RSGC
Royal St. George's College

The Young Researcher

2025 Volume 9 | Issue 1

When AI Joins the Scope: Canadian Endoscopists' Perceptions of NodeAI Versus Conventional Methods for Identifying Lymph Node Malignancies in EBUS Imaging

Ria Datta

Recommended Citation
Datta, R. (2025). When AI Joins the Scope: Canadian endoscopists' perceptions of NodeAI versus conventional methods for identifying lymph node malignancies in EBUS imaging. *The Young Researcher,* 9(1), 196-213. http://www.theyoungresearcher.com/papers/datta.pdf

ISSN: 2560-9815 (Print) 2560-9823 (Online) Journal homepage: http://www.theyoungresearcher.com

When AI Joins the Scope: Canadian Endoscopists' Perceptions of NodeAI Versus Conventional Methods for Identifying Lymph Node Malignancies in EBUS Imaging

Ria Datta

Abstract: This study investigates the potential impact of integrating NodeAI, an AI-assisted tool, into endobronchial ultrasound transbronchial needle aspiration (EBUS-TBNA) procedures for cancer staging and lymph node (LN) biopsy. Using a convergent parallel design mixed-methods approach, nine experienced participants, including thoracic surgeons and pulmonologists in North America, were surveyed and participated in a focus group to assess their perspectives on AI integration into EBUS-TBNA cases. The baseline and endline surveys measured shifts in opinions regarding NodeAI's usability in diagnostic accuracy, procedure time, and ease of use. Qualitative insights were gathered through open-ended questions during the focus group to explore clinicians' views on AI's potential role, while quantitative data was captured using scales/ratings. The study found that, while most participants expressed satisfaction with current EBUS-TBNA practices, concerns around over-reliance on AI, data privacy, and the technology's accuracy surfaced during discussion. However, following exposure to NodeAI, participants' views became more favorable, with an increased likelihood of incorporating AI into clinical practice. Key benefits identified included improved diagnostic speed, reduced false positives/negatives, and potential cost savings. The findings suggest that AI tools like NodeAI could enhance decision-making, reduce procedure time and resources, while also presenting challenges related to workflow integration and over-reliance, especially for less experienced individuals.

Keywords: artificial intelligence, NodeAI, EBUS-TBNA, lymph node diagnostic imaging, clinician perceptions, convergent parallel mixed methods

Introduction

As artificial intelligence (AI) increasingly transforms healthcare, its potential to revolutionize diagnostic procedures elicits both excitement and skepticism among providers. Since its introduction in the early 2000s, endobronchial ultrasound transbronchial needle aspiration (EBUS-TBNA) has transformed the diagnosis and staging of lung cancer, providing a minimally invasive method to assess lymph nodes (LN) in the thoracic region. EBUS is a minimally invasive technique for assessing LNs in the thoracic region, aiding in the diagnosis of lung cancer (Navani et al., 2015). By combining a bronchoscope with ultrasound technology, EBUS enables real-time imaging of LNs, allowing clinicians to evaluate their size, shape, and location, which is crucial for accurate cancer staging (Czarnecka-Kujawa & Yasufuku, 2017). TBNA involves a needle that passes through the bronchial wall to sample tissue from LNs or other structures within the chest. This combination is effective in diagnosing a variety of conditions, including lung cancer, infectious diseases, and airway lesions, while also facilitating the staging of diseases through mediastinal LN analysis (Aziz, 2012). In Canada, the Canada Lymph Node Score (CLNS) is used to assess EBUS-TBNA biopsies (Hylton et al., 2020; He et al., 2023). This introduces potential variability in diagnoses due to the subjective nature of human interpretation, highlight-

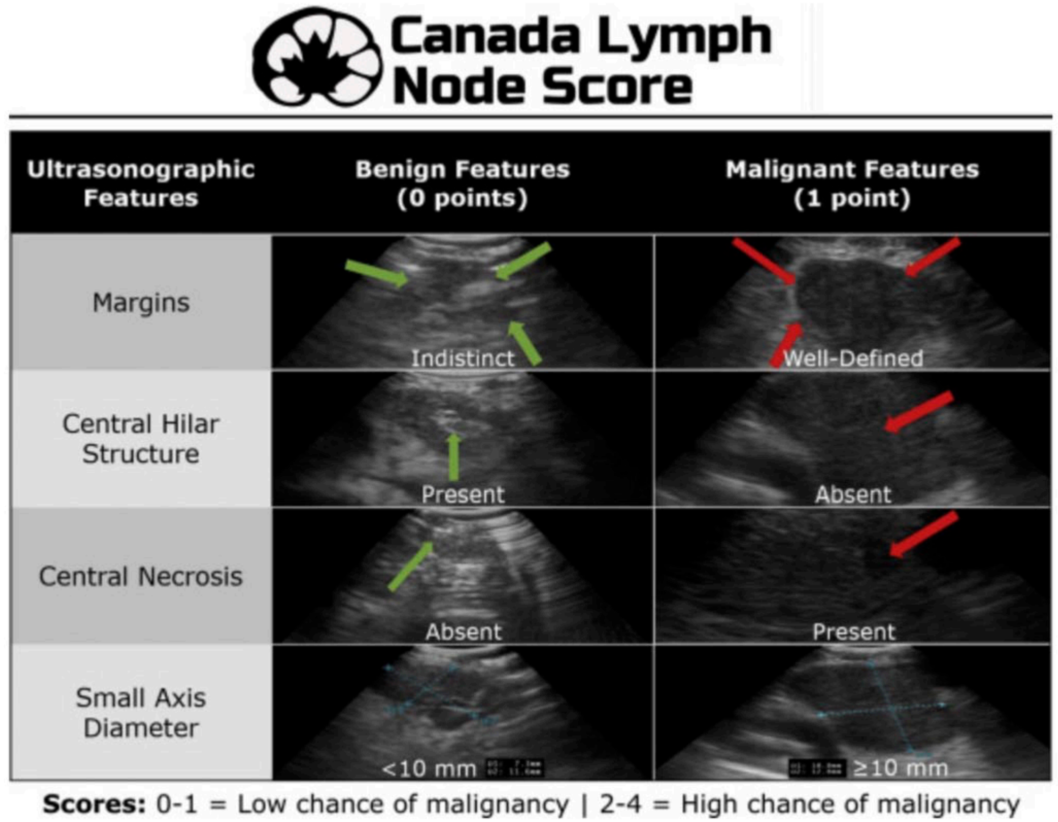

Figure 1. Comparison of benign and malignant versions of ultrasonographic feature that comprise the Canada Lymph Node Score.

ing the need for more standardized and reliable methods to assess LN involvement in lung cancer.

To address this issue, a team at McMaster Hospital developed NodeAI, formally known as NeuralSeg, an AI program that uses GPU processing to automatically segment LN images during EBUS procedures (Churchill, 2022; NeuralSeg, n.d.). By automating the scoring process, NodeAI aims to reduce human error and provide a more consistent interpretation of the images. This has the potential to standardize the staging while using GPU processing for rapid image analysis, ultimately enhancing diagnostic consistency, accuracy, and timeliness.

