Classification of CT Pulmonary Opacities as Perifissural Nodules: Reader Variability

Schreuder, A.; Ginneken, B. van; Scholten, E.T.; Jacobs, C.; Prokop, M.; Sverzellati, Nicola; Devaraj, Anand; Schaefer-Prokop, C.M.

2018, Article / Letter to editor (Radiology, 288, 3, (2018), pp. 867-875)

Doi link to publisher: https://doi.org/10.1148/radiol.2018172771

Classification of CT Pulmonary Opacities as Perifissural Nodules: Reader Variability

Anton Schreuder, MD • Bram van Ginneken, PhD • Ernst T. Scholten, MD, PhD • Colin Jacobs, PhD • Mathias Prokop, MD, PhD • Nicola Sverzellati, MD, PhD • Sujal R. Desai, MD • Anand Devaraj, MD • Cornelia M. Schaefer-Prokop, MD, PhD

From the Diagnostic Image Analysis Group, Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Geert Grooteplein Zuid 10, 6525 GA Nijmegen, the Netherlands (A.S., B.v.G., E.T.S., C.J., M.P., C.M.S.); Fraunhofer MEVIS, Bremen, Germany (B.v.G.); Department of Radiology, Azienda Ospedaliero-Universitaria de Parma, Parma, Italy (N.S.); Department of Radiology, Royal Brompton and Harefield National Health Service Foundation Trust, London, England (S.R.D., A.D.); and Department of Radiology, Meander Medisch Centrum, Amersfoort, the Netherlands (C.M.S.). Received November 29, 2017; revision requested January 24, 2018; final revision received March 6; accepted March 29. **Address correspondence to** A.S. (e-mail: *antoniusschreuder@gmail.com*).

Conflicts of interest are listed at the end of this article.

See also the editorial by Goo in this issue.

Radiology 2018; 288:867–875 • https://doi.org/10.1148/radiol.2018172771 • Content codes: CH CT OI

Purpose: To study interreader variability for classifying pulmonary opacities at CT as perifissural nodules (PFNs) and determine how reliably radiologists differentiate PFNs from malignancies.

Materials and Methods: CT studies were obtained retrospectively from the National Lung Screening Trial (2002–2009). Nodules were eligible for the study if they were noncalcified, solid, within the size range of 5 to 10 mm, and scanned with a section thickness of 2 mm or less. Six radiologists classified 359 nodules in a cancer-enriched data set as PFN, non-PFN, or not applicable. Nodules classified as not applicable by at least three radiologists were excluded, leaving 316 nodules for post-hoc statistical analysis.

Results: The study group contained 22.2% cancers (70 of 316). The median proportion of nodules classified as PFNs was 45.6% (144 of 316). All six radiologists uniformly classified 17.7% (56 of 316) of the nodules as PFNs. The Fleiss κ was 0.50. Compared with non-PFNs, nodules classified as PFNs were smaller and more often located in the lower lobes and attached to a fissure ($P < .001$). Thirteen (18.6%) of 70 cancers were misclassified 21 times as PFNs. Individual readers' misclassification rates ranged from 0% (0 of 125) to 4.9% (eight of 163). Of 13 misclassified malignancies, 11 were in the upper lobes and two were attached to a fissure.

Conclusion: There was moderate interreader agreement when classifying nodules as perifissural nodules. Less than 2.5% of perifissural nodule classifications were misclassified lung cancers (21 of 865) in this cancer-enriched study. Allowing nodules classified as perifissural nodules to be omitted from additional follow-up in a screening setting could substantially reduce the number of unnecessary scans; excluding perifissural nodules in the upper lobes would greatly decrease the misclassification rate.

Of all cancers worldwide, lung cancer has the highest mortality (1). In a screening setting, the National Lung Screening Trial (NLST) showed that CT is superior to chest radiography in decreasing lung cancer mortality and overall mortality (2). The NLST was nevertheless associated with a substantial number of false-positive findings because of the identification of patients with benign lung nodules initially labeled as suspicious, leading to unnecessary CT follow-up and patient anxiety.

Various pulmonary nodule follow-up guidelines for clinical or screening settings have recently been published, taking advantage of evidence from screening trials (3–6) with the aim of minimizing unnecessary CT surveillance or further investigations. The Lung Imaging Reporting and Data System (Lung-RADS), Fleischner Society, and National Comprehensive Cancer Network (NCCN) guidelines state that noncalcified solid nodules less than 6 mm in diameter do not require additional follow-up and nodules larger than 8 mm should be further investigated (4,5). For clinical settings, the British Thoracic Society (BTS) additionally recommended annual follow-up (one step above no follow-up) for 5- to 6-mm nodules (3). Across the four guidelines, follow-up recommendations range from 3- to 24-month intervals for nodule diameters between 5 and 10 mm.

Some of these guidelines also describe the entity of benign intrapulmonary lymph nodes (IPLNs), for which further follow-up or investigation is not recommended (3,4,6). Over the past 3 decades, increasing numbers of studies have pathologically confirmed that some pulmonary nodules seen at CT are IPLNs (7–17). An autopsy survey of 2337 South African miners revealed an IPLN prevalence of 3.7%, although this was considered a minimum estimate (16). CT studies have also demonstrated the negligible risk of malignancy in nodules typical for IPLNs. Because IPLN is a pathologic diagnosis, Ahn and colleagues (18) coined the term *perifissural nodule* (PFN) to describe nodules composed of IPLN features at CT; de Hoop et al (19) decided to additionally make the distinction between "typical" and "atypical" PFNs.

If radiologists could reliably identify IPLNs, especially those between 5 and 8 mm in diameter, these nodules can automatically be deemed benign without unnecessarily

Abbreviations

BTS = British Thoracic Society, CI = confidence interval, IPLN = intrapulmonary lymph node, Lung-RADS = Lung Imaging Reporting and Data System, NCCN = National Comprehensive Cancer Network, NLST = National Lung Screening Trial, PFN = perifissural nodule

Summary

The interreader agreement for classifying solid opacities of 5 to 10 mm in diameter as perifissural nodules was moderate at best. Misclassifications of malignancies as perifissural nodules occurred mostly for nodules located in the upper lobe.

