The Utility of Nodule Volume in the Context of Malignancy Prediction for Small Pulmonary Nodules

Hiren J. Mehta, MD; James G. Ravenel, MD; Stephanie R. Shaftman, MSc, MS; Nichole T. Tanner, MD, FCCP; Luca Paoletti, MD, FCCP; Katherine K. Taylor, MS; Martin C. Tammemagi, PhD; Mario Gomez, MD; Paul J. Nietert, PhD; Michael K. Gould, MD, FCCP; and Gerard A. Silvestri, MD, FCCP

Background: An estimated 150,000 pulmonary nodules are identified each year, and the number is likely to increase given the results of the National Lung Screening Trial. Decision tools are needed to help with the management of such pulmonary nodules. We examined whether adding any of three novel functions of nodule volume improves the accuracy of an existing malignancy prediction model of CT scan-detected nodules.

Methods: Swensen's 1997 prediction model was used to estimate the probability of malignancy in CT scan-detected nodules identified from a sample of 221 patients at the Medical University of South Carolina between 2006 and 2010. Three multivariate logistic models that included a novel function of nodule volume were used to investigate the added predictive value. Several measures were used to evaluate model classification performance.

Results: With use of a 0.5 cutoff associated with predicted probability, the Swensen model correctly classified 67% of nodules. The three novel models suggested that the addition of nodule volume enhances the ability to correctly predict malignancy; 83%, 88%, and 88% of subjects were correctly classified as having malignant or benign nodules, with significant net improved reclassification for each ($P < .0001$). All three models also performed well based on Nagelkerke R^2 , discrimination slope, area under the receiver operating characteristic curve, and Hosmer-Lemeshow calibration test.

Conclusions: The findings demonstrate that the addition of nodule volume to existing malignancy prediction models increases the proportion of nodules correctly classified. This enhanced tool will help clinicians to risk stratify pulmonary nodules more effectively.

CHEST 2014; 145(3):464-472

Abbreviations: AUC = area under the curve; LDCT = low-dose CT; NRI = net reclassification improvement

The pulmonary nodule is a single, spherical, well-circumscribed, radiographic opacity that measures < 3 cm in diameter and is completely surrounded by aerated lung. There is no associated atelectasis, hilar enlargement, or pleural effusion.¹ Approximately 150,000 such nodules are identified each year according to dated estimates.^{2,3} The incidence is likely much higher than this because of the increasing use of chest CT scan for the evaluation of a myriad of pulmonary symptoms and disorders. The National Lung Screening Trial has shown screening patients with low-dose CT (LDCT) scanning led to a relative risk reduction in death from lung cancer by 20%.⁴ Over the 3-year screening period, however, 39.1% of the participants

in the LDCT scanning group had a nodule discovered, of which (96.4%) were benign.⁴ Currently, 7 million Americans meet the National Lung Screening Trial screening criteria.^{4,5} Even if only one-fourth of those eligible are screened, a possible 680,000 new nodules could be discovered over 3 years.

For editorial comment see page 440

When faced with a patient with a pulmonary nodule, it is incumbent on the clinician to differentiate benign cases from cancer. Although most clinicians use clinical experience to estimate the probability of malignancy in pulmonary nodules, some rely on one or

more quantitative models for assistance.^{3,6} The American College of Chest Physicians guidelines on pulmonary nodule evaluation recommend the use of models to predict the probability of malignancy.⁷ Several such models have been proposed.^{3,8} Of particular interest is the Swensen model,⁸ which has been widely cited⁹ and externally validated^{10,11} and is commonly used in clinical practice. We undertook the present study to examine whether the addition of a measurement of nodule volume to the Swensen model would enhance its ability to predict malignancy in patients who present with small pulmonary nodules.

MATERIALS AND METHODS

Subjects

Consecutive subjects with at least one newly diagnosed pulmonary nodule were recruited between February 1, 2006, and May 1, 2010, from our institution's pulmonary clinic and were included if they had a thoracic CT scan, at least one nodule < 15 mm in diameter detected on the scan, and images satisfactory for volume assessment (slice thickness $\leq$ 2.5 mm). Although we had no lower limit cutoff per se, on the basis of detectability, no nodules were < 3 mm in size. The Institutional Review Board at the Medical University of South Carolina (HR#15125) approved this prospective study.