Through this research, I investigate endoscopists' perceptions of NodeAI efficacy in improving diagnostic accuracy for cancer staging compared to conventional methods. As AI continues to shape healthcare, successful integration into clinical practice depends on healthcare professionals' acceptance and trust. Understanding these perspectives is crucial for broader adoption of AI in lung cancer diagnostics. In order to explore expert perceptions of AI integration into EBUS procedures, the next section reviews existing literature on EBUS and addresses the gap in research. This is followed by a discussion of the design and methodology of this study. The results of the study and discussion of findings are provided thereafter, followed by the conclusion.

Literature Review

This literature review explores current research on the application of AI in EBUS imaging, the accuracy and efficiency benefits AI provides, the perceptions of healthcare professionals towards its adoption, and identifies the gap in research that this study aims to address.

2.1 Current Research on Identifying LN Malignancies through EBUS

In a 2018 narrative review published in *Lung Cancer*, Danielle A. Hylton, from Thoracic Surgery, McMaster University, and colleagues examined ultrasonographic features of LNs to predict malignancy during EBUS procedures. McMaster University is considered Canada's top health sciences research university and #36 in the world in Clinical and Health (Hylton et al., 2018; U15 Group of Canadian Research Universities, 2023) by the Times Higher Education World University Ranking (McMaster University, 2024). The review found that certain features, such as the absence of a central hilar structure and presence of central necrosis (dead tissue), were often linked to malignancy. Hylton and her colleagues believed that by understanding which features are linked to malignancy, this assessment has the potential to reduce the number of EBUS procedures. Building on this, Hylton et al. (2020) developed the CLNS as a predictive clinical tool to determine the likelihood of malignant LNs and to guide biopsy decisions during EBUS procedures. It is based on four characteristics: distinct margins, absent central hilar structure, present central necrosis, and a large axis diameter (size). A score of 0–2 suggests a lower chance of cancer, while a score of 3–4 suggests a higher chance of cancerous LNs. The CLNS was validated through the analysis of 300 LNs from 140 patients by 12 endoscopists, demonstrating its potential as a reliable predictive tool for identifying malignant LNs during EBUS procedures (Hylton et al., 2020). Richard He, Thoracic Surgery, University of Alberta, further demonstrated the validity of the CLNS. Over one year, CLNS scores for 367 LNs biopsied during endobronchial ultrasound were linked to malignancy outcomes. Higher scores (≥3) showed 84.4% specificity, while 10.1% of nodes with scores < 2 and negative CT/PET scans were still malignant (He et al., 2023).

Triple-normal LNs are those that appear normal on CT scanning, PET scanning, and the EBUS procedure, with a CLNS of less than two (Hylton et al., 2021b). An observational study was conducted by Hylton et al. (2021b) to determine whether LNs classified as triple-normal require routine biopsy. The study assessed 143 triple-normal LNs from 57 patients and found that they had a specificity of 60% and a negative predictive value of 93.1%, with only 5.6% of the nodes being malignant upon pathologic examination. This study suggests that routine biopsy is not required for patients, allowing for a more targeted approach to biopsy in EBUS staging. Sanz-Santos et al. (2022) further affirm this by hypothesizing that "targeted sampling (TS), which omits biopsy of triple-normal LNs during endobronchial ultrasound, is not an inferior staging strategy to systematic sampling (SS) of all lymph nodes" (Hylton et al., 2021a; Hylton et al., 2021b; Sanz-Santos et al., 2022). EBUS patients were randomized to TS or SS. Results showed that TS had a faster procedure time (3 minutes vs. 19 minutes for SS) and missed only 5.45% of cancer cases, below the 6% threshold. This suggests that TS may offer a more efficient and focused approach to staging, minimizing unnecessary biopsies and associated risks, while still ensuring accurate cancer detection.

2.2 Advancements of AI in Diagnostic Imaging Regarding Procedures

The integration of AI into diagnostic imaging has become increasingly significant, with studies demonstrating its ability to enhance both image quality and diagnostic accuracy. For example, Dr. Hosny and colleagues from Harvard Medical School (2018) found that AI improves medical imaging by enhancing image quality, increasing diagnostic accuracy, and reducing interpretation time. To validate this, Isabella Churchill and her team at McMaster Hospital (Churchill et al., 2022) conducted a two-phase study using the NodeAI algorithm, achieving 73% accuracy in Phase A by training and comparing on past and known LN images. In Phase B, accuracy increased to 76% when assessed on unseen new images, improving NodeAI's potential for real-time diagnostic scoring.

NodeAI is an advanced machine learning algorithm that evaluates EBUS images and provides real-time diagnostic scores based on the CLNS. It is also

used in other imaging procedures, like MRI, to automatically segment images, thereby reducing time and costs. Leveraging deep learning and GPU processing, NodeAI speeds up analysis from days of manual work to just minutes (Gatti, 2024; NeuralSeg, n.d.).

Dr. Anthony Gatti, a graduate student from Stanford Medical and McMaster University, first developed NodeAI (Gatti, 2018) to automate femoral cartilage segmentation from high-resolution MRI data. The study included 172 MRI scans from 86 individuals, split into training, testing, and validation sets. Segmentation accuracy was assessed using the Dice Similarity Coefficient (DSC), comparing automated results to manual segmentations. NodeAI achieved a mean accuracy of 88.3% for 28 images, indicating high segmentation accuracy. The average segmentation time was 56 seconds, demonstrating the method's efficiency (Gatti, 2018; Gatti, 2024). Furthermore, NodeAI was used in a study by Yogita Patel and colleagues from McMaster University, including Gatti, aimed to validate a stiffness area ratio from endobronchial ultrasound elastography images for diagnosing mediastinal LN malignancy in non-small cell lung cancer. NodeAI assessed bronchial tissue stiffness and created a map to differentiate tissue layers through shades of blue based on elasticity. They analyzed 210 LN images from 124 patients and found 70.59% accuracy, 43.04% sensitivity, 90.74% specificity (Patel et al., 2024). This reveals that NodeAI is adaptable, versatile and beneficial to various medical procedures, building credibility in its incorporation in EBUS procedures.

2.3 Perceptions of AI Among Healthcare Professionals

Despite AI's potential, its integration into healthcare has been staggered due to the controversy and barriers that come with it. The implementation of AI in healthcare faces several ethical barriers, primarily concerns about, "privacy, trust, consent, and conflicts of interest," (Ahmed et al., 2023). Molla Ahmed, Pediatric Respiratory Medicine, University Hospitals of Leicester and colleagues noted that among 59 articles reviewed, 20 highlight confidentiality as a substantial concern. The General Data Protection Regulation (GDPR) emphasizes patient control over their data and the need for informed consent when sharing data with AI developers (Marcu et al., 2019).