Implications for Patient Care

- Interreader agreement for classifying small solid nodules as perifissural nodules was only moderate.
- There is a higher risk of misclassifying malignancies as perifissural nodules when there are atypical features, when they are not attached to a fissure, and when they are located in the upper lobes; however, all of the readers were able to correctly identify the majority of perifissural nodules as benign.
- In a lung cancer CT screening setting, exempting middle- and lower-lobe nodules classified as perifissural nodules from additional follow-up would considerably reduce the number of false-positive follow-up studies, with a smaller risk of missing malignancies.

subjecting participants to further surveillance or tests and thus substantially reducing false-positive rates in screening.

Despite the fact that the PFN entity is mentioned by the BTS and Fleischner Society guidelines and is under consideration by the Lung-RADS and NCCN guidelines, no study has yet shown radiologists' accuracy rate in identifying benign nodules as PFNs in a data set with a large number of cancers, nor the extent to which radiologists agree with each other (3–6). We hypothesized that our observer study would show a considerable observer variation in the classification of pulmonary nodules as PFNs, but that regardless, no cancers would be deemed benign.

Materials and Methods

Scans and Data

Approval was obtained to utilize scans and data from the NLST (August 2002 to December 31, 2009) for our retrospective study by the National Cancer Institute Cancer Data Access System (approved project ID: NLST-296). The NLST had obtained approval by an institutional review board at each of the 33 centers involved and had obtained informed consent from all participants. The NLST data were anonymized when provided; thus, no institutional review board approval was needed.

Details and outcomes of the NLST have been reported elsewhere (1). The estimated average volume CT dose index was 2.9 mGy ± 1.0 (standard deviation), and the average dose-length product was 100 cm (20). The NLST recorded the nodules in all studies but recorded only the pulmonary lobe location and a section number; however, it was not always clear from which location in the reconstruction this number was derived. For our study, we received all NLST scans from 6528 participants; this included all 1089 participants given a diagnosis of lung cancer and a random sample of 5439 participants without lung cancer. For each participant, there were up to three total studies (one baseline and two follow-up). A general radiologist with 32 years of experience and a focus on chest radiology, including 5 years of lung cancer screening experience (E.T.S.), and a chest radiologist with more than 20 years of experience in thoracic imaging (C.M.S.) attempted to find and annotate the malignant nodules using all available information such as incidence date, lobe location, nodule growth over time, and highly suggestive morphologic criteria such as spiculation. If available, malignant nodules in earlier studies were included in the Microsoft Excel 2013 spreadsheet. Nodules were annotated with the help of dedicated in-house software that included computer-aided detection (CIRRUS Lung Screening; Diagnostic Image Analysis Group, Nijmegen, the Netherlands); this is a non–open-source prototype version of the U.S. Food and Drug Administration–and CE approved software Veolity (MeVis Medical Solutions, Bremen, Germany). Nodule volume and diameter (including long-axis and perpendicular short-axis diameters) were measured semiautomatically; location and fissural or pleural attachments were manually recorded by a trained medical researcher (A.S.).

Selection and Inclusion Criteria

A nodule was eligible for our study if it was annotated by an experienced radiologist (E.T.S.) as a noncalcified solid and the diameter was between 5 and 10 mm (including nodules with diameters rounded from 4.6 to 10.4 mm, respectively). Diameters were measured by using the same semiautomatic software mentioned earlier (CIRRUS Lung Screening). Of the NLST-reported cancers, only solitary lesions located in the specified lobe with potentially malignant features were included.

Ultimately, our study group included all eligible cancer nodules and a random selection (though limited to one nodule per study) of about three times as many benign nodules; 10% extra benign nodules were included to account for post-hoc analysis (see the Statistical Analysis section). An overview of the nodule selection process is displayed in Figure 1.

Observer Study

Six radiologists—C.M.S., E.T.S., M.P., N.S., S.R.D., and A.D.—with at least 15 years of experience in reading chest CT studies and an interest in lung cancer screening participated in our observer study. All readers were familiar with the morphology of PFNs as described in the literature. Although there was no extra training before the study was performed, guidance was provided in the form of definitions listing the morphologic features of typical and atypical PFNs, as well as non-PFNs. Readers were asked to judge the nodules as they would in clinical routine practice following these definitions. User-friendly in-house developed software, another Veolity prototype, was used to read the studies (CIRRUS observer, Diagnostic Image Analysis Group). Image-viewing features included zoom, ruler, and switching between the three anatomic planes. In short, all included nodules were displayed in a randomized sequence. The task involved independently classifying the nodule in question as one of four options; all readers were blinded to patient information and the post-imaging outcomes. Reader D had a 1-week interval between readout sessions; all other readers completed the observer study within 1 day.

