Approaches to lung nodule risk assessment: clinician intuition versus prediction models

Adam H. Fox[1], Nichole T. Tanner[1,2]

[1]Division of Pulmonary, Critical Care, Allergy and Sleep Medicine, Medical University of South Carolina, Charleston, SC, USA; [2]Health Equity and Rural Outreach Innovation Center (HEROIC), Ralph H. Johnson Veterans Affairs Hospital, Charleston, SC, USA
Contributions: (I) Conception and design: All authors; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: None; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.
Correspondence to: Nichole T. Tanner, MD, MSCR. Associate Professor of Medicine, Division of Pulmonary, Critical Care, Allergy and Sleep Medicine, Medical University of South Carolina, 96 Jonathan Lucas Street, CSB Suite 816, MSC 630, Charleston, SC 29425, USA. Email: tripici@musc.edu.

Abstract: Pulmonary nodules are increasingly identified on imaging exams performed for a number of clinical presentations and can pose a diagnostic problem for clinicians. Guideline-directed management algorithms are structured on nodule pre-test probability of malignancy. The risk of malignancy can be clinician-assigned or calculated utilizing validated risk prediction calculators. Once pre-test probability of cancer is estimated, nodule management options range from a conservative approach with serial imaging to more invasive measures including biopsy procedures or surgical resection. Here we review pulmonary nodule management with a focus on methods for assigning malignancy risk and highlight novel ways currently under active research to improve nodule risk assessment and management.

Keywords: Solitary pulmonary nodule; nodule risk prediction; nodule management; lung cancer

Introduction

Pulmonary nodules pose a frequent diagnostic challenge for clinicians and have the potential to cause distress in patients (1). Prior to the advent of lung cancer screening, an estimated 1.6 million pulmonary nodules were detected annually in the United States (2). Further, data from the largest lung cancer screening trial published to date found that 25% of those undergoing screening had a screen-detected nodule (3), and recent data from a large screening project conducted within the Veterans Administration (VA) Hospital System had a nodule detection rate of 60% (4). These numbers highlight the increasing prevalence of pulmonary nodules and further identify the need for sound methods for evaluation and management. While the decision to pursue tissue diagnosis or expectant management depends on many factors, including patient preference and the risk of tissue sampling, the inherent risk of malignancy is often a driving factor.

Clinicians must weigh the risks of a surgery or procedure to diagnose a nodule against the likelihood that the nodule is benign and requires no interventions. Ideally, invasive procedures are reserved for patients with nodules at the highest risk of malignancy and avoided where risk of cancer is low. Professional society guidelines for nodule management provide some structure for clinicians and institutions to follow but appropriately allow flexibility for individual patient decision-making (5,6). While there are variations in guidelines, all hinge on risk assessment of the nodule's pre-test probability or likelihood of malignancy.