Trust in AI is another significant barrier, noted in 25 studies (Ahmed et al., 2023). Healthcare professionals often lack training to evaluate AI tools, and the lack of rigorous, randomized controlled trials creates skepticism (Ahmed et al., 2023; Chen & See, 2020). Many AI algorithms operate as "black boxes," where users can see the input and output but lack insight into the process in between, and how AI arrived at those results (Sakamoto et al., 2020). This reduces confidence in validation and puts strain on doctor-patient relationships, raising concern and hesitance about AI's integration into healthcare treatment (Brady & Neri, 2022).

2.4 Gap Analysis

Despite research on AI efficacy in EBUS procedures, there is a gap in understanding clinicians' perspectives on adopting these technologies. Current literature often overlooks the factors influencing acceptance among medical providers, impacting the successful implementation of AI in diagnostics (Churchill et al., 2022). My research aimed to explore how Canadian endoscopists perceive the effectiveness, usability, and potential barriers to incorporating NodeAI into EBUS procedures. The findings of this study hope to provide valuable insights into the barriers and facilitators of AI adoption in clinical settings, ultimately guiding the successful integration of AI technologies like NodeAI into routine diagnostic practices (Koseoglu et al., 2023). By understanding clinicians' perspectives, this research could inform strategies to improve AI adoption, enhancing diagnostic accuracy and efficiency in EBUS procedures.

3. Methodology

This study used a convergent parallel mixed methods (CPMM) design, incorporating a qualitative nominal group technique (NGT) alongside baseline and endline surveys to capture shifts in opinions and facilitate discussion among expert participants. The following methodology section will detail the study design, participant selection, data collection process, and ethical considerations.

3.1 Design and Approach

This study employs a CPMM design developed by Creswell (2009), which "consists of taking quantitative and qualitative data collection and analysis and comparing the two and then interpreting them" (Harvard Catalyst, 2014). Mixed methods combine both quantitative and qualitative approaches, collecting both numerical and non-numerical data, integrating numerical data with in-depth insights, offering a fuller understanding. The CPMM design collects and analyzes two types of data simultaneously but separately (Damyanov, 2023). Also, it allows for cross-validation of results, as findings from qualitative and quantitative sources can support or contrast with each other, increasing the reliability of the conclusions (Ahmed et al., 2024). In CPMM, qualitative data often comes from interviews, focus groups, or observations, while quantitative data is typically derived from surveys, tests, or statistical measures (Damyanov, 2023).

This design is particularly useful when researching a new phenomenon with limited existing knowledge and literature. The CPMM is primarily used in healthcare and medical research to evaluate new interventions, technologies, and treatment strategies, especially when there is minimal prior research (Tomasi et al., 2018). It is beneficial in contexts where both broad patterns and individual experiences are valuable, as it enables researchers to compare different data types (Alele & Malau-Aduli, 2023). An example of a CPMM design is a study by Rosenkranz, Wang, and Hu from the University of Western Sydney School of Medicine (2015), which aimed to explore what motivates and demotivates medical students to pursue research. The study collected quantitative data through surveys and qualitative data via semi-structured interviews. The data were analyzed separately, and the results were then compared and integrated. Since this study aims to analyze perceptions, which, like motivations, are driven by values and are subjective, this design is particularly well-suited to capture these complex factors.

Alternatives, such as the Delphi method, which relies on iterative rounds of surveys to build consensus among experts, are often time-consuming and can be challenging for busy healthcare professionals to fully engage with. Additionally, the lack of real-time interaction in the Delphi method limits the opportunity for in-depth discussions, making it harder to explore the complex implications of AI in clinical practice and how it could impact decision-making (Nasa et al., 2021). Similarly, the NGT, which generates consensus through structured face-to-face discussions, is effective in gathering diverse perspectives but does not capture shifts in opinions over time. Without a baseline to measure changes, the NGT falls short in understanding how expert views evolve, particularly when new information or perspectives are introduced (Burke et al., 2019). These limitations make the CPMM a complete and nuanced understanding by combining qualitative insights from focus groups with quantitative survey data. This approach informs AI implementation and future research, addressing the research question of the impact of NodeAI on diagnostic accuracy and efficiency.

3.2 Sampling and Recruitment

For this study, purposeful sampling was employed to select nine thoracic surgeons and pulmonologists with expertise in EBUS and NodeAI technology (Patton, 2015). This sampling strategy ensures diversity while maintaining relevance to the research question, ensuring knowledgeable and insightful feedback. Recruitment was conducted via email, using an email list provided by an external advisor, a cardiothoracic surgeon. Once participants were confirmed, they were asked to commit to the duration of the study to ensure consistent and reliable participation throughout the research process.

3.3 Data Collection

3.3.1 Surveys

The first questionnaire was sent to participants to gather baseline data for the study. The survey included both open and closed-ended questions to explore demographic and professional information. The demographic section collected data on the expert's name, institutional affiliations, and credentials, helping to gauge the diversity of perspectives and the expertise represented. The professional section focused on the participants specialization, years of experience with EBUS and satisfaction with biopsy yields. Closed-ended questions, including Likert scale ratings, collected quantitative data on practices such as LN selection and the use of endosonographic scoring systems (like CLNS).

Following the focus group discussion, an endline survey was administered to capture shifts in expert opinions. The survey assessed changes in perspectives on the use of NodeAI in diagnostics, including perceived improvements in diagnostic accuracy, procedure time, and ease of use. Questions measured the likelihood of incorporating NodeAI into clinical practice. This allows for comparison with baseline data, providing insights into how group discussions have influenced individual views. This design is advantageous as it offers a clear evaluation of how specialist perceptions evolve, enhancing the study's understanding of AI's impact on EBUS procedures.

3.3.2 Focus Group

The guide aimed to explore participants' approaches to mediastinal staging, LN biopsy criteria, and their opinions on AI's potential role. The focus group included open-ended questions that provide deep qualitative data, such as how ultrasound features and the CLNS influence their decisions, and how they perceive AI's use. After an online demonstration of NodeAI, participants shared their initial thoughts, assessed its usefulness, and discussed how it might be incorporated into clinical practice. These qualitative findings were recorded in the quantitative survey that measured shifts in participants' views on AI immediately at the end of the focus group discussion. The integration of both qualitative and quantitative methods allows for a more comprehensive understanding of NodeAI's impact on clinical practices and its potential for implementation.

3.4 Data Analysis

Following my data collection, participants responses were synthesized into a comprehensive report. Qualitative data analysis was used to identify any changes in expert opinions and determine if a consensus had emerged regarding the use of NodeAI in EBUS procedures. Thematic analysis was applied to the open-ended responses: qualitative data was coded into categories, such as expected themes, unexpected findings, and significant insights, to uncover underlying patterns and trends, while quantitative data, including binary responses and Likert scales, was summarized and expressed in graphs and figures. Less emphasis was placed on statistical analysis compared to the qualitative insights, which form the core of the findings.