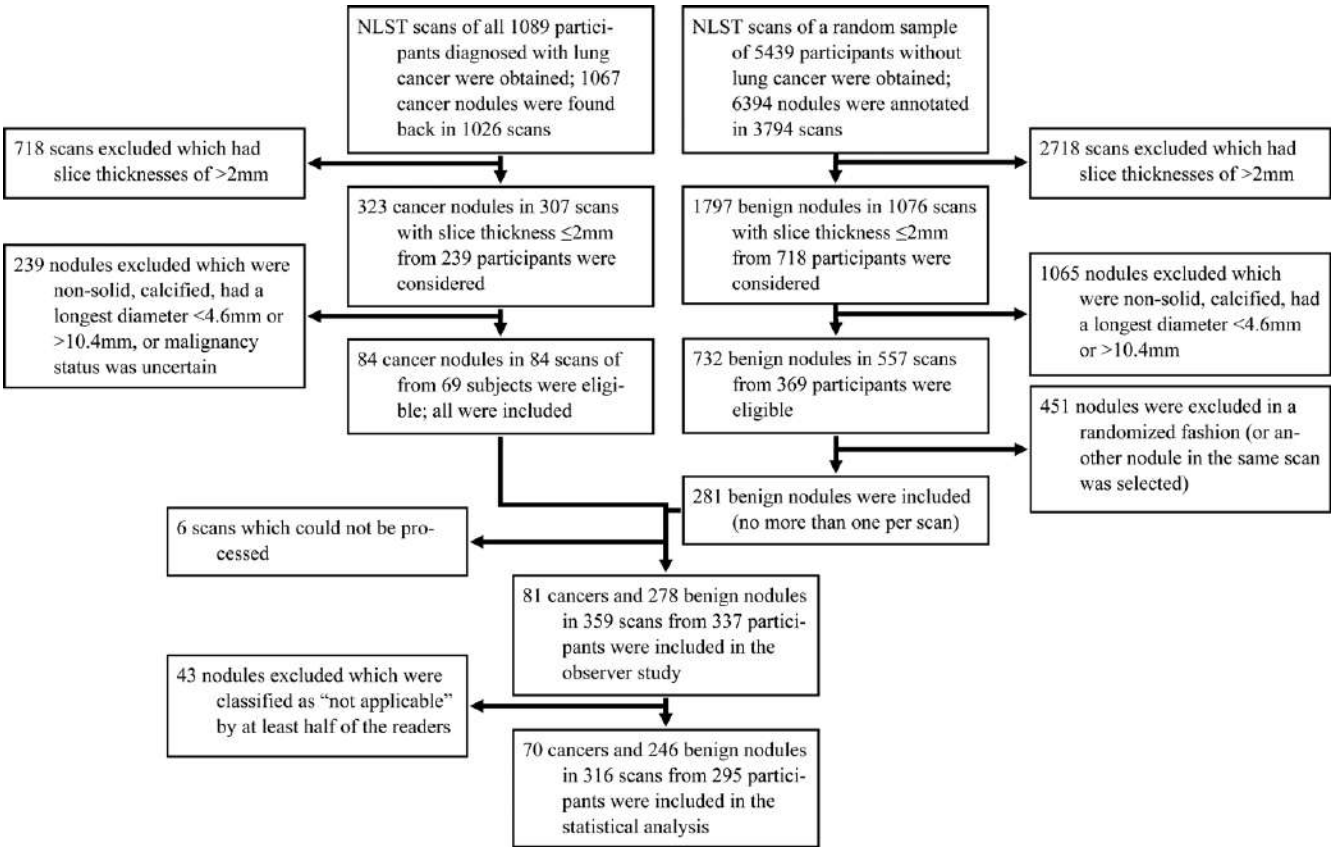

Figure 1: Nodule selection process. Nodules were classified as "not applicable" when they were considered to be subsolid, completely calcified, not nodular, or not visible. *NLST* = National Lung Screening Trial.

A "typical PFN" was defined as a nodule that had a lentiform, triangular, or polygonal shape, was located on or within 10 mm of the visceral pleura or lung fissure (major, minor, or accessory), and had extending linear densities. An "atypical PFN" met two of the three key criteria that define a typical PFN. A "non-PFN" was a noncalcified solid nodule that did not fit in the previous two categories or that showed at least one of the following features: spiculation, irregular shape, unsharp borders, and distortion of the pleura or fissure. Finally, the "not applicable" classification was given to any abnormality that, despite its initial inclusion in our study, was deemed to be subsolid, completely calcified, not nodular, or not visible during our study. Nodules that were classified as not applicable by at least half of the readers were excluded from the post-hoc analysis. The readers were requested to record the reason for classifying a nodule as not applicable.

Statistical Analysis

Microsoft Excel 2013 was used as a database for pulmonary nodules and our observer study outcomes; arithmetic operations were used for calculating percentiles and proportions. To approximate how "flat" a nodule was, the following two measures were calculated: the l/s diameter ratio is the ratio of the nodule diameter long axis and perpendicular short axis (a larger value implies a flatter nodule); percentage spherical is the percentage difference in volume between a perfect sphere volume derived from the nodule diameter ($\frac{4}{3}\pi r^3$) and the actual nodule volume (a smaller value implies a flatter nodule).

Statistical program R, version 3.4.1 *(https://www.r-project.org/)* was used to compare groups by using the Mann-Whitney *U* test for non–normally distributed continuous variables and the Pearson χ^2 test for two categoric outcomes. Interrater agreement was calculated by using the Fleiss κ for comparing all readers and the Cohen κ for comparing reader pairs. After nodules classified by at least three readers as not applicable were excluded, the remaining not applicable values were converted to non-PFN values for the post-hoc analyses.

Fleiss and Cohen κ values were calculated for the original classification system ("typical PFN," "atypical PFN," and "non-PFN") and for a simplified classification system in which "typical PFN" and "atypical PFN" were merged as "PFN." Interreader agreement was also calculated for cancer and benign nodule subgroups.

Results

Descriptive Statistics

Our final study group included 359 nodules from 359 studies in 337 participants; 81 nodules were labeled as cancers (Fig 1). At least three readers classified 12.0% (43 of 359) of the nodules as not applicable, leading to their exclusion; the main reasons given were "scar," "non-nodular consolidation," "atelectasis," and "part-solid nodule." About one-fourth (25.6% [11 of 43]) of these exclusions were cancers, of which 90.9% (10 of 11) were located in the upper lobes. The characteristics of the remaining 316 nod-

Table 1: Nodule Characteristics

Characteristic	Benign Nodules (n = 246)	Cancer Nodules (n = 70)	P Value*	Total (n = 316)
Diameter (mm)	5.7 (5.0–6.7)	7.7 (6.3–8.8)	<.001	6.0 (5.2–7.5)
Volume (mm^3)	98 (66–158)	236 (128–355)	<.001	114 (72–219)
l/s Diameter ratio	1.46 (1.24–1.71)	1.40 (1.24–1.65)	.60	1.45 (1.24–1.70)
Percentage sphericity	51.5 (37.5–68.3)	49.2 (37.9–67.2)	.72	51.1 (37.7–67.9)
Upper lobe location†	101 (41.1)	49 (70.0)	<.001	150 (47.5)
Attached to a fissure†‡	47 (19.1)	5 (7.1)	.02	52 (16.5)
Attached to the pleura†	84 (34.1)	15 (21.4)	.04	99 (31.3)
CT section thickness ≤ 1.3 mm†	40 (16.3)	13 (18.6)	.65	316 (16.8)

Note.—Unless otherwise specified, data are medians, with 25th–75th percentile ranges in parentheses. On average, benign nodules were smaller, more often located in the upper lobes, and more often attached to a fissure or pleura than cancer nodules. There was no statistically significant difference in the proportion of thinner CT section thicknesses between these groups, excluding a potential source of bias. l/s Diameter ratio was calculated by dividing the nodule's long axis by its perpendicular short axis. Percentage sphericity was the percentage difference in volume between the perfect sphere volume derived from the nodule long axis and the actual nodule volume. There was no difference in the l/s diameter ratio and percentage sphericity between benign nodules and cancer nodules.

