Evaluating the efficacy of bronchoscopy for the diagnosis of early stage lung cancer

David M. DiBardino[1], Anil Vachani[1], Lonny Yarmus[2]

[1]Section of Interventional Pulmonology, Division of Pulmonary, Allergy and Critical Care Medicine, University of Pennsylvania, Philadelphia, PA, USA; [2]Section of Interventional Pulmonology, Division of Pulmonary and Critical Care Medicine, Johns Hopkins University, Baltimore, MD, USA

Contributions: (I) Conception and design: L Yarmus, DM DiBardino; (II) Administrative support: DM DiBardino; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: DM DiBardino; (V) Data analysis and interpretation: DM DiBardino; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: David DiBardino, MD. 3400 Spruce Street, HUP Gates 821, Philadelphia, PA 19104, USA.
Email: david.dibardino@pennmedicine.upenn.edu.

Abstract: Novel diagnostic techniques for lung cancer are rapidly evolving. Specifically, several novel changes to bronchoscopy are reaching clinical evaluation. It is critical to think about historical standards for evaluating new diagnostic testing, and put those concepts into the framework of lung cancer. Often a thorough evaluation of new technology is not performed as a part of regulatory marketing clearance. Therefore, we must consider how to best study novel testing beyond these regulatory minimums. There are several methodological principles that can achieve this goal such as using a control arm, more thorough reporting of enrolled patients, consecutive patient enrollment, and adequate sample size. We hope clinicians, particularly those performing bronchoscopy for lung nodules, will feel empowered to critically appraise the evaluation of new diagnostic testing for lung cancer moving forward.

Keywords: Lung cancer; bronchoscopy; lung nodule

Introduction

New and exciting modalities now exist for the diagnosis of lung cancer in the form of biomarkers and new biopsy devices (1,2). Much of this technology is aimed at the evaluation of indeterminate pulmonary nodules in order to establish a definitive diagnosis of early stage lung cancer or an alternative etiology. As highlighted below, a full evaluation of the clinical efficacy of these medical advances is not necessarily part of the Federal Drug Administration's (FDA) regulatory approval process. Rather, more than 90% of the devices seeking FDA approval are utilizing the 510(k) pathway which emphasizes a device's similarities with a legally marketed technology. Therefore, products may come to market before a rigorous and statistically sound evaluation has been performed (3).

The number of small peripherally located indeterminate nodules, identified either incidentally or from the expansion of lung cancer screening, is very likely to increase (4). As such, more patients are identified as candidates for this innovative testing during their diagnostic evaluation. Now is an important time to consider the clinical evaluation of such novel technology and to improve the methodologies used to assess bronchoscopic outcomes for the diagnosis of lung cancer.

Adapting traditional standards to lung cancer

The concept of sensitivity, specificity, likelihood ratio positive, likelihood ratio negative, post-test probability, and receiver operating curves (AUC) have long been described in laboratory testing (5). It is important to think about the application of these standard test characteristics to novel biopsy modalities and biomarkers for lung cancer

Table 1 Traditional measurements used to evaluate diagnostic tests

Measurement	Definition	Application to lung cancer diagnosis	Practical use
Sensitivity	True positives/(true positives + false negatives)	How many lung cancer diagnoses were confirmed	Requires accurate false negative rate
Specificity	True negatives/(true negatives and false positives)	Cases of benign lung nodules that had a diagnostic test suggesting cancer	Very rarely important in histopathology-based testing
Likelihood ratio positive	Sensitivity/(1-specificity)	Odds that someone has lung cancer after positive test	Infinite for positive pathology
Likelihood ratio negative	(1-sensitivity)/specificity	Odds that someone has lung cancer after negative test	Requires accurate false negative rate
Post-test probability	PPV or NPV depending on test outcome	The probability someone has lung cancer given a certain test result	Near 100% PPV for histopathology, NPV requires accurate sensitivity
Receiver operating curve	Plot of true positives against false positives	May have use in biomarker study where a threshold value is being calculated	No relevance to a bronchoscopy study
Diagnostic yield or accuracy	Variable	Often used in bronchoscopy studies as (true positives + true negatives)/total N	Requires accurate true negative rate

PPV, positive predictive value; NPV, negative predictive value.