3.5 Ethics Memorandum

As this research did not involve the use of patient data or information, patient consent was not required. All participants were adults aged 25 and above. Before data collection, all participants signed an informed consent document that details their rights as participants. To ensure confidentiality, all results collected from the survey remained anonymous and were not used to identify any of the respondents. Participants were given pseudonyms, and their data is kept in password-protected files. This study was approved by the school's Internal Ethics Review Board and all procedures adhere to the ethical standards of the institution.

4. Results and Discussion

The study included nine participants with diverse specialties, professional experience levels, and practice settings (Table 1). This range of backgrounds provides a well-rounded perspective on the potential role of AI-assisted tools like NodeAI in EBUS procedures.

4.1 Concerns or Factors Influencing Confidence of AI and EBUS

Data collected from the baseline form, endline form, and focus group uncovered trends regarding the factors influencing people's perceived confidence in EBUS-TBNA and AI, specifically in NodeAI. In the baseline and endline surveys, participants were asked to distinguish their concern level regarding the accuracy of AI algorithms in medical diagnostics.

As shown in Figure 1, the double bar graph shows that after the focus group, participants' concerns surrounding the accuracy of AI algorithms increased. The number of participants who were somewhat unconcerned dropped to zero, while those who were somewhat concerned and concerned rose, reflecting increased awareness and skepticism about AI's reliability. Figure 2 reveals specific reasons.

Table 1. Subject Information

Participant	Practice	Affiliated organization(s)	Years of Practice
1	Thoracic Surgery	Academic or university hospital, research practice	>10
2	Thoracic Surgery	Academic or university hospital, research practice	6-10
3	Interventional Pulmonology	Academic or university hospital, community practice	0-5
4	Both	Academic or university hospital	6-10
5	Interventional Pulmonology	Academic or university hospital, research practice	0-5
6	Thoracic Surgery	Academic or university hospital	>10
7	Thoracic Surgery	Community practice, research practice	0-5
8	Thoracic Surgery	Academic or university hospital	>10
9	Interventional Pulmonology	Academic or university hospital	>10

Figure 2 highlights the primary concerns clinicians have regarding AI in medical diagnostics, specifically in EBUS-TBNA. The most significant concern, identified by 78% of participants, is the potential for over-reliance on technology, indicating a fear of losing human oversight in decision-making. Participant 4 expressed their concern for over-reliance on AI among less experienced bronchoscopists like trainees: "I am afraid that they become so dependent on it [NodeAI] they can't read lymph nodes anymore," referring to trainees potentially losing critical diagnostic skills. While initially viewed as a limitation, this concern was later reframed as an educational opportunity: "For those just starting, NodeAI could be a game-changer in helping them decide which to biopsy." With NodeAI real-time feedback and training, participant 4 highlighted that it could be used to teach trainees the differences between malignant and benign LNs, which could improve future clinicians and help reduce operator variability.

Ahmed and his colleagues' concern about "privacy, trust, consent, and conflicts of interest" appears to be common among the participant sample with 33% of participants believing that reliability of AI-generated results, data privacy and security issues, and cost of implementation were of concern (Ahmed et al., 2023). The concern about lack of training and support was the least prevalent, at only 11%, suggesting that clinicians may feel less concerned about educational gaps than about the tangible risks and challenges associated with implementing AI.

Additionally, challenges remain regarding accuracy, with one participant commenting, "Accuracy (79%) is not great, and if NodeAI is supposed to be the final source to decide on whether to biopsy a LN, we need higher accuracy." This concern underscores the need

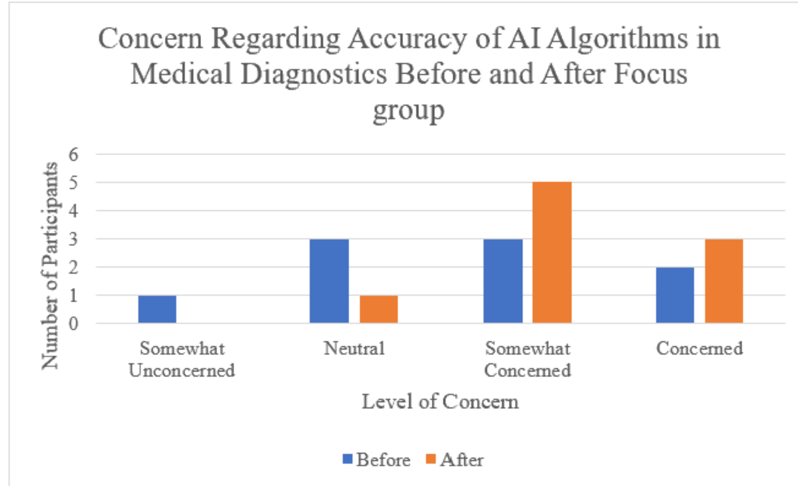

Figure 1. Double bar graph displaying concern regarding the accuracy of AI algorithms in medical diagnostics before and after focus group.

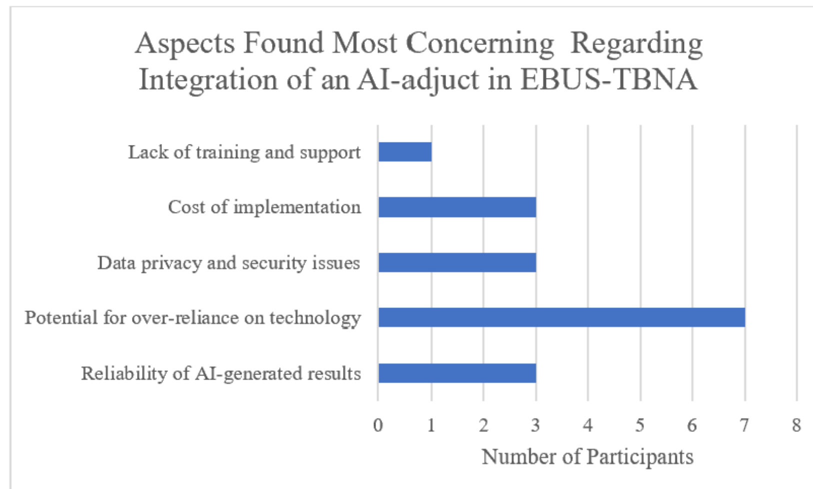

Figure 2. Bar graph of aspects found most concerning regarding accuracy of AI algorithms integrated EBUS-TBNA cases.