* For comparison between benign and cancer nodules, the Mann-Whitney U test was used for non-normal continuous outcomes and the Pearson χ^2 was used for for binary outcomes.

† Data are numbers of nodules, with percentages in parentheses.

‡ Includes major, minor, and accessory fissures (excludes the pleura).

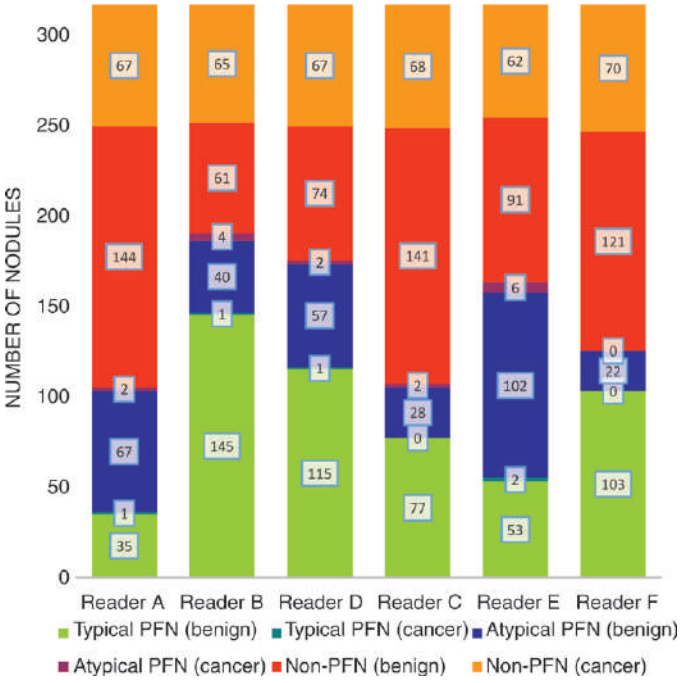

Figure 2: Graph shows classification outcomes of all six readers. *PFN* = perifissural nodule.

Table 2: Interreader Agreement for 316 Nodules

No. of Readers in Agreement	Typical PFN	Atypical PFN	PFN	Non-PFN
≥2	134 (42.4)	101 (32.0)	192 (60.8)	218 (69.0)
≥3	112 (35.4)	31 (9.8)	159 (50.3)	185 (58.5)
≥4	70 (22.2)	7 (2.2)	131 (41.5)	157 (49.7)
≥5	33 (10.4)	1 (0.3)	98 (31.0)	124 (39.2)
6	7 (2.2)	0	56 (17.7)	87 (27.5)

Note.—Data are the number of nodules that the specified number of readers classified in each category, with percentages in parentheses. A typical perifissural nodule (PFN) was defined as a nodule that was lentiform, triangular, or polygonal in shape, located on or within 10 mm of the visceral pleura or lung fissure, and had extending linear densities; an atypical PFN was defined as a nodule that lacked one of the three key criteria defining a typical PFN; and a non-PFN was defined as a noncalcified solid nodule that did not fit in the previous two categories or that showed at least one of the following features: spiculation, irregular shape, unsharp borders, and distortion of the pleura or fissure. In the table, PFN encompasses both typical and atypical PFN classifications. There was variable agreement between radiologists in the classification of nodules as PFNs or non-PFNs and even less when subdividing PFNs into typical and atypical PFNs.

ules are summarized in Table 1. The median nodule diameter and volume were 6.0 mm (25th–75th percentile, 5.2–7.5 mm) and 114 mm^3 (25th–75th percentile, 72–219 mm^3), respectively. The median l/s diameter ratio and percentage sphericity were 1.45 (25th–75th percentile, 1.24–1.70) and 51.1% (25th–75th percentile, 37.7%–67.9%), respectively. Almost half of the nodules (47.5% [150 of 316]) were located in the upper lobes, and 16.5% (52 of 316) were attached to a fissure. Less than one-fifth of the nodules (16.8% [53 of 316]) were analyzed with a section thickness between 1.0 and 1.3 mm. Table 1 includes descriptive statistics for benign and malignant nodules separately.

Interreader Variability

The median proportion of nodules classified as typical or atypical PFNs was 45.6% (144 of 316). Reader A was the most conservative, with 33.2% PFN labels (105 of 316), and reader B was the most liberal, with 60.1% PFN labels (190 of

316) (Fig 2). The proportions of nodules classified by all six readers to at least two readers as PFNs ranged from 17.7% (56 of 316) to 60.8% (192 of 316) (Table 2); the Fleiss κ was 0.50 and the Cohen κ ranged from 0.32 (95% confidence interval [CI]: 0.23, 0.41) to 0.63 (95% CI: 0.55, 0.72) (Table 3). Compared with nodules classified by at least half of the readers as non-PFNs, PFNs were found to be smaller and more often located in the lower lobes and attached to a fissure ($P < .001$) (Table 4). The l/s diameter ratio, percentage sphericity, and pleural attachment were not good distinguishing factors ($P = .15$, $P = .65$, and $P = .54$, respectively).

Interreader agreement was lower for distinguishing the typical and atypical PFN subgroups: The Fleiss κ was 0.35, and the Cohen κ ranged from 0.17 (95% CI: 0.11, 0.23) to 0.53 (95% CI: 0.46, 0.61) (Table 3). Only seven (2.2%) of 316 nodules were uniformly classified by all six readers as typical PFNs (Fig 3a); only one nodule was classified as an atypical PFN agreed on by five readers (Fig 3b). Results of a subanalysis of typical PFNs (according to at least three readers) directly attached to a fissure are included in Table 4. Table 2 shows the number of readers in agreement for each nodule's classification; the sum of typical and atypical PFNs may be greater than the total number of nodules because some nodules were classified as both subtypes by different readers.

Table 3 also shows the pairwise agreement when analyzing the interreader agreement for cancer and benign nodule subgroups. For cancer nodules, the only statistically significant Cohen κ values found were between readers C and E at 0.52 (95% CI: 0.16, 0.86) and 0.52 (95% CI: 0.16 to 0.87) for the original and simplified classification systems, respectively.

