accuracy							91.9	82.7
D. Finding of “Possibly” by Expert Panel								
Benign	7	4	0	0	2	13		
Malignant	2	8	1	4	4	19		
Accuracy for group							47.3	84.6
Cumulative accuracy							88.3	82.8

*Tissue obtained for biopsy that comes from a distant source such as a mediastinal node, metastasis, or sputum.

Supplemental Table 2B

Impact on Estimate of Diagnostic Accuracy of CT of Inclusion of Participants Without Complete Reference Standard

	Reader Interpretation						Diagnostic Statistic (%)	
Reference Standard	Definitely Benign	Probably Benign	Indeter-minate	Probably Malignant	Definitely Malignant	Total	Sensitivity	Specificity
A. Confirmed Sample								
Benign	30	35	55	29	11	160	95.7	40.6
Malignant	1	7	30	105	41	184		
B. Diagnosis Based Upon Remote Site*								
Benign	1	0	1	2	1	5		
Malignant	2	0	3	8	3	16		
Accuracy for group							87.5	20.0
Cumulative accuracy							95.0	40.0
C. Finding of “Probably” by Expert Panel								
Benign	1	4	7	2	0	14		
Malignant	0	1	1	6	3	11		
Accuracy for group							90.9	35.7
Cumulative accuracy							94.8	39.7
D. Finding of “Possibly” by Expert Panel								
Benign	3	2	4	4	0	13		
Malignant	2	2	3	8	4	19		
Accuracy for group							78.9	38.5
Cumulative accuracy							93.5	39.6

*Tissue obtained for biopsy that comes from a distant source such as a mediastinal node, metastasis, or sputum.