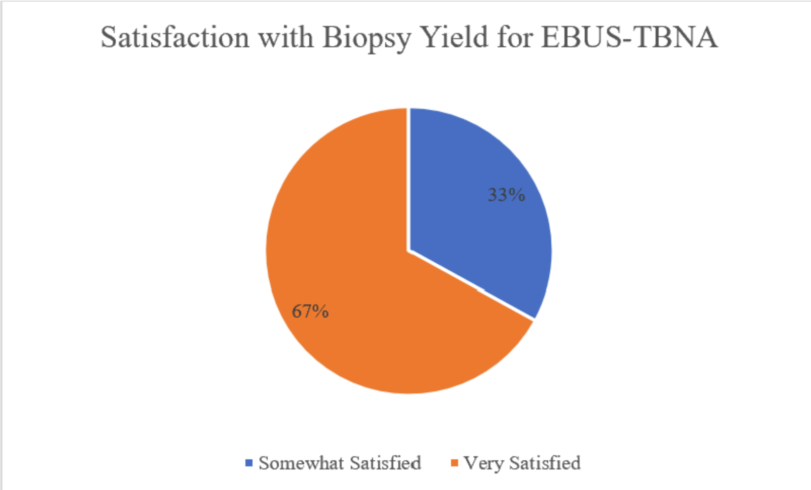

Figure 3. Participants satisfaction with biopsy yield for current EBUS-TBNA practices.

for continuous refinement of the AI model to ensure its practical applicability in clinical decision-making.

Despite the many concerns, overall satisfaction determined in the baseline survey seems positive.

In question 4 of the baseline survey, participants were asked to rate their satisfaction with the current EBUS-TBNA practice. As shown in Figure 3, 67% of participants reported that they were very satisfied with the biopsy yield from EBUS-TBNA, and 33% were somewhat satisfied. This indicates that most clinicians felt that the procedure was effective in yielding satisfactory results. Yet, most clinicians have concerns regarding current methods, as shown in Figures 1 and 2.

4.2 Current Methods

Figure 4 reveals that the majority of participants (78%) prefer using both CT/PET and endosonographic criteria for LN selection in EBUS-TBNA. In contrast, a smaller group (22%) prefers to biopsy all LNs regardless of imaging criteria or pre-test probability. This suggests a preference for a more targeted sampling as described by Hylton et al. (2021b). Guidelines recommend systematic staging, but in practice, many clinicians choose a more targeted approach, particularly for small or triple-normal LNs. According to three participants during the focus group, there is a desire to avoid unnecessary sampling of low-yield or likely benign nodes, supporting Sanz-Santos et al.'s (2022) theory that TS is not inferior to SS. However, in the focus group, Participant 2 was strongly against targeted sampling despite its 80% accuracy rate, stating, "I think more information is better… If you don't try, you don't know for sure, and you're still saying that 5% of triple-negative LNs can have cancer in them. It's still not 0%, right?" Participant 7 challenged this perspective, arguing, "I think these imaging devices will be extremely important because there are so many lymph nodes. We're not just talking about one lymph node within one station; you might see two or three lymph nodes, and we cannot sample all of them. You really want to look at the ultrasound image and sample the lymph nodes that are the most suspicious." Like Sanz-Santos and Hylton, supporters of targeted sampling emphasized that it saves time and resources, ultimately reducing waiting times, human availability, and costs.

Figure 5 shows a wide range of usage of endosonographic scoring systems, with 44% of participants reporting they always use the system, while 22% never use it. A small number of respondents rarely (11%) or occasionally (11%) use the system, and 11% use it fre-

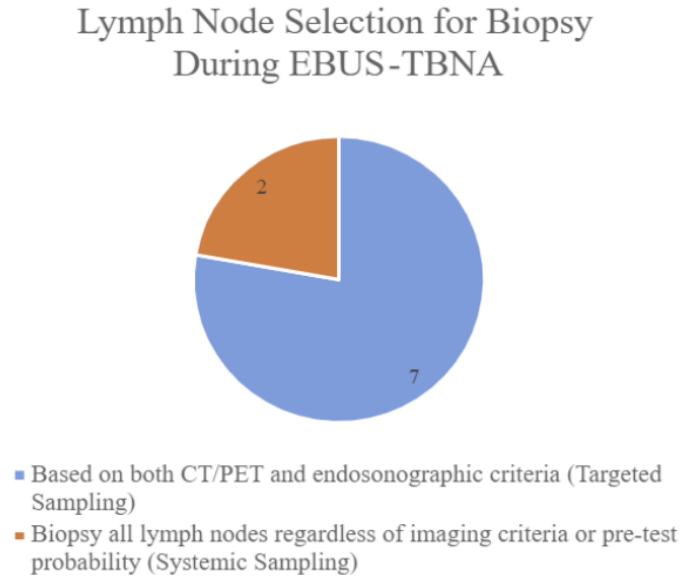

Figure 4. Pie graph of participants selection process for biopsy during EBUS-TBNA.

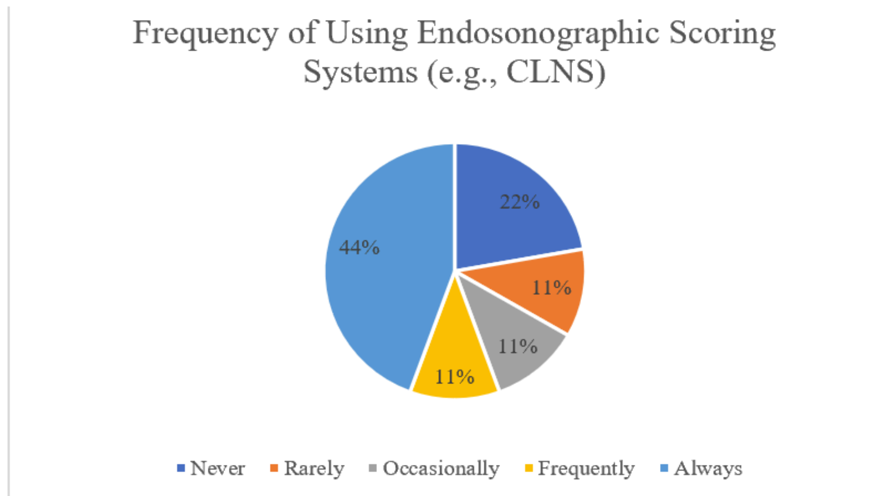

Figure 5. Pie graph showing participants frequency of using endosonographic scoring systems (e.g. CLNS) during EBUS-TBNA cases.

quently. This variability suggests that while some clinicians consistently use scoring systems, others may either lack familiarity or face barriers to incorporating them into routine practice. This is further supported by Participant 5 from Edmonton, who, after completing respirology training, noted that they did not use a scoring system like CLNS. Instead, their approach was to wait for substantial LN biopsies to return as non-diagnostic, and if that occurred, they would typically repeat the procedure. However, after coming to McMaster Hospital in Hamilton and learning about the CLNS, their approach changed significantly. Additionally, Participant 9 from Oshawa, Southern Ontario, reported not following the CLNS, highlighting regional variations in standardized protocols.

