using the most predictive ultrasound parameters of LNs that the authors identified (short-axis diameter, absence of a central hilar structure, delineation of well-defined margins, and presence of central necrosis).

Of note, central necrosis was not predictive of malignancy in their model (only 14 patients had this observation) but was kept in the scoring system, given the historical strong association with malignancy. The prevalence of an EBUS-positive LN was 18%. After the authors constructed the prediction model to discriminate malignant from benign LNs, the c-statistic was 0.72, which is quite reasonable. The analysis demonstrates that LNs scoring ≥3 were predictive of malignancy. The authors suggest that a high Canadian Lymph Node Score may inform improved decision-making for repeat biopsy, particularly for inconclusive cytopathology. In my opinion, this may make sense for hilar (N1) LNs, where one should consider switching to a smaller, more diagnostic needle (22 or 25 gauge), but mediastinal LNs should then be assessed by mediastinoscopy and not EBUS.

Constructing an algorithm to inform the user of a high malignant potential of a target LN, particularly for lung cancer staging, is attractive; however, this proposed scoring system has clear limitations, including the absence of external validation to substantiate the findings. The authors acknowledge that the inter-reviewer reliability was poor across the 12 screeners, but this weakens the translatability of ultrasonographic assessment of LNs across institutions except in the hands of expert EBUS users. Furthermore, although the authors were trying to assess LNs with sonographic characteristics alone, I think the scoring algorithm could be further enhanced by including positron emission tomography characteristics of mediastinal LNs since positron emission tomography/computed tomography is a care standard for staging lung cancer across all centers with access to such technology.

Wide adoption of a scoring system requires simplicity of use, reproducibility across institutions and EBUS practitioners, and validation in different patient populations. Since this Canadian Lymph Node Scoring system is based solely on sonographic features of LNs, adoption will be challenging due to wide observer variability. Nonetheless, this multi-institutional study provides a foundation for constructing an improved scoring algorithm that may include clinical parameters to enhance its sensitivity.

Reference

Commentary: Predicting diagnostic outcomes may aid clinical decision making
Farhood Farjah, MD, MPH, FACS

Hylton and colleagues developed and internally validated a prediction score for malignant mediastinal lymph node
disease—called the Canada Lymph Node Score (CLSN)—using imaging findings from endobronchial ultrasound (EBUS). The CLSN provides guidance on whether to pursue repeat invasive staging when initial EBUS results are uninformative. Key strengths of this study include a prospective study design, multicenter (national) patient recruitment, robust evaluation of variability in imaging interpretation, and a rigorous statistical analysis.

There are 2 types of errors associated with diagnostic tests (and strategies)—false-positives (FPs) and false-negatives (FNs). A CLSN of 3 has a FN and FP rate of 69% and 4%, respectively. The potential impact of these errors depends on how one uses the CLNS to guide decisions. Presumably, a high-risk of nodal disease (ie, CLNS ≥3) would result in repeat invasive staging, whereas a low risk of nodal disease would result in pursuit of surgical therapy. Under this scenario, the CLNS-guided selective strategy would rarely subject patients without nodal disease to undergo repeat invasive staging. It would, however, omit repeat invasive staging for at least 69% of truly positive nodes. Most patients and clinicians would be uncomfortable with this degree of error because the potential consequences of missing true nodal disease are great (ie, inappropriate treatment selection).

As a next step, the investigators—led by Dr Waël Hanna—are assessing the feasibility of randomizing patients with lung cancer initially evaluated by computed tomography, positron emission tomography, and EBUS to routine versus selective (ie, CLNS-guided) repeat invasive staging (NCT03859349). In this feasibility trial, the CLNS-guided strategy uses a CLNS ≥1 (ie, lower threshold for risk of true nodal disease) to pursue repeat invasive staging. Consequently, FN and FP rates should be 17% and 63%, respectively. Lower chances of missing true nodal disease would likely be more acceptable for clinicians and patients, even though it results in more repeat invasive staging in patients who will not benefit from it. Fortunately, severe procedure-related adverse events are extraordinarily rare and comparatively this strategy should still result in fewer procedures than a strategy of routine repeat invasive staging for uninformative initial EBUS results. Although convenient to think of diagnostic strategies in terms of their errors, we have no information on the impact of these errors on patients. A randomized trial is the best way to evaluate the impact of prediction models on patient outcomes.

In addition to Dr Hanna’s trial, another one to watch for is the MEDIAS Trial. Twenty Dutch centers and 1 Belgian hospital will randomize patients with lung cancer to endosonography only versus endosonography with selective mediastinoscopy (for negative endosonography results). Although these trials differ somewhat with respect to the diagnostic algorithms they examine, they both use an experimental rather than observational design to tackle tough questions about competing diagnostic strategies. Trials such as these are sorely needed in our field. Dr Hanna and his team are congratulated for not only developing a prediction score for nodal disease but also evaluating its impact using a randomized study design.

References