Study Variables

Clinical data (demographics, smoking status, and cancer history) and nodule characteristics (long-axis diameter, volume, spiculation, location, and nodule count) were recorded. Subjects were prospectively followed for a minimum of 24 months. A definitive diagnosis of malignancy was established either by surgical resection, CT image-guided needle biopsy, or serial follow-up CT scans demonstrating stability for at least 2 years. The study protocol included the collection of patient-specific variables (age, sex, race/ethnicity, FEV₁ % predicted, smoking history, secondhand

smoke exposure, personal history of malignancy, family history of malignancy, and history of asbestos or radon exposure). A single thoracic radiologist blinded to outcomes reviewed all the images to detect calcified lymph nodes, noncalcified enlarged lymph nodes, nodule cavitation, nodule density (solid, part solid, or ground glass), edge characteristics (smooth, lobulated, or spiculated), presence or absence of pleural tail, and location by lobe. Nodules were assessed by a semiquantitative volume program (Lung VCAR; General Electric Co). Diameter was measured in the transverse plane by electronic calipers and recorded as the average of the long-axis measurement and short axis measured perpendicular to the long axis. Volume was obtained by single mouse click on the nodule (thus invoking the Lung VCAR algorithm), and the segmented volume was recorded. No manual adjustments were made. See Figure 1 for CT image reconstruction of a pulmonary nodule.

Statistical Analysis

Simple descriptive statistics were used to characterize subjects and nodules with and without malignancy. Wilcoxon rank sum tests, *t* tests, χ^2 tests, and Fisher exact tests were initially used to highlight variables that differed significantly (ie, $P < .05$) between subjects in whom malignancy did and did not develop and between nodules that did and did not ultimately become malignant.

Following descriptive analyses, the Swensen model⁸ was applied to all nodules to obtain a preliminary probability of malignancy for each individual nodule. We then investigated in three separate models whether the addition of nodule volume (model 1), volume to diameter ratio (model 2), or sphericity index (model 3) resulted in improved prediction performance. The sphericity index of a nodule is a metric developed by the study investigators for this project and is the observed automated volume measurement divided by the volume of a sphere with a diameter equal to that of the nodule. Generalized linear mixed models with restricted cubic splines were used to construct and assess the novel prediction models¹² and accounted for random subject effects and the dependence of measurements made within the same individual. The use of splines provided a means for nonlinear modeling with effects that vary across predictor values. Spline knot locations (which define the segments) were based on recommendations by Harrell.¹³ The significance of adding the various functions of nodule volume and the accompanying splines to the prediction models underwent likelihood ratio tests. SAS, version 9.3 (SAS Institute, Inc) software was used for model construction, and the regression modeling strategies package within R, version 2.15.0,¹⁴ was used for internal validation.

Prediction models cannot be compared by a single statistic; therefore, we used an approach similar to the multifaceted framework for performance assessment outlined by Steyerberg et al.¹⁵ Prediction model performance measures used in the analysis included Nagelkerke R^2 , area under the curve (AUC), discrimination slope, Hosmer-Lemeshow goodness-of-fit statistics, and net reclassification improvement (NRI).¹⁶ Nagelkerke R^2 is a coefficient of determination and represents how well the model predicts outcomes on the basis of the information provided in the model. The AUC represents an overall measure of how well the model discriminates between those with and without the condition in question. Discrimination slope is the difference in mean probabilities between nodules with and without malignancy. NRI quantifies improvement in classification (ie, malignant vs benign) at a selected cutoff (eg, 0.5) for the predicted probability of a model when compared with a standard model (ie, the Swensen model). Internal validation was conducted with Harrell's optimism correction technique using 2,000 bootstrapped resamples,^{15,17} which we implemented with the regression modeling strategies package in R. The optimism correction internal validation process provides an estimate of the degree of bias introduced by potential model

Manuscript received March 22, 2013; revision accepted June 25, 2013.