(*Table 1*). For example, when considering the sensitivity and specificity for lung cancer we are not truly evaluating "diagnostic yield" or "diagnostic accuracy". Rather, we are *a priori* evaluating for the diagnosis of lung cancer alone and need to predetermine our definition of a true negative and false negative. We can largely accept a false positive rate near 0% and a true positive rate near 100% as cytology and pathology are rarely misclassified as malignant when diagnosing lung cancer (6). Diagnostic yield is a separately valuable endpoint but requires careful consideration as to what histopathologic findings constitute an accurate diagnosis of benign disease rather than a false negative in a lung cancer patient.

With this in mind, we must scrutinize the implication of a negative biopsy result (i.e., no evidence of malignancy) obtained by bronchoscopy and the appropriate approach needed to accurately estimate diagnostic performance in this setting. Given the less specific nature of benign cytology and pathology, a negative lung biopsy is not often specific for a benign alternative diagnosis to lung cancer (7,8). Further complicating matters, granulomatous inflammation can co-exist with cancer and further confuse clinicians deciding on the diagnosis of lung cancer (9-11). Therefore, long term clinical follow-up is essential in deciding which patients had a truly negative diagnostic test for lung cancer. A false negative can be confirmed by further invasive testing such as a subsequent lung biopsy, newly metastatic disease seen on follow-up imaging, or interval worsening of a lung

nodule then empirically treated as lung cancer. This type of clinical follow-up may be rigorous, but can realistically be obtained in a clinical trial or observational setting with medical record review extending 6 to 12 months beyond the biopsy. A true negative is even more challenging to confirm. Similar to false negatives, medical record review is necessary, but will likely take up to 1–2 years to confirm. Confirmation may include stability of a lesion on serial imaging, further invasive testing re-affirming benign pathology, or resolution of the abnormality on subsequent follow-up imaging. As inflammatory conditions can evolve over time, there may be nuance to the clinical record review that requires comparing final diagnoses between multiple investigators to validate the evaluation.

Despite the uncertainty created by non-malignant histopathologic findings on biopsy, the use of longitudinal assessment is essential to determine the test characteristics of novel technologies as diagnostic tools for lung cancer.

Moving beyond standard study outcomes measures and regulatory minimums

Traditional study designs designed to describe traditional characteristics of test performance may neglect the opportunity to conduct more patient-centered research (12). There are other potential endpoints that are both clinically relevant and meaningful to patients being evaluated for possible lung cancer. Novel testing should be conducted

Table 2 Common pitfalls and solutions with studies aimed at diagnosing lung cancer

Common study design	Pitfall	Solution
Single-arm study with new device	No clear comparison arm to judge new device's efficacy	Parallel trial design with a control arm
No clear power calculation	Unclear if the study can statistically fulfill the aim	Consideration of the study goals and pre-emptive power calculations
Highly selecting patients for novel diagnostic test	Lack of generalizability	Offer trial enrollment to consecutive patients being worked up for lung cancer
Expert centers only	Lack of generalizability	Multi-center design
Limited demographic and descriptive reporting of biopsy procedure	Lack of generalizability	Careful reporting of lung cancer prevalence in the study population and detailed reporting of nodule characteristics
Lack of confirmation for true negative biopsies	Cannot calculate sensitivity for lung cancer for a technology	Adequate clinical follow-up for all non-malignant biopsies

in a controlled setting where the morbidity required to perform the test can be carefully recorded. This will commonly include procedural complications, but we should also consider the issue of time. The length of time required to plan and execute a novel diagnostic test, and how long it takes to obtain a final diagnosis of lung cancer are often hugely important to patients. Furthermore, it is plausible that minimizing the time from diagnosis to treatment of lung cancer can improve outcomes although this association has not been consistently seen in the literature (13). Ultimately, an association between a novel diagnostic test and an earlier stage of lung cancer at the time of treatment initiation could have prognostic implications.