Misclassified Lung Cancers

For 70 cancers, at least one reader misclassified 13 (18.6%) as a PFN, and two (2.9%) cancers were mis-

Table 3: Cohen κ Values between Each Pair of Readers

Reader	Reader A	Reader B	Reader C	Reader D	Reader E	Reader F
All Nodules (n = 316)						
A	1.00	0.32 (0.23, 0.41)	0.35 (0.26, 0.44)	0.47 (0.37, 0.58)	0.41 (0.32, 0.50)	0.51 (0.41, 0.61)
B	0.17 (0.11, 0.23)	1.00	0.64 (0.56, 0.73)	0.45 (0.37, 0.53)	0.59 (0.50, 0.67)	0.56 (0.47, 0.64)
C	0.20 (0.13, 0.27)	0.50 (0.43, 0.58)	1.00	0.47 (0.39, 0.56)	0.57 (0.48, 0.66)	0.63 (0.55, 0.71)
D	0.27 (0.20, 0.35)	0.39 (0.32, 0.46)	0.39 (0.32, 0.47)	1.00	0.45 (0.36, 0.54)	0.63 (0.55, 0.72)
E	0.32 (0.23, 0.40)	0.38 (0.31, 0.44)	0.33 (0.26, 0.40)	0.26 (0.19, 0.33)	1.00	0.55 (0.46, 0.64)
F	0.29 (0.22, 0.36)	0.50 (0.42, 0.57)	0.53 (0.46, 0.61)	0.51 (0.43, 0.59)	0.36 (0.29, 0.43)	1.00
Cancer Nodules (n = 70)						
A	1.00	−0.06 (−0.10, −0.01)	0.30 (−0.20, 0.80)	−0.04 (−0.07, −0.00)	0.13 (−0.18, 0.44)	0.00
B	−0.04 (−0.08, −0.01)	1.00	−0.06 (−0.10, −0.01)	0.26 (−0.19, 0.70)	0.24 (−0.10, 0.58)	0.00
C	0.14 (−0.11, 0.38)	−0.04 (−0.08, −0.01)	1.00	0.38 (−0.18, 0.93)	0.52 (0.16, 0.87)	0.00
D	−0.03 (−0.06, 0.00)	0.26 (−0.18, 0.70)	0.38 (−0.17, 0.93)	1.00	0.16 (−0.16, 0.49)	0.00
E	0.04 (−0.10, 0.19)	0.25 (−0.08, 0.59)	0.52 (0.16, 0.88)	0.17 (−0.16, 0.49)	1.00	0.00
F	0.00	0.00	0.00	0.00	0.00	1.00
Benign Nodules (n = 246)						
A	1.00	0.20 (0.11, 0.29)	0.22 (0.12, 0.32)	0.42 (0.31, 0.54)	0.34 (0.23, 0.44)	0.44 (0.33, 0.55)
B	0.09 (0.03, 0.15)	1.00	0.52 (0.40, 0.64)	0.33 (0.24, 0.42)	0.48 (0.36, 0.59)	0.44 (0.34, 0.54)
C	0.12 (0.04, 0.19)	0.40 (0.31, 0.49)	1.00	0.36 (0.27, 0.46)	0.43 (0.31, 0.55)	0.53 (0.44, 0.63)
D	0.23 (0.14, 0.31)	0.30 (0.22, 0.37)	0.30 (0.22, 0.38)	1.00	0.37 (0.27, 0.48)	0.58 (0.48, 0.68)
E	0.26 (0.17, 0.35)	0.29 (0.21, 0.36)	0.21 (0.13, 0.30)	0.20 (0.12, 0.27)	1.00	0.48 (0.37, 0.58)
F	0.23 (0.15, 0.31)	0.41 (0.32, 0.49)	0.46 (0.37, 0.54)	0.45 (0.36, 0.54)	0.30 (0.22, 0.38)	1.00

Note.—Data are interreader Cohen κ values between each pair of readers, with 95% confidence intervals in parentheses. The table is divided into three sections: from top to bottom, the analysis was done for all included nodules, for only cancer nodules, and for only benign nodules. In each table section, the interreader agreement when nodules were classified into two classes (perifissural nodule [PFN] or non-PFN) is displayed above the diagonal; the values when nodules were divided into three classes (typical PFN, atypical PFN, and non-PFN) are shown below the diagonal. In the bottom half of the table, a typical PFN had a lentiform, triangular, or polygonal shape, was located on or within 10 mm of the visceral pleura or lung fissure, and had extending linear densities; an atypical PFN lacked one of the three key criteria defining a typical PFN. A PFN could not be spiculated, have an irregular shape, have unsharp borders, or distort the pleura or fissure.

Table 4: Characteristics of Nodules Classified as PFNs, Non-PFNs, or Typical PFNs Attached to a Fissure by at Least Half of the Readers

Characteristic	Non-PFNs (n = 185)	PFNs (n = 159)	P Value*	Typical PFNs Attached to a Fissure (n = 42)
Diameter (mm)	6.7 (5.5–8.0)	5.5 (4.9–6.3)	<.001	5.4 (4.8–6.5)
Volume (mm^3)	155 (86–270)	86 (62–128)	<.001	113 (61–268)
l/s Diameter ratio	1.42 (1.23–1.67)	1.49 (1.28–1.71)	.15	1.69 (1.46–1.82)
Percentage sphericity	51.6 (37.9–59.6)	51.2 (37.6–67.4)	.65	45.0 (30.2–63.7)
Upper lobe location†‡	111 (60.0)	51 (32.1)	<.001	20 (47.6)
Attached to a fissure†	14 (7.6)	42 (26.4)	<.001	42 (100.0)
Attached to the pleura†	56 (30.3)	53 (33.3)	.54	5 (11.9)
CT section thickness ≤ 1.3 mm†	36 (19.5)	19 (11.9)	.06	5 (11.9)

Note.— Unless otherwise specified, data are medians, with 25th–75th percentile ranges in parentheses. Note that there is an overlap of nodules in the first two groups, because the presence of nodules with a classification agreed on by three readers signifies that the other three readers agreed on the other classification. On average, nodules classified as perifissural nodules (PFNs) were smaller, were more often located in the upper lobes, and were more often attached to a fissure (but not the pleura) than non-PFNs. There was no statistically significant difference in the proportion of thinner CT section thicknesses between these groups, excluding a potential source of bias. l/s Diameter ratio was calculated by dividing the nodule's long axis by its perpendicular short axis. Percentage sphericity was the percentage difference in volume between the perfect sphere volume derived from the nodule long axis and the actual nodule volume. Although there was no difference between non-PFNs and PFNs, most typical PFNs attached to fissures were "flatter" in shape than other nodules.