Affiliations: From the Division of Pulmonary, Critical Care, and Sleep Medicine (Dr Mehta), University of Florida College of Medicine, Gainesville, FL; Department of Radiology and Radiological Sciences (Dr Ravenel), Division of Biostatistics and Epidemiology (Ms Shaftman and Dr Nietert), and Division of Pulmonary and Critical Care Medicine (Drs Tanner, Paoletti, and Silvestri and Ms Taylor), Department of Medicine, Medical University of South Carolina, Charleston, SC; Department of Community Health Sciences (Dr Tammemagi), Brock University, St. Catharines, ON, Canada; Pulmonary & Sleep Center of the Valley (Dr Gomez), Weslaco, TX; and Department of Research and Evaluation (Dr Gould), Kaiser Permanente Southern California, Pasadena, CA.
Funding/Support: This study was supported by the Department of Defense [award W81XWH-05-1-0378], the National Cancer Institute [award 5K24CA120494], and the National Center for Research Resources [award 5UL1RR029882].

Correspondence to: Gerard A. Silvestri, MD, FCCP, Division of Pulmonary and Critical Care Medicine, Department of Medicine, Medical University of South Carolina, 96 Jonathan Lucas St, CSB 810, Charleston, SC 29425; e-mail: silvestri@musc.edu

© 2014 American College of Chest Physicians. Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details.

DOI: 10.1378/chest.13-0708

FIGURE 1. A, CT image of a lung nodule. B, Volumetric reconstruction of the same nodule.

overfitting; in other words, it quantifies the extent to which the predictive ability of a model can be expected to degrade when applied to a different set of data. Because of the nature of diameter and volume being inherently correlated, sensitivity analyses were performed in which models were constructed and assessed without nodule diameter being included.

RESULTS

Two hundred thirty-three nodules were observed among 221 subjects, with 37% being malignant. The average $\pm$ SD age was 62.4 ± 10.2 years, and 79% were classified as ever smokers. Prior extrathoracic cancers were noted in 34% of subjects. The average nodule diameter and volume were 9.2 ± 3.2 mm (range, 3.0–14.9 mm) and 723 mm^3 , respectively. Table 1 shows a comparison between study subjects with

benign nodules and those with malignant nodules; no significant differences between groups were noted regarding age, sex, race, or smoking status. Among ever smokers, pack-years were significantly higher in those with malignant nodules (mean, 67.6 ± 46.0) vs those with benign nodules (mean, 42.7 ± 42.0). The prevalence of nonpulmonary cancer history was also significantly higher among subjects whose nodules were ultimately determined to be malignant (47.5% vs 26.2%). Comparisons in nodule characteristics among malignant and benign nodules are shown in Table 2. Malignant nodules tended to have significantly different edge characteristics than benign nodules, with malignant nodules more likely than benign nodules to be considered lobulated or spiculated. Malignant nodules were also significantly more likely to have a larger diameter, volume, volume to diameter ratio, and sphericity

Table 1—Baseline Clinical Characteristics, Smoking Status, and Cancer History of Study Subjects

Clinical Characteristic	No Malignant Nodules Present (n = 141)	≥ 1 Malignant Nodule Present (n = 80)
Age, y	62.2 ± 10.4	62.9 ± 9.9
Male sex	46 (32.6)	32 (40.0)
Race		
White, non-Hispanic	115 (81.6)	61 (76.3)
Black, non-Hispanic	21 (14.9)	13 (16.3)
Hispanic	1 (0.7)	1 (1.3)
Other	1 (0.7)	2 (2.5)
Not reported	3 (2.1)	3 (3.8)
Smoking status		
Current or former smoker	108 (76.6)	66 (82.5)
Pack-y among current and former smokers	42.7 ± 42.0	67.6 ± 46.0^a
Secondhand smoke exposure		
Ever lived with a smoker	107 (75.9)	70 (87.5) ^a
Ever worked with a smoker	88 (62.4)	60 (75.0)
Reported asbestos exposure	11 (7.8)	8 (10.0)
Reported radon exposure	1 (0.7)	0 (0.0)
Pulmonary function testing		
FEV ₁ , L	2.1 ± 0.9 (n = 76)	1.9 ± 0.8 (n = 57)
FEV ₁ , % predicted	80.4 ± 29.6 (n = 76)	76.2 ± 26.4 (n = 57)
Nonpulmonary cancer history	37 (26.2)	38 (47.5)

Data are presented as mean $\pm$ SD or No. (%).