4.3 Perceptions and Potential Surrounding AI Integration and EBUS-TBNA

Figure 6 shows a slight increase in clinicians' perceptions of NodeAI's benefits for predicting malignancy in LNs during EBUS-TBNA. In the pre-survey, 33% of participants found it not beneficial, 33% viewed it as slightly beneficial, and 33% considered it beneficial. Post-survey results showed an increase in the number of participants who rated NodeAI as slightly beneficial (44%), a decrease in those who rated it not beneficial (22%), while beneficial remained unchanged (33%). This shift suggests that, through further understanding and discussion of AI adjuncts, clinicians may develop a more positive view of its application.

Like positive perceptions, the likelihood to use an AI-adjunct (NodeAI or other) for predicting Malignancy in EBUS-TBNA cases also increased. Figure 7 shows an increase in the likelihood of using an AI-adjunct (NodeAI or similar) for predicting malignancy in EBUS-TBNA cases after exposure to the technology. Initially, 78% of participants were likely to use AI, with only 22% neutral, and none were unlikely. After exposure, the number of participants who were likely to use AI rose to 89%, while only one participant indicated neutrality. This suggests an overall more positive perception of NodeAI specifically as supported by Figure 8.

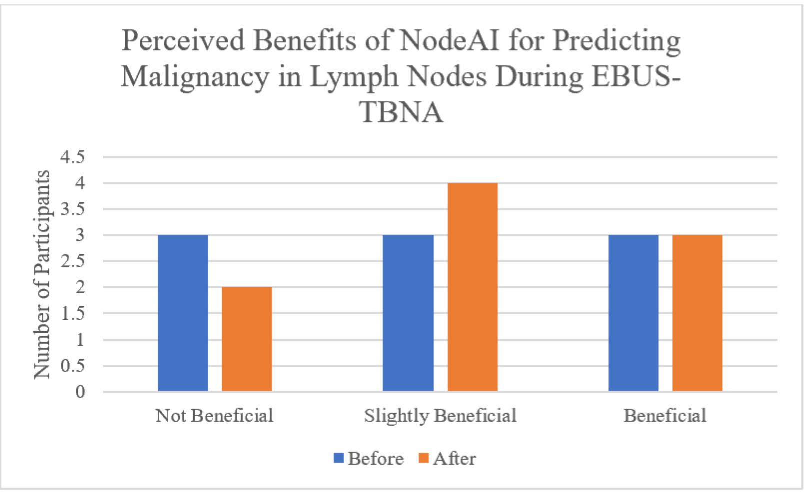

Figure 6. Graph of participant's perceived benefits of NodeAI for predicting malignancy in lymph nodes during EBUS-TBNA.

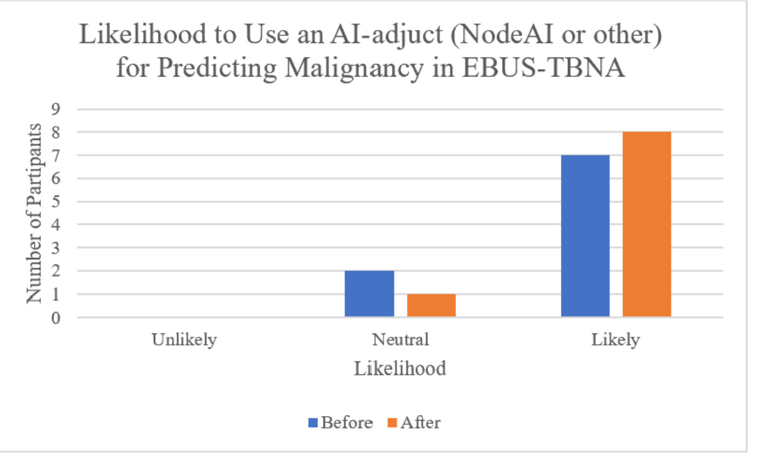

Figure 7. Graph showing participants likelihood to use an AI-adjunct (NodeAI or other) for predicting malignancy on EBUS-TBNA before and after the focus group.

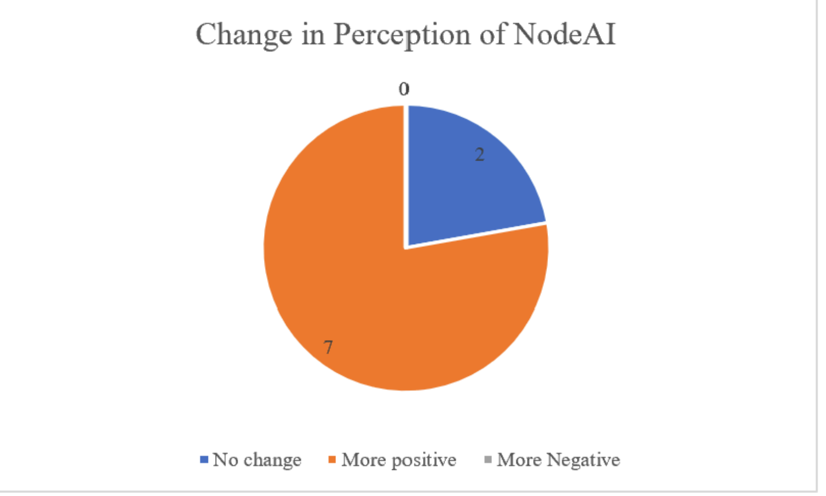

Figure 8. Graph displaying participants change in perceptions of NodeAI after the focus group.

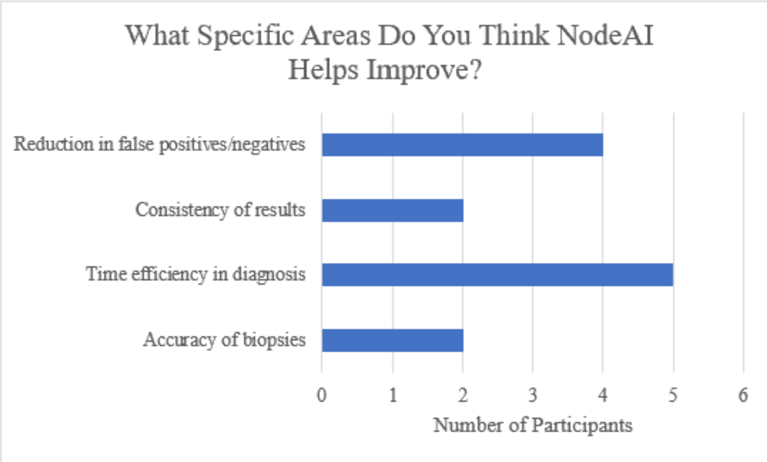

Figure 9. Graph showing specific areas NodeAI is perceived to improve after focus group by participants.

Determined by a question in the endline survey, Figure 8 reveals that after exposure to NodeAI, 78% of participants reported a somewhat more positive perception, while 22% experienced no change in their views. None of the participants reported a more negative perception, suggesting that exposure to NodeAI had a generally positive influence on clinicians' attitudes. Many factors and perceived benefits were both reported during the survey and discussed in the focus group.