It is worth mentioning that diagnostic testing has not traditionally been held to this standard. In fact, outcomes indirect to clinical care and surrogate outcomes are commonly used (14). As it applies to biopsy devices to diagnose lung cancer, U.S. FDA marketing clearance is often obtained via section 510(k) of the *Food*, *Drug*, *and Cosmetic Act* whereas a device gains marketing approval after demonstrating equivalent safety and efficacy as a legally marketed similar device (15). Therefore, it becomes difficult to imagine the motivation for the commercial developers of these technologies to fund costly, time-consuming prospective controlled studies to generate high quality evidence on diagnostic performance and clinical efficacy when it is unnecessary by regulatory standards. Furthermore, this pathway does appropriately hasten the availability of new technology for patients suffering from serious disease states such as lung cancer. This potential disconnect between regulatory minimums, clinically

important outcomes, and encouraging medical advances has recently come under scrutiny outside of the lung cancer space (16). Perhaps safety concerns are less relevant for new devices aimed at diagnosing lung cancer, but the motivations presented by 510(k) clearance will clearly affect new developments in bronchoscopy as all recent innovations in this technology have utilized this pathway (17).

Study designs

Novel diagnostic tests are often studied in a prospective single-arm, descriptive design on account of the regulatory issues aforementioned, cost, and simplicity (*Table 2*). This study design is becoming less and less useful or accepted as more options to diagnose lung cancer are developed. The tradition in lung cancer diagnosis has been single-armed studies of transthoracic or bronchoscopic biopsy procedures aimed at evaluating for the traditional test characteristics of sensitivity and diagnostic yield (6,8,18-25). This study design must be interpreted with caution and is not the ideal format for studying novel diagnostic testing for lung cancer (26). These descriptive studies generally report test characteristics that are specific to the exact population studied and can allow for an informal comparison to historical controls of diagnostic rates. Single-arm studies also cannot be used to compare a novel test to an established diagnostic approach as there is no ability to avoid bias or control for important confounders that can influence diagnostic yields.

The main concern in this setting is selection bias (27). Many single-arm studies highly select patients that

are "ideal" for a diagnostic test rather than enrolling consecutive, un-selected patients. This is also in contrast to multi-center randomized controlled trials where confounding can at least be equally distributed between study arms and institutions, and data should be rigorously collected (28). Single-arm studies have the potential to produce a favorable result if patients are consciously or unconsciously selected or reported based on factors that lead to a successful procedure or incomplete data collection. While this can be measured to some degree (i.e., nodule size, lobe, bronchus sign, systematic data collection, etc.), there are a number of factors relating to nodule characteristics, including size, density, and orientation, that can influence the ability to achieve a high-quality biopsy and many barriers to complete data collection. No prior studies have collected imaging tests and attempted to report a more sophisticated analysis of factors that may have influenced diagnostic yield. As an example, a large-scale meta-analysis of all guided bronchoscopy techniques have aggregated data from such single-arm studies describing diagnostic yield (29). Most patients in these studies had lung cancer and a diagnostic yield of roughly 70% was seen regardless of exact technique used. Now contrast that to a large registry study describing less highly selected patients undergoing guided bronchoscopy with a yield of 54%, or a prospective randomized clinical trial of bronchoscopy where both arms have a diagnostic yield of less than 50% (8,30). In fact, this discrepancy was further validated recently by an update of the meta-analysis discussed above which showed a decrease in diagnostic yield of almost 10% over the past decade despite advances in technologic development (31). At least some of this effect is attributed to the field maturing in its definition of nonmalignant disease. Requiring stricter criteria to distinguish between true- and false-negative bronchoscopy results invalidated many of the earlier claims of diagnostic yield.