^a $P < .05$ by t test, χ^2 test, or Fisher exact test, as appropriate, compared with subjects with no malignant nodules present.

Table 2—Baseline Characteristics of Nodules Among Study Subjects

Nodule Characteristic	Benign Nodules (n = 148)	Malignant Nodules (n = 85)
Nodule edge	^a	^a
Entirely smooth	74 (50.0)	17 (20.0)
Lobulated	31 (21.0)	21 (24.7)
Spiculated	29 (19.6)	38 (44.7)
Ill defined	14 (9.5)	9 (10.6)
Cavitation	3 (2.0)	3 (3.5)
Location		
Central	19 (12.8)	8 (9.4)
Peripheral	129 (87.2)	77 (90.6)
RUL	27 (18.2)	26 (30.6)
RML	14 (9.5)	7 (8.2)
RLL	41 (27.7)	18 (21.2)
LUL	19 (12.8)	15 (17.7)
Lingula	7 (4.7)	5 (5.9)
LLL	40 (27.0)	14 (16.5)
Diameter, cm	0.83 ± 0.31 ^a	1.08 ± 0.28 ^a
Novel measures		
incorporating nodule volume		
Volume, cm ³	0.48 ± 0.54 ^a	1.14 ± 0.80 ^a
V:D ratio, cm ²	0.48 ± 0.42 ^a	0.99 ± 0.60 ^a
SI	1.37 ± 1.20 ^a	1.73 ± 1.46 ^a

Data are presented as No. (%) or mean ± SD. LLL = left lower lobe; LUL = left upper lobe; RLL = right lower lobe; RML = right middle lobe; RUL = right upper lobe; SI = sphericity index; V:D = volume:diameter. ^a*P* < .05 comparing the two groups by *t* test, χ^2 test, or Fisher exact test, as appropriate.

index. Table 3 shows frequency of malignancy based on the size of the nodules. None of the nodules < 4 mm in size and 47% of nodules between 8 and 15 mm in diameter were malignant.

The specific parameters for the Swensen model⁸ and the three novel multivariable models are listed in Table 4. Table 5 shows the adjusted ORs for the variables of each model. Compared with a model without any information about nodule volume, each novel model had significantly (*P* < .001) improved fit as evidenced by the likelihood ratio χ^2 tests. The new models also confirm the significant influence of cancer history on malignancy prediction. The influences of age, smoking status, spiculation, and upper lobe location on the probability of malignancy were slightly attenuated in the new models compared with the Swensen model.

Table 3—Malignancy Rates Based on the Size of the Nodule

Diameter Range	No. Nodules Included	Malignant Nodules
< 4	4	0 (0.00)
≥ 4 to < 6	42	3 (7.10)
≥ 6 to < 8	41	13 (31.70)
≥ 8 to < 15	146	69 (47.30)

Data are presented as No. (%) unless otherwise indicated.

All three novel models suggest that nodule volume may enhance the ability to predict malignancy, as is evident by the summary of model performance in Table 6. All models have *R*² values, AUC values, and discrimination slopes higher than the Swensen model (applied to the present data), even after adjusting for optimism. The Hosmer-Lemeshow tests for all the models were nonsignificant, suggesting sufficient calibration. Figure 2 illustrates the improvement in the AUC in models 1 to 3 compared with the Swensen model. Figure 3 illustrates the differences in the predicted probabilities of the models for malignant and benign nodules; a greater divergence between the malignant and benign nodules in Figures 3B-D compared with 3A is indicative of greater discrimination ability in each novel model. Overall, the model performance measures were relatively unchanged in the sensitivity analyses in which nodule diameter was omitted from the models.