Figure 9 highlights the areas where clinicians believe NodeAI could improve EBUS-TBNA. The most commonly identified benefit was time efficiency in diagnosis, with 56% of participants noting this as a key improvement. This was consistent in the focus group, as many participants believed that NodeAI could reduce procedure times as it speeds up decision making. However, Participant 1 thought that it could prolong procedure time due to its unfamiliarity, yet this concern was quickly diminished after the proposal of creating training modules and tutorials on NodeAI. Furthermore, anesthesia complications, such as sedation issues, have been reported. Participant 2, a clinician from Quebec, mentioned, "We sedate patients, but sometimes they cough and fight back, so we have to change plans. We can't biopsy all the nodes, and we have to skip some." Participant 9 said, "These difficulties worsen in settings where anesthesia availability is limited which reduces ability to access LNs and procedure time on patient due to their consciousness," further emphasizing the importance of speed and targeted sampling to ensure the procedure's success.

This led one participant, currently practicing in California but with prior training in Ontario, to explain, "In Canada and other public systems, reducing procedure time is a major priority. But in the United States, time and reimbursement concerns are different." Despite these differences, Participant 3 emphasized that overall cost savings, such as fewer pathology tests and repeated procedures, are still valuable. Thus, shortening procedure time should not be the main selling point if NodeAI wants to reach the US as well.

Reducing false positives/negatives was also a significantly desired area of improvement, cited by 44% of respondents. Accuracy of biopsies and consistency of results were selected by 22% of participants. This smaller percentage may be due to NodeAI's role in aiding decision-making, but its limitations, such as not assisting with scope movement, capturing clear ultrasound images of the LN, and ensuring conclusive biopsy yield, were also factors. This suggests that Canadian clinicians see the most immediate value of NodeAI in enhancing diagnostic speed and accuracy, with other valuable, yet less obvious benefits that may attract a broader range of users.

Also, achieving high-quality ultrasound images of LNs is objectively challenging without proper experience. This is true for those who are less experienced, learning and have trouble visualizing the LNs. Furthermore, obtaining a sufficient biopsy yield is challenging, but confirming takes even longer. Turnaround time for pathology (especially in the absence of on-site cytology, such as ROSE) can delay treatment decisions. As Participant 2 shared, "If we don't have ROSE [rapid on-site evaluation] the delay in getting pathology results can really impact treatment timelines." While NodeAI cannot directly help with the technicalities of EBUS-TBNA, it can help minimize variability when determining abnormal or normal LNs, a critical concern agreed by the majority of the sample during the focus group. This variability can result in inaccurate LN assessments, making it difficult to determine the best nodes for biopsy.

To improve EBUS-TBNA, the integration of AI tools like NodeAI offers many benefits. According to Figure 6, 44% of participants saw NodeAI as at least slightly beneficial and 33% of participants as beneficial for predicting malignancy. For beginners and less experienced clinicians, AI can assist in identifying the right LNs to biopsy, enhancing decision-making. An AI tool (like NodeAI) that predicts malignancy in real time could help clinicians decide whether or not to biopsy a particular LN. This might reduce passes, limit overall sedation time, and spare pathology resources.

5. Future Research

Looking forward, the future of AI in medical procedures like EBUS-TBNA depends heavily on its integration into work-flow and its ease of use. According to Figure 7, after exposure to NodeAI, 89% of participants were likely to use AI in predicting malignancy, indicating interest and a willingness to adopt the technology. However, participants agreed that testing is crucial to ensuring that clinicians are comfortable using the technology. As Participant 8 explained, "We need to test it out in our individual practices to ensure it is feasible to integrate." All participants in the focus group agreed that the tool must be adaptable to various clinical environments to ensure widespread adoption. Moreover, while ease of use is important, NodeAI must continue to evolve to meet the accuracy standards required for clinical decision-making to build confidence among users. Long-term testing and the integration of training modules will be key in overcoming initial hesitations. As the tool undergoes further development, NodeAI has the potential to improve diagnostic accuracy, reduce procedure times, and offer valuable educational opportunities for clinicians, ultimately enhancing EBUS-TBNA cases and patient outcomes.

6. Limitations

Limitations of this study include groupthink, in which participants align their opinions to the dominant narrative, limiting the diversity of feedback and authenticity of the data (Jhangiani, 2022). To address this, clear expectations were set at the beginning of the session, encouraging participants to express their own views and actively asking for differing opinions. Additionally, since AI is a relatively objective topic and less likely to evoke emotional responses compared to other subjects, it reduces the likelihood of personal biases or discomfort affecting the discussion. Lastly, to prevent any one participant from dominating the conversation and ensure all voices are heard, a round-robin approach was used for certain questions. This ensures that everyone had an equal opportunity to speak and that all perspectives were represented in the discussion.

Moreover, the results are limited by the perspectives of the nine recruited experts. With all participants being either interventional pulmonologists, thoracic surgeons, or both, their views are largely from a scientific perspective. While it is no doubt important to have scientific voices represented, gathering experts from outside the scientific and medical spheres could yield a greater range of perspectives. For instance, incorporating EBUS-TBNA patients could provide valuable insight into the patient viewpoint, including their concerns, expectations, and perspectives on AI-assisted diagnostics. Additionally, another limitation is the overrepresentation of participants from the

Greater Toronto Area, which may not capture practices in other regions. Expanding the study to include participants from diverse geographic areas, such as rural centers or regions with varying levels of technology, workflow and resources, could provide a more comprehensive understanding of how NodeAI might be implemented in different clinical settings.

7. Conclusion

This study aimed to examine the extent to which clinicians perceive NodeAI as a valuable tool in comparison to traditional diagnostic methods. Research revealed current use of sonographic methods such as the CLNS to determine abnormal lymph nodes, and shifting towards a more targeted approach as opposed to systemic as supported by many of the studies and participants. A CPMM approach was used, combining both qualitative and quantitative data from surveys and focus group discussions with nine experienced endoscopists across North America. Through data collection, I found that while all participants are currently satisfied with EBUS-TBNA, all had significant concerns including biopsy yield and nondiagnostic samples, decision making (identifying which nodes to biopsy), and variability in operator experience and diagnosing. After the discussion of NodeAI and AI-adjuncts, participants reported NodeAI's potential, including improved targeted sampling, support for decision making, and education, with most participants citing reduced time and resources. These findings align with existing literature that suggests AI has the potential to improve diagnostic outcomes, but also underscores the importance of workflow integration and proper training to prevent overreliance on technology.

Future research should investigate patients' perceptions of integrating NodeAI and other AI-adjuncts into medical procedures. Further test-trials and development to increase NodeAI's accuracy (currently at 79%) must be done, as all participants identified that as a critical barrier. In addition, determining its value proposition to reach a large audience, as other countries (such as the United States) may be less motivated solely by time savings. Also, participants unanimously expressed a desire to pilot the device in real-time clinical settings to demonstrate workflow feasibility and design larger trials to prove clinical impact, cost savings, and user adoption. Overall, the findings communicate the importance of balancing technological innovation with clinical expertise to deliver the most effective patient care.