Ideally new biopsy technologies will be evaluated using rigorous comparative designs that incorporate a relevant control group. Some randomized, experimental studies have in fact compared different bronchoscopy techniques for the diagnosis of lung cancer (30,32,33). A recent study was the first to directly compare three existing advanced bronchoscopic techniques within a cadaver model. This study directly compared the use of radial probe endobronchial ultrasound (r-EBUS) with an ultrathin scope to electromagnetic navigational bronchoscopy (EMNB) to robotic bronchoscopy in a randomized fashion in human cadavers. The study showed superiority of the robotic approach over either r-EBUS or EMNB (34). The advantages of this trial design are growing as multiple navigational bronchoscopy platforms are now available. The ideal control arm represents a widely available standard of care that clinicians are aiming to improve on. Specifically, when evaluating novel bronchoscopic techniques the ideal control arm varies depending on a clinician's available options. Unfortunately, there remains equipoise in the guided bronchoscopy literature, and there is no widely agreed upon standard of care. The widely available technology reported in most registry studies and numerous single-arm descriptive trials is fluoroscopy, EMNB, virtual bronchoscopy (VB), and r-EBUS (8,18-22,24,25,29,30). Making things more complicated, these modalities can be combined to improve success (23,32,35). The exciting potential of guided bronchoscopy was reflected in the 2013 the American College of Chest Physicians guideline statement recommending the use of r-EBUS or EMNB to diagnose pulmonary nodules when available (36). This background creates several quality options for control arms. For example, an institution currently using an EMNB platform will be most interested to know how novel devices compare to EMNB. Furthermore, they can internally collect data on the diagnostic performance for consecutive patients at their center and compare this number to the control arm of this proposed trial. This, along with the demographics table, will give the reader a sense of how generalizable the study is to their practice. Generally speaking, r-EBUS is the most widely available and well-studied technique that presents itself as an attractive control arm. Another desirable aspect of r-EBUS is the ability to add this intervention to the experimental arm (i.e., the novel device arm) in order to integrate the limited comparative data supporting combining guided modalities (23,32).

Bronchoscopy trials also need to account for proficiency bias and maximize generalizability. Single-center studies involving well known procedure experts may not be generalizable to the majority of physicians charged with diagnosing lung cancer. Utilizing a multi-centered design and trying to include diverse practice settings may be of benefit (37). The first large multicentered clinical trial investigating a clinical algorithmic approach (ALL IN ONE trial) in combination with advanced bronchoscopic technologies is underway with anticipated completion in 2019. Final results are expected in 2020 after a robust clinical and radiographic follow-up period with pathologic confirmation to ensure confidence in diagnoses labeled as benign (38).

Reporting of the study population and intervention

The importance of the patient population enrolled in a study designed to evaluate lung cancer cannot be overstated. Physicians struggle to understand when diagnostic tests apply to their clinical practice and must frame all medical testing for lung cancer with this in mind. How exactly patients were chosen for enrollment in a biopsy trial allows for insight into generalizability. Other than well-defined inclusion and exclusion criteria, were these consecutive patients undergoing lung nodule evaluation who needed further investigation? If not consecutive, how did the investigators decide to offer enrollment? If patients declined enrollment, was there any systematic reason that could influence diagnostic yield?

Investigators should clearly state the patient population successfully enrolled with a thorough demographics table. Specifically, for a bronchoscopy device all known factors associated with an increased diagnostic yield for guided bronchoscopy should be reported for each study arm (when applicable) such as nodule size, location of the nodule, distance from pleura, and presence of a bronchus sign on compute tomography (CT) scan. Less well understood factors may also be useful in an effort to communicate how difficult nodules are to biopsy with traditional bronchoscopy such as apical-medial location in the upper lobes, and apical position in the superior segment of the lower lobes. It may also be useful to know what proportion of patients were also eligible for another diagnostic test such as a transthoracic needle biopsy or surgical lung biopsy. This could lead to hypothesis generating subgroup analyses for patients that may or may not have had other options.