Guidelines begin by categorizing their recommendations according to nodule size, which is itself a known predictor of malignancy and included in all prediction models discussed below. As an example, small nodules (<8 mm) are less likely to be malignant, are often not suitable for sampling and are below the size limit in which positron emission tomography
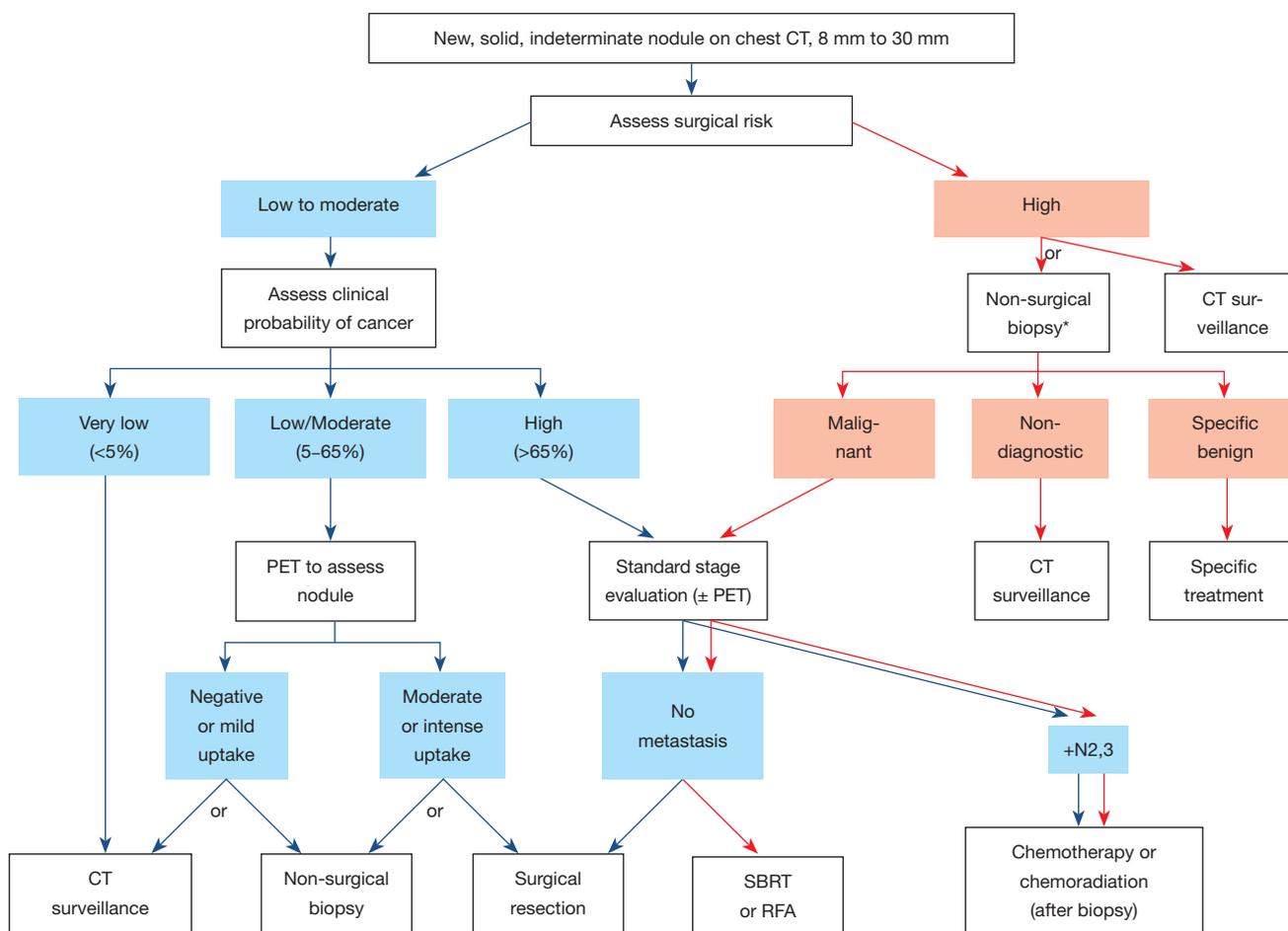
Figure 1 CHEST guidelines: management algorithm for indeterminant pulmonary nodules, 8–30 mm (5). *, among individuals at high risk for surgical complications, CHEST guidelines recommend either CT scan surveillance (when the clinical probability of malignancy is low to moderate) or nonsurgical biopsy (when the clinical probability of malignancy is moderate to high).

(PET) is useful. Management of nodules <8 mm involves a decision on timing of serial imaging. In the Fleischner Society Guidelines, for example, the length of the interval for subsequent imaging is driven by risk inherent to the patient, with the lowest-risk patients having no interval CT scan and the highest-risk patients having repeat imaging in 3 months (7). In addition to a patient's inherent risk of malignancy such as age and smoking history, there other clinical factors specific to the nodule that lower the risk of malignancy to near zero. Examples of typical benign nodules include certain calcification patterns, stability for greater than two years, resolution on imaging, or density consistent with benign processes such as a hamartoma; these nodules have a low enough risk of malignancy that serial imaging is typically not warranted (5,7-9). Nodules that lack a specific benign pattern are often labelled "indeterminant".

In larger, indeterminant nodules (≥8 mm), society guidelines recommend using clinical judgment or a validated prediction model to assess the risk of malignancy (5,9). Management algorithms are then further stratified by this assessment of risk as illustrated in *Figure 1* from the American College of Chest Physicians (CHEST) guidelines (5). This article will review different strategies in assessing risk of malignancy, including physician gestalt and risk prediction modeling.

Risk prediction models

The CHEST and the British Thoracic Society (BTS) guidelines recommend the use of a validated risk prediction model to assess nodule risk, with the BTS guidelines being more prescriptive in naming which model to use (5,9).