With use of a 0.5 cutoff value for the model probability to classify nodule malignancy, the Swensen model correctly classified 67% of nodules in the present dataset. Models 1, 2, and 3 correctly classified 83%, 88%, and 88% of nodules, respectively. However, a different cutoff threshold yielded slightly varied results. For example, with a 0.2 cutoff for the probability of malignancy, the Swensen model correctly classified 68% of nodules, compared with 73%, 72%, and 73% for models 1, 2, and 3, respectively. With a 0.8 cutoff, the Swensen model correctly classified 63% of nodules, compared with 69% for models 1, 2, and 3. NRI analyses suggested that each novel model yielded a significant (*P* < .0001) net improved reclassification. Compared with the Swensen model at the 0.5 cutoff, the NRIs for models 1, 2, and 3 were 46.5%, 45.4%, and 41.2%, respectively. For all three models, the largest gains in classification improvement came from correctly reclassifying low-risk nodules (according to the Swensen model) as high risk (model 1, 50.6%; model 2, 49.4%; model 3, 45.9%) at only a slight cost of incorrectly reclassifying a few high-risk nodules as low risk (model 1, -4.1%; model 2, -4.1%; model 3, -4.7%).

DISCUSSION

This study is one of the first, to our knowledge, to focus on malignancy prediction in nodules measuring < 15 mm in diameter and has several important findings. First, the addition of nodule volume to the Swensen model significantly enhances its predictive ability. The novel models incorporating nodule volume, volume to diameter ratio, and sphericity index performed relatively similar to one another. All three outperformed the original Swensen model and correctly classified more subjects as having malignant lesions than did the Swensen model. The addition of nodule volume

Table 4—Multivariate Predictive Models: Model Parameter Values

Logistic Regression Variable	Swensen Model (Published)	Model 1	Model 2	Model 3
Model intercept	−6.8272	−2.8183 ^a	−4.0617 ^b	−5.7364 ^c
Age	0.0391	−0.0106	−0.0084	−0.0062
Ever smoker (vs never smoker)	0.7917	0.5512	0.5070	0.5394
History of cancer (vs no history)	1.3388	0.9196 ^a	0.9506 ^a	0.9137 ^a
Nodule diameter, cm	1.2740 ^d	−0.3034	0.6357	3.0762 ^e
Spiculation present (vs absent)	1.0407	0.6446	0.6610	0.6537
Upper lobe location	0.7838	0.6222	0.5577	0.6122
Volume, cm ³		^f		
Spline 1: volume (linear)	...	4.1695	...	...
Spline 2: (volume − 0.067) ³	...	−2.1404	...	...
Spline 3: (volume − 0.447) ³	...	2.6403	...	...
Spline 4: (volume − 1.624) ³	...	1.5317	...	...
V:D ratio			^f	
Spline 1: V:D ratio (linear)	...	...	5.0497	...
Spline 2: (V:D ratio − 0.1197) ³	...	...	−4.6532	...
Spline 3: (V:D ratio − 0.4836) ³	...	...	7.1544	...
Spline 4: (V:D ratio − 1.4000) ³	...	...	−0.6621	...
SI	...			^f
Spline 1: SI (linear)	...	...	...	1.0376
Spline 1: (SI − 0.54) ³	...	...	...	−0.0687
Spline 2: (SI − 1.78) ³	...	...	...	0.1674
Spline 3: (SI − 2.40) ³	...	...	...	−0.0947

See Table 2 legend for expansion of abbreviations.

^a $P < .05$.

^b $P < .01$.

^c $P < .001$.

^dThe original publication expressed diameter in millimeters; because the novel models use centimeters, we converted the parameter estimate to reflect this change.

^e $P < .0001$.

^f $P < .001$ by likelihood ratio χ^2 test.