Detailed reporting of the biopsy methods is required if the novel diagnostic test is a bronchoscopy procedure. Combining traditional sampling methods such as cytology brushes, fine-needle aspiration, transbronchial biopsies, and bronchial lavage will maximize the sensitivity for lung cancer without adding futile maneuvers to the procedure. For studies involving r-EBUS, reporting the r-EBUS view obtained during the procedure (eccentric versus concentric versus no view) can inform proceduralists about the expected diagnostic yield and lead to informative subgroup analyses (19). Other patient safety related outcomes should be reported such as procedure time, pneumothorax, and clinically meaningful bleeding.

Sample size

A statistically sound method should be used to determine the sample size needed to evaluate a novel diagnostic test for lung cancer. For studies with a control arm and traditional experimental design this power calculation will be more straight forward. As above, the sensitivity for lung cancer or diagnostic yield can be compared as proportions using a Chi-square test. The sensitivity for lung cancer or diagnostic yield will need to be assumed depending on the investigators' interpretation of existing data. For example, if registry and clinical trial data is used to estimate the success in the control arm, one may choose an expected diagnostic yield of 50%. The novel technology's diagnostic yield can be estimated based on pilot data while ensuring a clear clinical improvement on this baseline rate of 50%. One may consider a diagnostic yield of 75% as a clinically significant goal for novel technology as it starts to approach the data supporting transthoracic needle biopsy in expert hands (6). Assuming a power of 80% to detect a difference and a standard 5% alpha error, this study would require 116 subjects. This enrollment goal will change dramatically if different assumptions are made. This number would rise to 322 if instead a novel technology aimed to improve on an assumed 70% diagnostic yield, as seen in prior single-arm descriptive studies, to an 85% diagnostic yield, with a 90% power to detect a difference.

Unfortunately, this type of power calculation is complicated for single-arm studies and can be done using any number of statistical philosophies (39,40). Ultimately these evaluations could use similar principles as controlled trials if a baseline diagnostic yield or sensitivity for lung cancer is assumed based on prior data. This inevitably requires the problematic assumption that the patient population undergoing the novel diagnostic test exactly matches the patient population previously studied.

The sample size required to assess safety has to do with the accuracy of an estimated complication rate. An adequate subject number should be chosen in order to narrow the confidence intervals around estimates based on previous data for pneumothorax and bleeding. For example, if an estimated pneumothorax rate is 2%, a study with 100 subjects would estimate that event with a fairly wide 95% confidence interval of 0.24% to 7%. A study with 300 subjects will have a much narrower 95% confidence interval of 0.74% to 4.3%.

Conclusions

There are obviously many competing interests when evaluating novel technology to diagnose lung cancer. Clinicians, patients consenting to testing, and our industry partners need to find common ground on what pragmatic study designs will help us all achieve our goals. Evaluation of these advancements may take place in steps. Often marketing approval will be obtained by the FDA before the technology's clinical efficacy evaluation is fully known. In the post-marketing phase, fairly large studies involving diverse practice settings and a carefully considered control arm will go a long way towards understanding the use of novel testing. These results should be carefully reported, include patient centered outcomes when possible, and focus on clinical follow-up to confirm true and false negative testing.

Thorough evaluation of novel diagnostic testing will set the stage for many future possibilities and guide continued innovation. By optimizing diagnostic testing we can further refine lung cancer screening algorithms, better understand the potential for endoscopic ablation of lung cancer, and mitigate the uncertainty patients face when diagnosed with an indeterminate pulmonary nodule.

Acknowledgments

Footnote

Provenance and Peer Review: This article was commissioned by the Guest Editors (Fabien Maldonado and Robert Lentz) for the series "Novel Diagnostic Techniques for Lung Cancer" published in *Journal of Thoracic Disease*. The article was sent for external peer review organized by the Guest Editors and the editorial office.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi. org/10.21037/jtd.2020.02.35). The series "Novel Diagnostic Techniques for Lung Cancer" was commissioned by the editorial office without any funding or sponsorship. AV reports personal fees from Johnson and Johnson, grants from Bronchus Medical, Inc., grants from MagArray, Inc., outside the submitted work. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all 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.