3298

Fox and Tanner. Approaches to lung nodule risk assessment

Table 1 Comparison of prediction model characteristics

| Model | Study type | # of study patients | Enrollment time | Study sites | Prevalence of cancer | Patient's baseline characteristics | | Model accuracy (AUC ROC) |
						Smoking status	Gender	
Swenson Model	Retrospective	629	1984–1986	1 center in the US	23%	68% ever smokers	51% male	0.83
Herder Model	Retrospective	106	1997–2001	1 center in the Netherlands	57.5%	74% ever smokers	58% male	0.92
Brock Model	Prospective, lung cancer screening trial	1,871 patients with 7,008 nodules	2008–2010	8 sites across Canada	5.5%	Only smokers	55% male	>0.9*
VA Model	Prospective	375	1999–2001	10 sites across the US	54%	93% ever smokers	97% male	0.79

*, depending on the model. VA, veterans administration.

Table 2 Variables included in prediction models

Model	Variables included in model
Swenson Model	Age, nodule diameter, smoking status, upper lobe location, and presence of spiculation
VA Model	Smoking status, age, nodule diameter, number of years since smoking cessation
Brock Model (parsimonious)	Sex, nodule size, upper lobe location and presence of spiculation
Brock Model (Full model)	Age, sex, family history of lung cancer, presence of emphysema, nodule diameter, nodule density, upper lobe predominance, number of nodules and presence of spiculation

VA, veterans administration.

Several popular and validated models include The Mayo (aka Swenson) Model, The Brock Model and the VA Model. Each of these models were developed in different populations and across different decades each with a different prevalence of malignancy. The Mayo Model was developed in a retrospective cohort of 629 patients with newly discovered nodules by chest radiograph between 1984 and 1986 at a single center (10). It was later externally validated in several retrospective cohorts (11,12). As PET imaging for nodule characterization became more frequently utilized, the Herder model was developed which added the level of PET avidity to the Mayo model, improving the model's performance (13). The Brock Model was built from a cohort of 1,090 patients from a computed tomography (CT) lung cancer screening population as part of the Pan-Canadian Early Detection of Lung Cancer Study (PanCan) across 8 Canadian health centers from 2008–2010 (14,15). While this model was developed in a screening cohort, it was later validated in non-screening cohorts (11,16). The VA Model was developed utilizing data from 375 patients across 10 VA sites as part of a prospective

study assessing the accuracy of CT compared to PET for the evaluation of lung nodules (17,18). Details of the study cohorts used to develop the models discussed here are reviewed in *Table 1*.

Prediction models offer a standardized and reproducible approach to nodule risk assessment. When considering use of a prediction model and selecting one over another for clinical use, one should determine whether the model or validation cohorts have comparable characteristics to the individual patient case in which the model is to be applied. For example, when evaluating a screen-detected nodule, utilizing the Brock model developed out of a screening cohort would be better suited than the Mayo model. In addition to the characteristics of the patients included to build each model, factors such as cohort size, a prospective nature of the study and presence of an external validation offer strength to statistical models. Model choice may also be indicated based on availability of required information. The variables used by each model differ and are outlined in *Table 2*. These include patient level factors (e.g., smoking status and age) and nodule radiographic features (e.g., size and border).

Functional imaging for nodule evaluation

PET scans are often utilized to evaluate indeterminant pulmonary nodules and commonly used for lung cancer staging. Guidelines generally suggest that PET be performed in patients with indeterminant nodules >8 mm in which the probability of malignancy is intermediate (e.g., 5–65% in Chest, ≥10% in BTS). The diagnostic utility of PET scans alone has been previously evaluated and several meta-analyses found sensitivities and specificities of 95% and 80%, respectively (19,20). PET scan evaluation may add value in malignancy prediction, as exemplified by the Herder model's inclusion of PET avidity which improved the area under the ROC curve by 13.6% (P=0.0003) compared to the Mayo model (21). However, there are limitations in the utility of PET scans in nodule evaluation. If the pre-test probability of malignancy is already high (e.g., >65%), PET is unlikely to change management and is better utilized for staging of a suspected lung cancer (e.g., to determine nodal or extrathoracic metastasis). Similarly, nodules <8 mm in diameter are below the size threshold in which PET is reliably useful. An important consideration when utilizing PET is its inability to distinguish malignancy from other inflammatory conditions such as fungal infections, tuberculosis, sarcoidosis, and other acute infections. One recent meta-analysis showed that the specificity of PET in areas with endemic granulomatous infections may drop as low at 61% (22). PET is best utilized in a patient with an indeterminant nodule >8 mm with a low risk for malignancy and geography without a significant burden of granulomatous infections.