Table 5—Multivariate Predictive Models: Model ORs and 95% CIs

Logistic Regression Variable	Swensen Model	Model 1	Model 2	Model 3
Model intercept	...	...	...	...
Age	1.04 (1.01-1.07)	0.99 (0.96-1.03)	0.99 (0.96-1.03)	0.99 (0.96-1.03)
Ever smoker (vs never smoker)	2.21 (1.17-4.16)	1.74 (0.73-4.14)	1.66 (0.69-3.99)	1.71 (0.72-4.11)
History of cancer (vs no history)	3.82 (1.39-10.5)	2.51 (1.22-5.16)	2.59 (1.25-5.34)	2.49 (1.21-5.14)
Nodule diameter, cm	3.58 (2.32-5.51) ^a	0.74 (0.10-5.34)	1.89 (0.37-9.56)	21.7 (6.16-76.2)
Spiculation present (vs absent)	2.83 (1.47-5.45)	1.91 (0.87-4.18)	1.94 (0.88-4.28)	1.92 (0.89-4.16)
Upper lobe location	2.19 (1.27-3.79)	1.86 (0.92-3.78)	1.75 (0.86-3.56)	1.84 (0.91-3.73)
Volume, cm ³				
Spline 1: volume (linear)	...	64.7 (1.8-2262)	...	...
Spline 2: (volume − 0.067) ³	...	^b	...	...
Spline 3: (volume − 0.447) ³	...	^b	...	...
Spline 4: (volume − 1.624) ³	...	^b	...	...
Volume:diameter (V:D) ratio				
Spline 1: V:D ratio (linear)	...	...	156 (2.5-9,906)	...
Spline 2: (V:D ratio − 0.1197) ³	...	...	^b	...
Spline 3: (V:D ratio − 0.4836) ³	...	...	^b	...
Spline 4: (V:D ratio − 1.4000) ³	...	...	^b	...
SI				
Spline 1: SI (linear)	...	...	...	2.82 (0.83-9.55)
Spline 1: (SI − 0.54) ³	...	...	...	^b
Spline 2: (SI − 1.78) ³	...	...	...	^b
Spline 3: (SI − 2.40) ³	...	...	...	^b

ORs are associated with a 1-unit increase in the variable of interest. See Table 2 legend for expansion of abbreviations.

^aThe original publication expressed diameter in millimeters; because the novel models use centimeters, we converted the parameter estimate to reflect such a change.

^bORs for splines are not provided because the change in odds for a unit change in a spline variable is not constant across the range of the spline variable.

Table 6—Model Performance Characteristics

Performance Measure	Swensen Model (as Applied to the Current Data)	Model 1 (Volume)	Model 2 (V:D Ratio)	Model 3 (SI)
Overall performance				
Nagelkerke R^2 (optimism corrected), %	55.4 (N/A)	59.6 (29.3)	61.4 (29.5)	59.3 (19.1)
Discrimination				
AUC (optimism corrected)	0.748 (N/A)	0.911 (0.786)	0.915 (0.784)	0.914 (0.780)
Discrimination slope	0.149	0.382	0.388	0.380
Calibration				
Hosmer-Lemeshow test	$P = .12$	$P = .29$	$P = .50$	$P = .24$

AUC = area under the curve; N/A = not applicable (optimism corrections were only determined for the novel models). See Table 2 legend for expansion of other abbreviations.

to parameters specified in the Swensen model enhanced the ability of the model to correctly predict malignancy. As evident in Figure 2, the novel models showed that 83%, 88%, and 88% of subjects were correctly classified as having malignant or benign nodules compared with 67% by the Swensen model, with a significant net improved reclassification for each ($P < .0001$).

Although it may seem that volume (which directly corresponds to diameter for a sphere) simply duplicates the size aspect of the Swensen model, it clearly adds more to the predictive value. A primary reason for this is that diameter is measured in the transverse plane, which may not be the on the axis of the nodule. In particular, diameter will underrepresent volume for lesions that are longer in the z-axis than in the x- and y-axes. Similarly, diameter does not perfectly predict volume for irregular or spiculated nodules. Thus, it is not surprising that knowledge of nodule volume pro-

vides a distinct value for malignancy prediction when added to transverse diameter.

The management of small pulmonary nodules continues to be a challenge for clinicians. Although there are no accurate estimates of the prevalence of CT scan-detected pulmonary nodules in a nonscreened population, the incidence is rising with the use of chest CT scanning in diagnosing a myriad of disorders. Should a national policy for screening for lung cancer with LDCT imaging be broadly implemented in the United States, the incidence of pulmonary nodules has the potential to rise dramatically. During the evaluation of a pulmonary nodule, clinicians usually decide on one of three management strategies: serial imaging, biopsy, or surgery. This decision is based on their initial assessment of the probability that the nodule is malignant and should take into account the patient's comorbidities and preferences.⁷ The goal should be

FIGURE 2. Improvement in area under the receiver operating characteristic curve in models 1 to 3 compared with the Swensen model. AUC = area under the curve.

