CHEST

Original Research

LUNG CANCER

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 cited[9] 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 ≤ 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, FEV_1 % 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 model[8] 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 models[12] 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

DOI: 10.1378/chest.13-0708