Clinician-assessed nodule risk evaluation

Pulmonary nodule management guidelines specify that expert clinician intuition, rather than risk prediction models, is an acceptable alternative to assessing nodule pretest probability of malignancy (5,9). In one survey of over 400 practicing pulmonologists and thoracic surgeons, only 28% reported routinely using nodule prediction calculators, while the majority relied on their intuition (23). In contrast to risk prediction models, a clinician can take into account more variables, past experiences and the nuances of an individual patient's case to assess the risk of malignancy. One study compared the risk assessment of 11 physician participants against that of the Mayo and VA models and found similar performance in predicting malignancy with area under the ROC curves of 0.73, 0.8,

and 0.8 for clinical judgement, the Mayo model and the VA model, respectively (24). A later, prospective study of 337 individual risk assessments showed that physician-assessed probability of malignancy performed better than prediction models with area under the ROC curve of 0.85, 0.75 and 0.78 for physician assessment, the VA model and the Mayo model, respectively (25). Of note, risk of malignancy in this study was assessed by physicians who routinely manage a large volume of patients with nodules.

Despite evidence that clinician-assessed risk is accurate in predicting likelihood of malignancy, adherence to society guidelines among pulmonologists, radiologists and surgeons is poor. A retrospective cohort evaluating guideline concordance in 300 patients with lung nodules in 15 VA hospitals found that 44% of patients received care inconsistent with guidelines and multivariate analysis revealed inappropriate radiologist recommendation was the strongest predictor of care inconsistent with guidelines (26). One retrospective multicenter study analyzed the management of 377 patients with nodules by pulmonologists in the community setting and found significant departures from guideline-recommended practice (27). Despite a low (<5%) pretest probability of malignancy, 36 patients underwent invasive procedures, 28% undergoing biopsy and 17% having surgical resection; all 36 patients had benign disease. Another prospective, multicenter study that included pulmonologists, surgeons, and oncologists evaluating 685 cases showed that while clinician intuition was excellent there was discordance between guidelines and the next management step in 61% of cases. For instance, low risk (<5%) patients received more aggressive treatment than guidelines would suggest in 13 (52%) of cases (28). While assessing pre-test probability of malignancy is a crucial first step in nodule management, these studies highlight that important outcomes such as reduction of procedures in benign disease, timely diagnosis of malignancy, and cost-effectiveness are unlikely to be achieved unless pre-test probability of malignancy then appropriately informs the next step in management.

Future of nodule risk assessment

Despite advances over recent decades in development of risk prediction models, cross-sectional imaging, and society guidelines that aid clinicians in nodule risk assessment, the current paradigm still has limitations. In an ideal scenario, after nodule discovery, cheap and non-invasive testing would more accurately delineate those with cancer from those

without, further optimizing invasive procedures for those with malignant lesions and alleviating the need for further repeat imaging or testing in those with benign nodules. Several approaches including biomarkers and radiomics are being investigated to achieve higher diagnostic accuracy. If a reliable biomarker for lung cancer could be developed from serum, sputum, or upper airway epithelium, this test could help to rule-in or rule-out malignancy in the right patient population. A biomarker for lung cancer could resemble the use of a negative d-dimer in a patient already at low risk for pulmonary embolism making the post-test probability so low that further testing is not required (29). The PANOPTIC trial investigated the performance of a plasma proteomic biomarker integrated with a risk prediction model in patients with nodules at lower risk for malignancy (≤50%) and found sensitivity of 97% and a negative predictive value of 98% (28). Another study evaluated a genomic classifier obtained from a mainstem bronchial brushing during a bronchoscopy performed to investigate a potentially cancerous lesion (30). Combining the genomic classifier along with the results of the bronchoscopy yielded a sensitivity ~97%. Once a biomarker achieves clinical validity with good performance characteristics, it must then be shown to have clinical utility (31). This means that test results alter clinician practice and alter outcomes. For example, in the case of a rule-out biomarker, a clinical utility study would demonstrate that test results changed clinician behavior and fewer subsequent testing (e.g., biopsies and procedures) were performed as a result.