FIGURE 3. Box plots of predicted probabilities associated with the Swenson model and three novel models. A, Swenson model. B, Model 1 (Swenson model plus nodule volume). C, Model 2 (Swenson model plus nodule volume:diameter ratio). D, Model 3 (Swenson model plus nodule sphericity index).

prompt diagnosis and treatment of malignant nodules and avoidance of invasive testing for those that are benign. However, one study showed that different geographic regions of the United States had vastly different biopsy rates for lung nodules (14.7-36.2 per 100,000 adults), suggesting uncertainty in how lung nodules should be managed.¹⁸ Applying our prediction models to estimate the probability of malignancy could potentially decrease the time to diagnosis and treatment of malignant pulmonary nodules. A goal of future prediction models would be to classify benign nodules more accurately to help lower rates of interventional diagnostic procedures for benign lesions.

Several models have been developed to predict the probability of malignancy in pulmonary nodules.^{8,9} These models are limited, however, by a restricted number of potential predictors, generally low overall predictive performance, and being based on larger

pulmonary nodules (> 15 mm in diameter). One unique aspect of this study compared with existing models is that it excluded pulmonary nodules ≥ 15 mm in diameter. This improves the applicability of the present study in clinical practice, where nodules < 15 mm in diameter are more common and pose a significant and greater diagnostic dilemma, especially the nodules between 8 and 15 mm in size. When used in conjunction with the existing algorithm proposed by the American College of Chest Physicians most recent lung cancer guidelines,⁷ the present models may help clinicians to be better able to predict malignancy in small pulmonary nodules. Incorporating the proposed models into clinical practice should help patients and clinicians to make better informed choices about options for evaluation.

This study, like those of all other prediction models, has some limitations. Generalizability is a potential

limitation because all subjects were observed at a single academic health center; the novel models need to be externally validated. We believe that the nodule volumes are a reasonable approximation of true nodule volume because we have previously validated its precision in a phantom model¹⁹; however, the volumetric software program used in the present study is an older version and not validated for part-solid or ground glass nodules. The results, therefore, may not be generalizable to other nodule volume platforms, which points to the need for a consistent dataset against which the accuracy of segmentation can be tested. Also important to recognize is that the novel models are only relevant during the decision-making process associated with the initial clinical assessment of an identified nodule; additional knowledge and data obtained by comparing serial CT scans could be very informative and have the potential to make even greater improvements in malignancy prediction performance.

There are also barriers to implementation. In this study, the workstation used was separate and distinct from our clinical radiology workstation. The process could potentially be more onerous and lengthy when attempting to integrate into daily work flow, although this problem should be relatively easy to overcome. Improved integration of picture archiving and communication system and other third-party workstations are needed to make this process efficient for the radiologist. In addition, there will need to be good communication between clinician and radiologist to establish the added value and necessity for volumetric measurements. Still, commercially available technology exists to perform volumetric measurements on indeterminate pulmonary nodules, and consideration should be given to how this can be integrated into clinical practice. Future studies should focus on validation of our model in other cohorts, including a population of screen-detected nodules.

With an increasing number of small pulmonary nodules being detected, clinicians will have a greater need for more accurate prognostic prediction. Others have reported that an estimate of nodule diameter and volume is helpful when a clinician decides to use serial imaging to follow a screen-detected pulmonary nodule.²⁰ The present study is the first, to our knowledge, to demonstrate that measuring different aspects of volume within a nodule at the time of nodule detection provides unique information and improves the ability of the prediction model to differentiate between benign and malignant disease.

In conclusion, the models presented in this article improve the ability to predict malignancy in a pulmonary nodule, which has an important clinical implication given the diagnostic dilemma associated with small pulmonary nodules (< 15 mm) and the evidence of substantial variation in current management. The

current practice of using clinical intuition to decide how to manage a pulmonary nodule can be augmented by the use of this model, thus, reducing unnecessary biopsy procedures in those with benign disease and promoting prompt management in those with malignancy.