* For comparison between non-PFNs and PFNs, the Mann-Whitney U test was used for non-normal continuous outcomes, and the Pearson χ^2 was used for binary outcomes.

† Data are numbers of nodules, with percentages in parentheses.

‡ Includes major, minor, and accessory fissures (excluding the pleura).

classified by three readers; none of the malignancies were misinterpreted by more than three readers. Hence, 81.4% (57 of 70) of the cancers were classified unanimously as non-PFNs. Per reader, the error rate ranged from 0% (0 of 125) to 4.9% (eight of 163) for all PFNs, 0% (0 of 103) to 3.6% (two of 55) for typical PFNs, and 0% (0 of 22) to 9.1% (four of 44) for atypical PFNs (Fig 2). For nodules attached to a fissure (n = 52), only one of the six readers misclassified one cancer as a typical PFN (one of 12), and only one other reader misclassified another cancer as an atypical PFN (one of 25).

The median diameter of misclassified malignancies was 7.0 mm (25th–75th percentile, 5.9–7.5 mm) (Table 5). Eleven of the 13 cancers misclassified at least once were located in the upper lobes; the remaining two were located in the middle lobe and the lower lobe. Two of 13 were attached to fissures between the middle and upper lobes; one fissure was incomplete, and the second was very subtly displaced. Out of all occasions a nodule was classified as a typical PFN, 0.9% (five of 533) turned out to be cancer; for atypical PFNs, the proportion was 4.8% (16 out of 332) (P < .001). Figure 3c and 3d are screenshots of two cancers misclassified by three readers as PFNs.

Discussion

Being able to correctly identify benign PFNs at CT can potentially have a large impact on reducing the number of false-positive findings in a lung cancer screening setting. However, such a setup requires a reliable and highly accurate distinction of PFNs from small solid malignancies. To our knowledge, this is the first study quantifying interreader variability in the classification of PFNs and the risk of misclassifying small solid malignancies as PFNs. In our study, the Fleiss κ of six experienced thoracic radiologists was fair to moderate (21). The Cohen κ for reader pairs ranged from slight to moderate agreement; among cancer nodules, statistically significant agreement was found between only one pair. This may reflect the lack of a uniform definition of PFNs in the literature, as well as a range of viewpoints from experienced radiologists as to what constitutes a PFN.

To date, to our knowledge, there have only been two published CT-based studies—by Ahn et al (18) and de Hoop and colleagues (19)—that focused on outcomes in PFNs. Each used different PFN definitions, but both showed a 0% risk of malignancy in PFNs. Additionally, there have been a number of small case series describing the CT features of histopathologically confirmed IPLNs (7–17). We decided to remove the distinction between fissures and pleura in our definition of PFN. Although definitions of PFNs had been provided as a guide, readers may have chosen to label PFNs based on their existing practice.

de Hoop and colleagues (19) were the first to make the distinction between typical and atypical PFNs on the basis of the complete or incomplete presence of features characterizing PFNs. We showed that it it was five times as likely that a nodule labeled as an atypical PFN turned out to be a cancer compared with a typical PFN (P < .001). Indeed, although the error rate of misclassified malignancies per reader for all PFNs was between 0% (0 of 125) and 4.9% (eight of 163), the rate was between 0% (0 of 103) and 3.6% (two of 55) when considering only typical PFNs. Consequently, this indicates that it may be important to distinguish between typical and atypical PFNs and thus formulate different follow-up protocols. With respect to direct fissural attachment, we found that almost 90% (42 of 47) of the benign nodules attached to a fissure were classified by at least three readers as a typical PFN. On the other hand, there were also five cancers attached to a fissure; one of the five was

classified as a typical PFN by one reader and another one was classified as an atypical PFN by one reader. Although fissural attachment is a strong predictor for classification as a typical PFN, it would not completely eliminate the risk of misclassification. Additionally, less than 20% (47 of 246) of benign noncalcified solid nodules would then be eligible for classification as IPLNs. Therefore, despite its nomenclature, we prefer to state that fissural attachment is not required for defining a PFN. It should be noted that our study examined a data set that was enriched with malignancies, and the misclassification rate is likely to be considerably lower in routine screening practice.

Our results support previous reports that lung cancer is more common in the upper lobes and that IPLNs are mostly detected in the lower lobes or inferior to the carina (2,7–13,17). Indeed, only one of the 13 cancers misclassified as a PFN was located in the lower lobe; another was located in the middle lobe. To reduce the number of potential misclassifications in a lung cancer screening setting, increased caution when classifying a nodule in the upper lobe as a PFN should be considered.