Radiomics is another area of interest for improving risk stratification of nodules. Radiomics, as it applies to lung nodule evaluation, is the use of quantifiable radiologic parameters to predict the malignant potential or future behavior of the nodule. Radiomics already plays a key role in nodule risk assessment as many radiologic features are key predictors of malignancy. As imaging modalities, resolution, data storage, and software continue to advance, the opportunity to identify new unique predictors of malignancy increases. One area of research, already being incorporated in BTS and Fleischner Society guidelines is the assessment of nodule volume (7,9). Assessing volume rather than diameter offers potential benefits in risk estimation and may improve prediction models (32). Diameter measurements do not account for the three-dimensional nature of nodules and do not account for growth in some dimensions but not others. Other variables such as nodule density, texture, contour, and other measurable characteristics on CT are also being investigated (33-35).

Conclusions

Evaluation of pulmonary nodules begins with assessment of pretest probability of malignancy followed by decisions driven to expedite treatment of early-stage lung cancers and avoid procedures in benign nodules. There is evidence that prediction models and clinician judgement perform similarly. In practice, physicians inherently use clinical judgement in nodule management decisions, whether or not they also use a prediction model. Physician gestalt is quick, intuitive, and allows for incorporation of the entire clinical picture including nuances and variables not included in prediction models. Alternatively, nodule prediction models provide physicians a definitive number derived from robust historical cohorts that is free of potential bias. In addition, the exact number can be utilized in patient discussions to aid in decision-making. Novel biomarkers, radiomics and other modalities to enhance nodule risk assessment are currently in various stages of discovery and validation but will need to be shown to have clinical utility over what is currently available.

Acknowledgments

Footnote

aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Slatore CG, Wiener RS. Pulmonary Nodules: A Small Problem for Many, Severe Distress for Some, and How to Communicate About It. Chest 2018;153:1004-15.
2. Gould MK, Tang T, Liu IL, et al. Recent Trends in the Identification of Incidental Pulmonary Nodules. Am J Respir Crit Care Med 2015;192:1208-14.
3. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395-409.
4. Kinsinger LS, Anderson C, Kim J, et al. Implementation of Lung Cancer Screening in the Veterans Health Administration. JAMA Intern Med 2017;177:399-406.
5. Gould MK, Donington J, Lynch WR, et al. Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013;143:e93S-120S.
6. Baldwin DR, Callister ME, Guideline Development G. The British Thoracic Society guidelines on the investigation and management of pulmonary nodules. Thorax 2015;70:794-8.
7. 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;284:228-43.
8. Erasmus JJ, Connolly JE, McAdams HP, et al. Solitary pulmonary nodules: Part I. Morphologic evaluation for differentiation of benign and malignant lesions. Radiographics 2000;20:43-58.
9. Callister MEJ, Baldwin DR, Akram AR, et al. British Thoracic Society guidelines for the investigation and management of pulmonary nodules: accredited by NICE.

