# Learning Curves in Electromagnetic Navigational Bronchoscopy: What Do They Tell Us?

**Darlene R. Nelson, M.D., M.H.P.E.[1], Andrea Leep Hunderfund, M.D., M.H.P.E.[2], and John J. Mullon, M.D.[1]**

[1]Division of Pulmonary and Critical Care and [2]Department of Neurology, Mayo Clinic, Rochester Minnesota

The field of interventional pulmonology (IP) has grown rapidly over the past two decades (1) as the number and complexity of procedures for treating patients with pulmonary disease have increased. This innovation brings with it the challenge of teaching new procedures and ensuring the competency of physicians using new techniques. Recent studies demonstrate variable rates of skill acquisition among learners and emphasize volume alone should not determine competency (2–4). Learning curves can help inform competency assessment by illustrating the rate of learner skill acquisition. Understanding learning curves for novices for various procedures is thus essential for program directors as they train the next generation of pulmonologists.

In this issue of *ATS Scholar*, Lee and colleagues describe a multicenter prospective evaluation of IP fellow learning curves for electromagnetic navigational bronchoscopy (ENB) (5). The first 20 consecutive ENB procedures of 26 fellows at 16 academic institutions in clinical settings were observed and scored by two independent faculty raters (via direct observation) and two blinded faculty raters (via asynchronous video review). Participating IP fellows had completed 3 years of pulmonary and critical care fellowship training and performed more than 150 flexible bronchoscopies before starting their IP fellowship. Raters used an assessment tool designed for use by faculty to measure ENB skill acquisition among learners in a simulation environment (6). This tool has published validity evidence supporting its use for this purpose and includes four domains: *1*) procedural planning; *2*) equipment set up and registration; *3*) navigation to the target lesion; and *4*) biopsy performance. Fellows who achieved a score greater than 12 (with a minimum score of 3 in each domain) on three consecutive procedures were considered competent (6). The authors generated learning curves for each fellow and organized the resulting curves into quartiles based on the number of procedures required to meet the predefined competency threshold. The rate of skill acquisition varied significantly between groups, with the first quartile achieving competence after a median of 2 procedures (range 1–2) and the fourth quartile achieving competence after a



ATS Scholar Vol 3, Iss 2, pp 171–174, 2022

DOI: 10.34197/ats-scholar.2022-0046ED

median of 9 procedures (range 6–15). Procedural complications occurred in 6 of 184 procedures (3.8%), with 4 occurring before competency attainment and 2 after competency attainment. Baseline characteristics (handedness, year of training, sex, familiarity with ENB) did not predict which fellows would fall into which quartile, and all fellows ultimately achieved the predefined competency standard. The authors conclude competency in ENB can be achieved among IP fellows after 15 cases and suggest learning curves can help educators recognize struggling or advanced learners.

Two additional aspects of this study deserve mention. First, scores generated by in-person and blinded faculty raters were highly correlated. This finding suggests blinded ratings of video recordings could be used adjunctively for procedural skill assessment, as is already being done in Europe (7). This approach could help eliminate sources of error such as rater bias due to prior knowledge or experience with the learner, although it also reduces opportunities for real-time teaching and feedback. Second, the learning curves of participating fellows showed an initial upward slope (reflecting ENB skill acquisition) but did not demonstrate a second slope (which is often associated with increasing mastery). The lack of a second slope suggests attainment of mastery may require significantly more than 20 procedures. Alternatively, there may be a potential ceiling effect to the current assessment tool.

This ambitious study adds to a growing body of literature regarding procedural learning curves in the field of IP (2, 4, 6, 8). The authors discuss numerous key issues regarding the use of learning curves in procedural training, including wide variation among trainees and the need for robust assessment tools to help identify struggling learners as well as learners who are ready to advance to the next skill. We share the authors' interest in using learning curves for this purpose and would like to highlight additional important considerations surrounding their use.

The validity of a learning curve refers to the degree that the result is justifiable and meaningful and depends on the validity of scores used to generate the curve (9, 10). Evidence for validity should be gathered from multiple sources, and the authors appropriately used an assessment tool with previously published evidence supporting the use of resulting scores to measure ENB skill acquisition (6). Using learning curves to classify learners and determine when competency standards are met can have important downstream consequences for learners, faculty, programs, and patients. These consequences comprise a source of validity evidence that has not (to our knowledge) been reported for scores derived from this assessment tool. For example, if a fellow is inappropriately placed in a low-performing group, unneeded instructor and trainee time may be spent. Conversely, if a trainee is misclassified in a high-performing group, procedural complications and patient risk may increase (11). These and other potential consequences of testing would be fruitful areas for future inquiry.