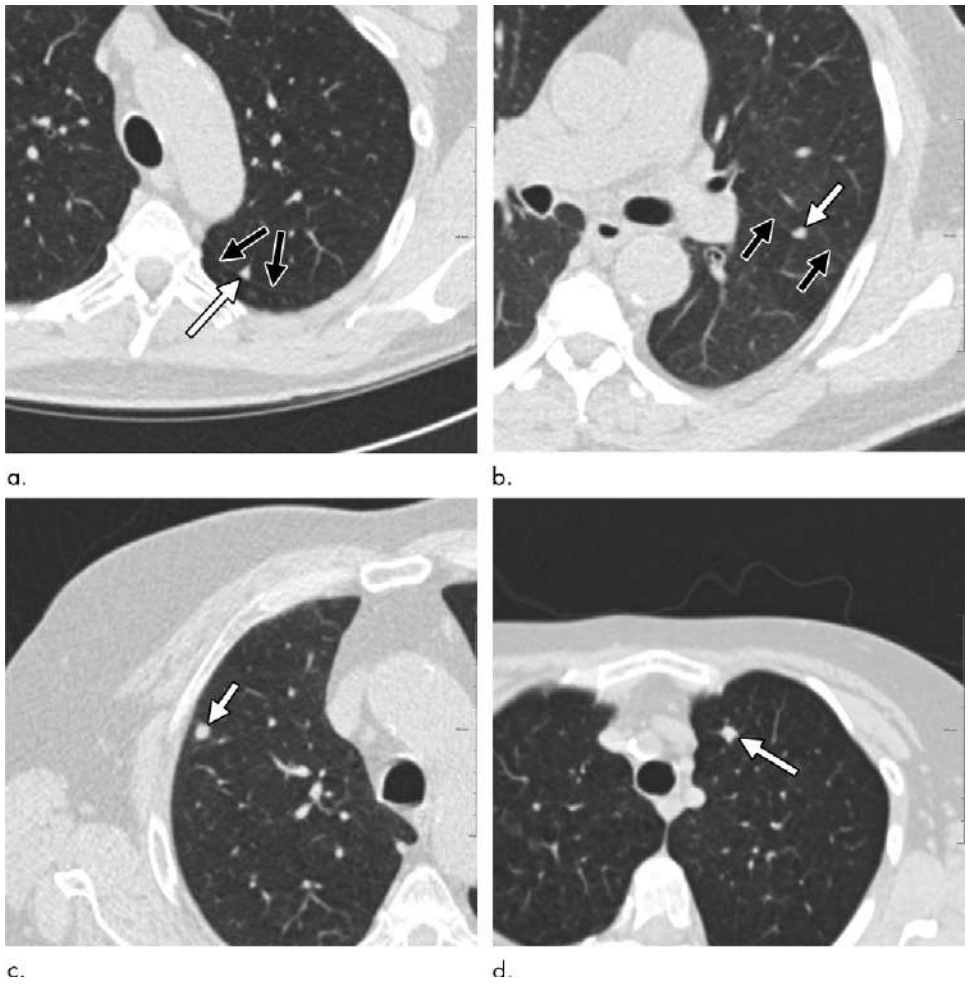

Figure 3: Axial CT images show **(a)** one of the seven nodules classified unanimously as a typical perifissural nodule (PFN), **(b)** an atypical PFN according to five of the six readers, **(c)** a cancer misclassified three times as an atypical PFN, and **(d)** a cancer misclassified twice as a typical PFN and once as an atypical PFN. White arrows = the nodule; black arrows = the pulmonary fissure.

Each nodule was presented in the center of the image, eliminating any search and thus any detection errors. The "non-PFN" category was provided for any type of solid opacity that would not fulfill the criteria for being a (typical or atypical) PFN, while the category "not applicable" was for calcified, nonnodular, or subsolid lesions. Malignancies could therefore be categorized in both categories dependent on their morphology. Among the 43 nodules being classified as not applicable by at least three readers, 11 were malignancies. This does not mean that readers did not recognize them as potential malignancies. Furthermore, 7.1% (five of 70) of all malignancies were attached to a fissure and 21.4% (15 of 70) were attached to the pleura; although this was significantly less than the rate among benign nodules, this finding supports the idea that proximity to a fissure or pleura does not guarantee benignity.

Despite definitions being provided for all readers, the readers were asked to perform as they would in a clinical setting. Readers were not asked to document on which features they based their classification. This was done for time constraints but also to illustrate that it was not the purpose of the study to test the diagnostic accuracy of or visual agreement on the presence of specific morphologic features. This would explain why there was no statistically significant difference in the proportion of pleurally attached nodules between nodules classified as PFNs and those classified as non-PFNs (P = .54, Table 4) by at least half of the readers. While it is obvious that the observers had a certain gestalt in their mind of what an IPLN represents, considerable observer variability exists with respect to the presence of one or more of the defining morphologic features. This observer variability with respect to strictly defined morphologic features has been described in previous publications (22,23). Therefore, we cannot draw any conclusions about the clinical reliability of specific morphologic features, either alone or in combination. Although there was no pathologic confirmation to prove that the nodules classified as PFNs were IPLNs, we are confident that these nodules were not cancers because of the 5- to 7-year follow-up period.

The main limitation of our study was that most of the CT studies had been obtained with a 2-mm section thickness, limiting the possibility for careful appraisal of nodule morphology in

Table 5: Characteristics of Lung Cancers Misclassified as PFNs by at Least One Reader

					Reader Classification†					
Diameter (mm)	Volume (mm³)	Lobe	Fissural Attachment*	Pleural Attachment	A	B	C	D	E	F
7.5	226	Left upper	No	No		A			A	
7.3	206	Left upper	No	No	A		T		T	
7.5	224	Right upper	No	No			A		A	
7.4	213	Right upper	No	No			A	A	A	
6.2	123	Right upper	No	No		A			A	
5.8	104	Left upper	No	No		A				
7.5	222	Right upper	No	No					T	
5.2	76	Right middle	Yes	No	T					
6.0	115	Right upper	No	No		A		A		
7.5	224	Left lower	No	No					A	
7.4	210	Left upper	No	Yes	A					
5.0	67	Left upper	No	No		T				
6.7	156	Right upper	Yes	Yes					A	

Note.—This table shows the nodule-specific characteristics of cancers misclassified as perifissural nodules (PFNs), including whether the anonymized reader had given the typical or the atypical PFN classification. A typical PFN had a lentiform, triangular, or polygonal shape, was located on or within 10 mm of the visceral pleura or lung fissure, and had extending linear densities; an atypical PFN lacked one of the three key criteria defining a typical PFN. A PFN could not be spiculated, have an irregular shape, have unsharp borders, or distort the pleura or fissure. There does not seem to be a clear misclassification pattern among the radiologists.

* Includes major, minor, and accessory fissures (excluding the pleura).

† A = atypical PFN; T = typical PFN.

the sagittal and coronal views; the BTS guidelines advise that a reconstruction thickness of no more than 1.25 mm should be used (grade C recommendation) (3). A future observer study comparing nodule classification in a mixture of CT studies obtained with 1-mm section thickness and CT studies resampled with 2- or 3-mm section thickness would help elucidate the effect of section thickness on the classification of PFNs. Although the results of this study are based exclusively on screening CT scans, we expect that they are also valid for interpreting nodular opacities in clinical CT scans if the clinical context (eg, history of malignancy) is also adequately taken into account.