References

1. Tanoue LT, Tanner NT, Gould MK, et al. Lung cancer screening. Am J Respir Crit Care Med 2015;191:19-33.
2. Shafiq M, Lee H, Yarmus L, Feller-Kopman D. Recent Advances in Interventional Pulmonology. Ann Am Thorac Soc 2019;16:786-96.
3. Rathi VK, Ross JS. Modernizing the FDA's 510(k) Pathway. N Engl J Med 2019;381:1891-3.
4. Qian F, Yang W, Chen Q, et al. Screening for early stage lung cancer and its correlation with lung nodule detection. J Thorac Dis 2018;10:S846-59.
5. Fardy JM. Evaluation of diagnostic tests. Methods Mol Biol 2009;473:127-36.
6. DiBardino DM, Yarmus LB, Semaan RW. Transthoracic needle biopsy of the lung. J Thorac Dis 2015;7:S304-16.
7. Baaklini WA, Reinoso MA, Gorin AB, et al. Diagnostic yield of fiberoptic bronchoscopy in evaluating solitary pulmonary nodules. Chest 2000;117:1049-54.
8. Ost DE, Ernst A, Lei X, et al. Diagnostic yield and complications of bronchoscopy for peripheral lung lesions. results of the aquire registry. Am J Respir Crit Care Med 2016;193:68-77.
9. Dagaonkar RS, Choong CV, Asmat AB, et al. Significance of coexistent granulomatous inflammation and lung cancer. J Clin Pathol 2017;70:337-41.
10. Khurana KK, Stanley MW, Powers CN, et al. Aspiration cytology of malignant neoplasms associated with granulomas and granuloma-like features: diagnostic dilemmas. Cancer 1998;84:84-91.
11. Gilbert CR, Abendroth C, Yarmus LB. The intranodal presence of coexisting granulomatous inflammation and carcinoma during transbronchial needle aspiration of intrathoracic lymphadenopathy. J Bronchology Interv

Pulmonol 2017;24:80-3.

12. Sacristán JA. Patient-centered medicine and patient-oriented research: improving health outcomes for individual patients. BMC Med Inform Decis Mak 2013;13:6.

13. Olsson JK, Schultz EM, Gould MK. Timeliness of care in patients with lung cancer: a systematic review. Thorax 2009;64:749-56.

14. Singh S, Chang SM, Matchar DB, et al. Chapter 7: grading a body of evidence on diagnostic tests. J Gen Intern Med 2012;27 Suppl 1:S47-55.

15. Federal Food, Drug, and Cosmetic Act, 21 § 9 [Internet]. 2012 [cited 2019 Oct 7]. Available online: https://uscode.house.gov/view.xhtml;jsessionid=30F7CC5B6000C257E67F0F879CEBCD8A?req=granuleid%3AUSC-prelim-title21-chapter9&saved=%7CZ3JhbnVsZWlkOlVTQy1wcmVsaW0tdGl0bGUyMS1zZWN0aW9uMzQy%7C%7C%7C0%7Cfalse%7Cprelim&edition=prelim

16. Pellerin C, Panchbhavi V, Janney CF. An analysis of foot and ankle device recalls by the food and drug administration. Cureus 2018;10:e3123.

17. FDA 510(k) Premarket Notification [Internet]. [cited 2019 Oct 7]. Available online: https://www.accessdata.fda.gov/scrIpts/cdrh/cfdocs/cfpmn/pmn.cfm

18. Hsia DW, Jensen KW, Curran-Everett D, et al. Diagnosis of lung nodules with peripheral/radial endobronchial ultrasound-guided transbronchial biopsy. J Bronchology Interv Pulmonol 2012;19:5-11.

19. Chen A, Chenna P, Loiselle A, et al. Radial probe endobronchial ultrasound for peripheral pulmonary lesions. A 5-year institutional experience. Ann Am Thorac Soc 2014;11:578-82.

20. Ali MS, Trick W, Mba BI, et al. Radial endobronchial ultrasound for the diagnosis of peripheral pulmonary lesions: A systematic review and meta-analysis. Respirology 2017;22:443-53.