ACKNOWLEDGMENTS

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

Dr Mehta: contributed to the study design, data collection, and manuscript writing.

Dr Ravenel: contributed to the study design, data collection, and manuscript writing.

Ms Shaftman: contributed to the study design, statistical analysis, and manuscript writing.

Dr Tanner: contributed to the study design, data collection, and manuscript writing.

Dr Paoletti: contributed to the study design, data collection, and manuscript writing.

Ms Taylor: contributed to the study design, data collection, and manuscript writing.

Dr Tammemagi: contributed to the study design, data collection, and manuscript writing.

Dr Gomez: contributed to the study design, data collection, and manuscript writing.

Dr Nietert: contributed to the study design, statistical analysis, and manuscript writing.

Dr Gould: contributed to the study design, data collection, and manuscript writing.

Dr Silvestri: contributed to the study design, data collection, and manuscript writing.

Financial/nonfinancial disclosures: The authors have reported to *CHEST* that no potential conflicts of interest exist with any companies/organizations whose products or services may be discussed in this article.

Role of sponsors: The National Center for Research Resources provided protected time for Dr Nietert to participate in this project. The National Institute of Health provided protected time for Dr Silvestri to mentor junior investigators involved on this project. The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the Department of Defense, the National Center for Research Resources, or the National Institutes of Health.

REFERENCES

1. Tuddenham WJ. Glossary of terms for thoracic radiology: recommendations of the Nomenclature Committee of the Fleischner Society. *AJR Am J Roentgenol*. 1984;143(3):509-517.
2. Holin SM, Dwork RE, Glaser S, Rikli AE, Stocklen JB. Solitary pulmonary nodules found in a community-wide chest roentgenographic survey; a five-year follow-up study. *Am Rev Tuberc*. 1959;79(4):427-439.
3. Lillington GA. Management of solitary pulmonary nodules. *Dis Mon*. 1991;37(5):269, 271-318.
4. Aberle DR, Adams AM, Berg CD, et al; National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365(5):395-409.
5. Goulart BH, Bensink ME, Mummy DG, Ramsey SD. Lung cancer screening with low-dose computed tomography: costs, national expenditures, and cost-effectiveness. *J Natl Compr Canc Netw*. 2012;10(2):267-275.
6. Gurney JW. Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis. Part I. Theory. *Radiology*. 1993;186(2):405-413.

7. 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(5_suppl):e93S-e120S.
8. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Arch Intern Med*. 1997;157(8):849-855.
9. Gould MK, Ananth L, Barnett PG; Veterans Affairs SNAP Cooperative Study Group. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. *Chest*. 2007;131(2):383-388.
10. Herder GJ, van Tinteren H, Golding RP, et al. Clinical prediction model to characterize pulmonary nodules: validation and added value of ¹⁸F-fluorodeoxyglucose positron emission tomography. *Chest*. 2005;128(4):2490-2496.
11. Schultz EM, Sanders GD, Trotter PR, et al. Validation of two models to estimate the probability of malignancy in patients with solitary pulmonary nodules. *Thorax*. 2008;63(4):335-341.
12. Marrie RA, Dawson NV, Garland A. Quantile regression and restricted cubic splines are useful for exploring relationships between continuous variables. *J Clin Epidemiol*. 2009; 62(5):511-517., e1.
13. Harrell FE Jr. *Regression Modeling Strategies*. New York, NY: Springer; 2001.
14. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Development Core Team; 2011.
15. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1): 128-138.
16. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27(2):157-172.
17. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361-387.
18. Wiener RS, Schwartz LM, Woloshin S, Welch HG. Population-based risk for complications after transthoracic needle lung biopsy of a pulmonary nodule: an analysis of discharge records. *Ann Intern Med*. 2011;155(3):137-144.
19. Ravenel JG, Leue WM, Nietert PJ, Miller JV, Taylor KK, Silvestri GA. Pulmonary nodule volume: effects of reconstruction parameters on automated measurements—a phantom study. *Radiology*. 2008;247(2):400-408.
20. van Klaveren RJ, Oudkerk M, Prokop M, et al. Management of lung nodules detected by volume CT scanning. *N Engl J Med*. 2009;361(23):2221-2229.