Thorax 2015;70:ii1-ii54.
10. Swensen SJ, Silverstein MD, Ilstrup DM, et al. The Probability of Malignancy in Solitary Pulmonary Nodules: Application to Small Radiologically Indeterminate Nodules. Archives of Internal Medicine 1997;157:849-55.
11. Al-Ameri A, Malhotra P, Thygesen H, et al. Risk of malignancy in pulmonary nodules: A validation study of four prediction models. Lung Cancer 2015;89:27-30.
12. 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;63:335-41.
13. 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;128:2490-6.
14. McWilliams A, Tammemagi MC, Mayo JR, et al. Probability of cancer in pulmonary nodules detected on first screening CT. N Engl J Med 2013;369:910-9.
15. Tammemagi MC, Schmidt H, Martel S, et al. Participant selection for lung cancer screening by risk modelling (the Pan-Canadian Early Detection of Lung Cancer [PanCan] study): a single-arm, prospective study. Lancet Oncol 2017;18:1523-31.
16. Chung K, Mets OM, Gerke PK, et al. Brock malignancy risk calculator for pulmonary nodules: validation outside a lung cancer screening population. Thorax 2018;73:857-63.
17. Gould MK, Ananth L, Barnett PG, et al. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. Chest 2007;131:383-8.
18. Kymes SM, Lee K, Fletcher JW, et al. Assessing diagnostic accuracy and the clinical value of positron emission tomography imaging in patients with solitary pulmonary nodules (SNAP). Clin Trials 2006;3:31-42.
19. Gould MK, Maclean CC, Kuschner WG, et al. Accuracy of Positron Emission Tomography for Diagnosis of Pulmonary Nodules and Mass LesionsA Meta-analysis. JAMA 2001;285:914-24.
20. Cronin P, Dwamena BA, Kelly AM, et al. Solitary Pulmonary Nodules: Meta-analytic Comparison of Cross-sectional Imaging Modalities for Diagnosis of Malignancy. Radiology 2008;246:772-82.
21. Herder GJ, van Tinteren H, Golding RP, et al. Clinical Prediction Model To Characterize Pulmonary Nodules: Validation and Added Value of 18 F-Fluorodeoxyglucose Positron Emission Tomography. Chest 2005;128:2490-6.
22. Deppen SA, Blume JD, Kensinger CD, et al. Accuracy of

3302

Fox and Tanner. Approaches to lung nodule risk assessment

FDG-PET to diagnose lung cancer in areas with infectious lung disease: a meta-analysis. JAMA 2014;312:1227-36.

23. Tanner NT, Brasher PB, Jett J, et al. Effect of a Rule-in Biomarker Test on Pulmonary Nodule Management: A Survey of Pulmonologists and Thoracic Surgeons. Clin Lung Cancer 2020;21:e89-98.

24. Gould MK, Simkovich S, Mestaz PJ, et al. editor. Predicting The Probability Of Malignancy In Patients With Pulmonary Nodules: Comparison Of Clinical Judgment With Two Validated Models. American Thoracic Society, 2012.

25. Tanner NT, Porter A, Gould MK, et al. Physician Assessment of Pretest Probability of Malignancy and Adherence With Guidelines for Pulmonary Nodule Evaluation. Chest 2017;152:263-70.

26. Wiener RS, Gould MK, Slatore CG, et al. Resource use and guideline concordance in evaluation of pulmonary nodules for cancer: too much and too little care. JAMA Intern Med 2014;174:871-80.

27. Tanner NT, Aggarwal J, Gould MK, et al. Management of Pulmonary Nodules by Community Pulmonologists: A Multicenter Observational Study. Chest 2015;148:1405-14.

28. Silvestri GA, Tanner NT, Kearney P, et al. Assessment of Plasma Proteomics Biomarker's Ability to Distinguish Benign From Malignant Lung Nodules: Results of the PANOPTIC (Pulmonary Nodule Plasma Proteomic Classifier) Trial. Chest 2018;154:491-500.

29. Jaconelli T, Eragat M, Crane S. Can an age-adjusted D-dimer level be adopted in managing venous thromboembolism in the emergency department? A retrospective cohort study. Eur J Emerg Med 2018;25:288-94.

30. Silvestri GA, Vachani A, Whitney D, et al. A Bronchial Genomic Classifier for the Diagnostic Evaluation of Lung Cancer. N Engl J Med 2015;373:243-51.

31. Mazzone PJ, Sears CR, Arenberg DA, et al. Evaluating Molecular Biomarkers for the Early Detection of Lung Cancer: When Is a Biomarker Ready for Clinical Use? An Official American Thoracic Society Policy Statement. Am J Respir Crit Care Med 2017;196:e15-e29.

32. Mehta HJ, Ravenel JG, Shaftman SR, et al. The utility of nodule volume in the context of malignancy prediction for small pulmonary nodules. Chest 2014;145:464-72.

33. Lee SH, Cho HH, Lee HY, et al. Clinical impact of variability on CT radiomics and suggestions for suitable feature selection: a focus on lung cancer. Cancer Imaging 2019;19:54.

34. Xu Y, Lu L, E LN, et al. Application of Radiomics in Predicting the Malignancy of Pulmonary Nodules in Different Sizes. AJR Am J Roentgenol 2019;213:1213-20.

35. Wilson R, Devaraj A. Radiomics of pulmonary nodules and lung cancer. Transl Lung Cancer Res 2017;6:86-91.