21. Khandhar SJ, Bowling MR, Flandes J, et al. Electromagnetic navigation bronchoscopy to access lung lesions in 1,000 subjects: first results of the prospective, multicenter NAVIGATE study. BMC Pulm Med 2017;17:59.

22. Eberhardt R, Anantham D, Herth F, et al. Electromagnetic navigation diagnostic bronchoscopy in peripheral lung lesions. Chest 2007;131:1800-5.

23. Chee A, Stather DR, Maceachern P, et al. Diagnostic utility of peripheral endobronchial ultrasound with electromagnetic navigation bronchoscopy in peripheral lung nodules. Respirology 2013;18:784-9.

24. Gex G, Pralong JA, Combescure C, et al. Diagnostic yield and safety of electromagnetic navigation bronchoscopy for lung nodules: a systematic review and meta-analysis. Respiration 2014;87:165-76.

25. Asano F, Eberhardt R, Herth FJF. Virtual bronchoscopic navigation for peripheral pulmonary lesions. Respiration 2014;88:430-40.

26. Manja V, Lakshminrusimha S. Epidemiology and clinical research design, part 1: study types. Neoreviews 2014;15:e558-69.

27. Grayling MJ, Mander AP. Do single-arm trials have a role in drug development plans incorporating randomised trials? Pharm Stat 2016;15:143-51.

28. Butcher NJ, Monsour A, Mew EJ, et al. Improving outcome reporting in clinical trial reports and protocols: study protocol for the Instrument for reporting Planned Endpoints in Clinical Trials (InsPECT). Trials 2019;20:161.

29. Wang Memoli JS, Nietert PJ, Silvestri GA. Meta-analysis of guided bronchoscopy for the evaluation of the pulmonary nodule. Chest 2012;142:385-93.

30. Tanner NT, Yarmus L, Chen A, et al. Standard bronchoscopy with fluoroscopy vs thin bronchoscopy and radial endobronchial ultrasound for biopsy of pulmonary lesions: A multicenter, prospective, randomized trial. Chest 2018;154:1035-43.

31. Thomas N, Tanner N, Nietert P, et al. Updated meta-analysis of guided bronchoscopy for the evaluation of pulmonary lesions. Chest 2019;156:A1702.

32. Eberhardt R, Anantham D, Ernst A, et al. Multimodality bronchoscopic diagnosis of peripheral lung lesions: a randomized controlled trial. Am J Respir Crit Care Med 2007;176:36-41.

33. Oki M, Saka H, Asano F, et al. Use of an Ultrathin vs Thin Bronchoscope for Peripheral Pulmonary Lesions: A Randomized Trial. Chest 2019;156:954-64.

34. Yarmus L, Akulian J, Wahidi M, et al. A Prospective Randomized Comparative Study of Three Guided Bronchoscopic Approaches for Investigating Pulmonary Nodules: The PRECISION-1 Study. Chest 2020;157:694-701.

35. Yarmus LB, Arias S, Feller-Kopman D, et al. Electromagnetic navigation transthoracic needle aspiration for the diagnosis of pulmonary nodules: a safety and feasibility pilot study. J Thorac Dis 2016;8:186-94.

36. Rivera MP, Mehta AC, Wahidi MM. Establishing the diagnosis of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013;143:e142S-65S.

37. Sprague S, Matta JM, Bhandari M, et al. Multicenter collaboration in observational research: improving generalizability and efficiency. J Bone Joint Surg Am 2009;91 Suppl 3:80-6.

38. Thiboutot J, Lee HJ, Silvestri GA, et al. Study design and rationale: A multicenter, prospective trial of electromagnetic bronchoscopic and electromagnetic transthoracic navigational approaches for the biopsy of peripheral pulmonary nodules (ALL IN ONE trial). Contemp Clin Trials 2018;71:88-95.

39. Bujang MA, Adnan TH. Requirements for minimum sample size for sensitivity and specificity analysis. J Clin Diagn Res 2016;10:YE01-YE06.

40. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform 2014;48:193-204.