How learner performance is measured has implications for consequences of testing and the validity of resulting learning curves. Different approaches can lead to different conclusions about the typical number of procedures needed to achieve competence. For example, Lee and colleagues found 15 ENB procedures were required for all fellows to achieve a competency standard based on checklist scores, whereas Toennesen and colleagues found an average of 40–50 ENB procedures were required for experienced bronchoscopists to achieve a competency standard based on diagnostic accuracy (12). The difference between the learning curves in these studies can be explained by differences in how performance was assessed (checklist vs. clinical outcome).

The standard used to define competence also has important implications for consequences of testing. The process by which standards are set thus deserves careful consideration. Various approaches to standard setting exist. The Angoff method, for example, involves subject experts making judgements about the difficulty of each item (as was done for the assessment tool used in this study) (13). However, prior studies have shown that faculty often make holistic judgments about learner competence that differ from those derived from checklist scores (14). Other approaches include the contrasting groups method, which identifies a pass/fail cut score where two groups (expert/novice) intersect and the borderline group method in which judges identify how a borderline candidate would perform and use that as the cut score (13). Each approach has merits and shortcomings. Given the consequences arising from competency standards, educators should be explicit about their approach and provide a defensible rationale informed by the purpose and stakes of the assessment.

Competency-based procedural education requires longitudinal assessments over time that incorporate (or predict) clinically important outcomes. Lee and colleagues have made an important contribution to this effort. We look forward to additional research exploring potential consequences of testing and how these can be optimized for learners, faculty, programs, and patients.

**Author disclosures are available with the text of this article at www.atsjournals.org.**

## REFERENCES

1. Aslam W, Lee HJ, Lamb CR. Standardizing education in interventional pulmonology in the midst of technological change. *J Thorac Dis* 2020;12:3331–3340.

2. Stather DR, Chee A, MacEachern P, Dumoulin E, Hergott CA, Gelberg J, *et al.* Endobronchial ultrasound learning curve in interventional pulmonary fellows. *Respirology* 2015;20:333–339.

3. Voduc N, Adamson R, Kashgari A, Fenton M, Porhownick N, Wojnar M, *et al.* Development of learning curves for bronchoscopy: results of a multicenter study of pulmonary trainees. *Chest* 2020;158:2485–2492.

4. Wahidi MM, Hulett C, Pastis N, Shepherd RW, Shofer SL, Mahmood K, *et al.* Learning experience of linear endobronchial ultrasound among pulmonary trainees. *Chest* 2014;145:574–578.

5. Lee HJ, Argento AC, Batra H, Benzaquen S, Bramley K, Chambers D, *et al.* A multicenter study assessing interventional pulmonary fellow competency in electromagnetic navigation bronchoscopy. *ATS Scholar* 2022;3:220–228.

6. Lee HJ, Lerner AD, Coleman B, Semaan R, Mallow C, Arias S, *et al.* Learning Electromagnetic Navigational Bronchoscopy and Percutaneous Transthoracic Needle Biopsy (LEAP): a pilot study. *J Bronchology Interv Pulmonol* 2019;26:55–61.

7. Candoli P, Ceron L, Trisolini R, Romagnoli M, Michieletto L, Scarlata S, *et al.* Competence in endosonographic techniques. *Panminerva Med* 2019;61:249–279.

8. Mahmood K, Wahidi MM, Shepherd RW, Argento AC, Yarmus LB, Lee H, *et al.* Variable learning curve of basic rigid bronchoscopy in trainees. *Respiration* 2021;100:530–537.

9. Pusic MV, Boutis K, Hatala R, Cook DA. Learning curves in health professions education. *Acad Med* 2015;90:1034–1042.

10. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med* 2006;119:166.e7–16.

11. Mema B, Mylopoulos M, Tekian A, Park YS. Using learning curves to identify and explain growth patterns of learners in bronchoscopy simulation: a mixed-methods study. *Acad Med* 2020;95:1921–1928.

12. Toennesen LL, Vindum HH, Risom E, Pulga A, Nessar RM, Arshad A, *et al.* Learning curves for electromagnetic navigation bronchoscopy using CUSUM analysis. *J Bronchology Interv Pulmonol* [online ahead of print] 27 Sep 2021; DOI: 10.1097/LBR.0000000000000815.

13. Yudkowsky R, Park YS, Downing SM, editors. *Assessment in health professions education*, 2nd ed. New York, NY: Routledge; 2020. pp. 5–7, 17–50.

14. Wood TJ, Pugh D. Are rating scales really better than checklists for measuring increasing levels of expertise? *Med Teach* 2020;42:46–51.