In summary, we showed that the interreader agreement for classifying small solid nodules as PFNs was moderate at best. Although it is a small risk, there is a risk of misclassifying malignancies as PFNs, in particular when there are atypical features, when the nodule is not attached to a fissure, and when the nodule is located in the upper lobes. However, all readers were able to correctly identify the majority of benign nodules as PFNs (≥97% [844 of 865]). To decrease the cancer misclassification rate, we suggest being more cautious with nodules classified as atypical PFNs and nodules located in the upper lobes; the latter guideline is objective and is therefore expected to be more effective. By correctly identifying PFNs, lung cancer CT screening readers have the possibility of substantially reducing false-positive follow-up studies while minimizing the risk of delaying cancer diagnoses.

Author contributions: Guarantors of integrity of entire study, A.S., C.J., C.M.S.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, A.S., B.v.G., C.J., A.D., C.M.S.; clinical studies, N.S., S.R.D.; experimental studies, B.v.G.; statistical analysis, A.S., C.M.S.; and manuscript editing, all authors

Disclosures of Conflicts of Interest: **A.S.** disclosed no relevant relationships. **B.v.G.** Activities related to the present article: institution has received a grant from KWF. Activities not related to the present article: disclosed no relevant relationships. Other relationships: disclosed no relevant relationships. **E.T.S.** disclosed no relevant relationships. **C.J.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: institution has received a grant and receives royalties from MeVis Medical Solutions. Other relationships: disclosed no relevant relationships. **M.P.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: has received research funding from Toshiba; is on the speakers bureaus of Bracco, Bayer, Toshiba, and Siemens. Other relationships: department spin-off (no personal financial interest) at Thiroux. **N.S.** disclosed no relevant relationships. **S.R.D.** disclosed no relevant relationships. **A.D.** disclosed no relevant relationships. **C.M.S.** disclosed no relevant relationships.

References

1. GLOBOCAN 2012: estimated cancer incidence, mortality and prevalence worldwide in 2012. France: IARC. http://globocan.iarc.fr. Published 2017. Accessed September 4, 2017.
2. National Lung Screening Trial Research Team, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365(5):395–409.
3. Callister MEJ, Baldwin DR, Akram AR, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules. Thorax 2015;70(Suppl 2):ii1–ii54 [Published correction appears in Thorax 2015;70(12):1188.].
4. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of incidental pulmonary nodules detected on ct images: from the Fleischner Society 2017. Radiology 2017;284(1):228–243.
5. Lung Imaging Reporting and Data System (Lung-RADS). American College of Radiology. https://www.acr.org/Quality-Safety/Resources/LungRADS. Published 2014. Accessed September 4, 2017.
6. NCCN Guidelines: lung cancer screening. National Comprehensive Cancer Network. Published 2014. https://www.tri-kobe.org/nccn/guideline/lung/english/lung_screening.pdf. Accessed October 3, 2017.
7. Kradin RL, Spirn PW, Mark EJ. Intrapulmonary lymph nodes: clinical, radiologic, and pathologic features. Chest 1985;87(5):662–667.
8. Bankoff MS, McEniff NJ, Bhadelia RA, Garcia-Moliner M, Daly BD. Prevalence of pathologically proven intrapulmonary lymph nodes and their appearance on CT. AJR Am J Roentgenol 1996;167(3):629–630.
9. Yokomise H, Mizuno H, Ike O, Wada H, Hitomi S, Itoh H. Importance of intrapulmonary lymph nodes in the differential diagnosis of small pulmonary nodular shadows. Chest 1998;113(3):703–706.
10. Miyake H, Yamada Y, Kawagoe T, Hori Y, Mori H, Yokoyama S. Intrapulmonary lymph nodes: CT and pathological features. Clin Radiol 1999;54(10):640–643.

11. Matsuki M, Noma S, Kuroda Y, Oida K, Shindo T, Kobashi Y. Thin-section CT features of intrapulmonary lymph nodes. J Comput Assist Tomogr 2001;25(5):753–756.
12. Sykes A-MG, Swensen SJ, Tazelaar HD, Jung S-H. Computed tomography of benign intrapulmonary lymph nodes: retrospective comparison with sarcoma metastases. Mayo Clin Proc 2002;77(4):329–333.
13. Oshiro Y, Kusumoto M, Moriyama N, et al. Intrapulmonary lymph nodes: thin-section CT features of 19 nodules. J Comput Assist Tomogr 2002;26(4):553–557.
14. Hyodo T, Kanazawa S, Dendo S, et al. Intrapulmonary lymph nodes: thin-section CT findings, pathological findings, and CT differential diagnosis from pulmonary metastatic nodules. Acta Med Okayama 2004;58(5):235–240.
15. Ishikawa H, Koizumi N, Morita T, Tsuchida M, Umezu H, Sasai K. Ultrasmall intrapulmonary lymph node: usual high-resolution computed tomographic findings with histopathologic correlation. J Comput Assist Tomogr 2007;31(3):409–413.
16. Honma K, Nelson G, Murray J. Intrapulmonary lymph nodes in South African miners--an autopsy survey. Am J Ind Med 2007;50(4):261–264.
17. Wang CW, Teng YH, Huang CC, Wu YC, Chao YK, Wu CT. Intrapulmonary lymph nodes: computed tomography findings with histopathologic correlations. Clin Imaging 2013;37(3):487–492.
18. Ahn MI, Gleeson TG, Chan IH, et al. Perifissural nodules seen at CT screening for lung cancer. Radiology 2010;254(3):949–956.
19. de Hoop B, van Ginneken B, Gietema H, Prokop M. Pulmonary perifissural nodules on CT scans: rapid growth is not a predictor of malignancy. Radiology 2012;265(2):611–616.
20. Larke FJ, Kruger RL, Cagnon CH, et al. Estimated radiation dose associated with low-dose chest CT of average-size participants in the National Lung Screening Trial. AJR Am J Roentgenol 2011;197(5):1165–1169.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–174.
22. Chung K, Jacobs C, Scholten ET, et al. Malignancy estimation of Lung-RADS criteria for subsolid nodules on CT: accuracy of low and high risk spectrum when using NLST nodules. Eur Radiol 2017;27(11):4672–4679.
23. van Riel SJ, Ciompi F, Winkler Wille MM, et al. Malignancy risk estimation of pulmonary nodules in screening CTs: comparison between a computer model and human observers. PLoS One 2017;12(11):e0185032.