NODE AI VERSUS CONVENTIONAL METHODS IN EBUS IMAGING

References

Ahmed, M. I., Spooner, B., Isherwood, J., Lane, M. A., Orrock, E., & Dennison, A. (2023). A systematic review of the barriers to the implementation of artificial intelligence in healthcare. *Cureus*, *15*(10). https://doi.org/10.7759/cureus.46454

Alele, F., & Malau-Aduli, B. (2023). Mixed methods study designs. *Jcu.pressbooks.pub*. https://jcu.pressbooks.pub/intro-res-methods-health/chapter/5-5-mixed-methods-study-designs/

Aziz, F. (2012). Endobronchial ultrasound-guided transbronchial needle aspiration for staging of lung cancer: a concise review. *PubMed*, *1*(3), 208–213. https://doi.org/10.3978/j.issn.2218-6751.2012.09.08

Brady, A. P., & Neri, E. (2020). Artificial Intelligence in Radiology—Ethical Considerations. *Diagnostics*, *10*(4), 231. https://doi.org/10.3390/diagnostics10040231

Burke, C. S., Georganta, E., & Marlow, S. (2019). A bottom-up perspective to understanding the dynamics of team roles in mission critical teams. *Frontiers in Psychology*, *10*(1322). https://doi.org/10.3389/fpsyg.2019.01322

Chen, J., & See, K. C. (2020). Artificial intelligence for COVID-19: A rapid review (Preprint). *Journal of Medical Internet Research*. https://doi.org/10.2196/21476

Churchill, I. F., Gatti, A. A., Hylton, D. A., Sullivan, K. A., Patel, Y. S., Leontiadis, G. I., Farrokhyar, F., Hanna, W. C. An artificial intelligence algorithm to predict nodal metastasis in lung cancer. *The Annals of Thoracic Surgery*, vol. 114, no. 1, July 2022, pp. 248–256, https://doi.org/10.1016/j.athoracsur.2021.06.082

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications.

Czarnecka-Kujawa, K., & Yasufuku, K. (2017). The role of endobronchial ultrasound versus mediastinoscopy for non-small cell lung cancer. *Journal of Thoracic Disease*, *9*(S2), S83–S97. https://doi.org/10.21037/jtd.2017.03.102

Damyanov, M. (2023). *Mixed Methods Research Guide with Examples*. Dovetail.com. https://dovetail.com/research/mixed-methods-research/

Gatti, A. A. (2018). NEURALSEG: state-of-the-art cartilage segmentation using deep learning – analyses of data from the osteoarthritis initiative. *Osteoarthritis and Cartilage*, *26*, S47–S48. https://doi.org/10.1016/j.joca.2018.02.110

Gatti, A. (2024). *NeuralSeg*. Nubinary.com. https://nubinary.com/case-studies/neuralseg

He, R. X., Hylton, D. A., Eric L.R. Bédard, Johnson, S., Laing, B., Azim Valji, Hanna, W. C., & Turner, S. R. (2023). Clinical validation of the Canada Lymph Node Score for endobronchial ultrasound. *The Annals of Thoracic Surgery*, *115*(6), 1456–1462. https://doi.org/10.1016/j.athoracsur.2021.11.071

Harvard Catalyst. (2014). *Basic Mixed Methods Research Designs*. Harvard.edu. https://catalyst.harvard.edu/community-engagement/mmr/hcat_mmr_sm-6090567e0f943-60905896c80af-60e5fdbc2399e-60e5fdd8057fc-610bf777da6a0-610bf7808de24-610bf792228a4-610bf8685d8f5-610bf-871cbea9-66be4af6dac9f/

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018b). Artificial intelligence in radiology. *Nature Reviews Cancer*, *18*(8), 500–510. https://doi.org/10.1038/s41568-018-0016-5

Hylton, D. A., Turner, J., Shargall, Y., Finley, C., Agzarian, J., Yasufuku, K., Fahim, C., & Hanna, W. C. (2018). Ultrasonographic characteristics of lymph nodes as predictors of malignancy during endobronchial ultrasound (EBUS): A systematic review. *Lung Cancer*, *126*, 97–105. https://doi.org/10.1016/j.lungcan.2018.10.020

Hylton, D. A., Turner, S., Kidane, B., Spicer, J., Xie, F., Forough Farrokhyar, Kazuhiro Yasufuku, Agzarian, J., & Hanna, W. C. (2020). The Canada Lymph Node Score for prediction of malignancy in mediastinal lymph nodes during endobronchial ultrasound. *Journal of Thoracic and Cardiovascular Surgery*, *159*(6), 2499-2507.e3. https://doi.org/10.1016/j.jtcvs.2019.10.205

Hylton, D. A., Turner, J., Shargall, Y., Finley, C., Agzarian, J., Yasufuku, K., Fahim, C., & Hann, A., & Meining, A. (2021a). Artificial intelligence in endoscopy. *Visceral medicine*, *37*(6), 471–475. https://doi.org/10.1159/000519407

Hylton, D. A., Kidane, B., Spicer, J., Turner, S., Churchill, I., Sullivan, K., Finley, C. J., Shargall, Y., Agzarian, J., Seely, A. J. E., Yasufuku, K., Hanna, W. C., & Canadian Association of Thoracic Surgery Research Group (2021b). Endobronchial Ultrasound Staging of Operable Non-small Cell Lung Cancer: Do Triple-Normal Lymph Nodes Require Routine Biopsy? *Chest*, *159*(6), 2470–2476. https://doi.org/10.1016/j.chest.2020.12.050

Jhangiani, R. (2022, September 26). *Group Decision Making – Principles of Social Psychology – 1st International Edition*. Opentextbc.ca. https://opentextbc.ca/socialpsychology/chapter/group-decision-making/

Koseoglu, F. D., Alıcı, I. O., & Er, O. (2023). Machine learning approaches in the interpretation of endobronchial ultrasound images: a comparative analysis. *Surgical endoscopy*, *37*(12), 9339–9346. https://doi.org/10.1007/s00464-023-10488-x

Marcu, L. G., Boyd, C., & Bezak, E. (2019). Current issues regarding artificial intelligence in cancer and health care. Implications for medical physicists and biomedical engineers. *Health and Technology*, *9*(4), 375–381. https://doi.org/10.1007/s12553-019-00348-x

McMaster University. (2024, March 3). *Rankings - Discover McMaster*. Discover McMaster. https://discover.mcmaster.ca/about/rankings/#:~:text=THE%20World%20University,Academic%20Rankings%20of%20World%20Universities

Nasa, P., Jain, R., & Juneja, D. (2021). Delphi methodology in healthcare research: How to decide its appropriateness. *World Journal of Methodology*, *11*(4), 116–129. https://doi.org/10.5662/wjm.v11.i4.116

Navani, N., Nankivell, M., Lawrence, D. R., Lock, S., Makker, H., Baldwin, D. R., Stephens, R. J., Parmar, M. K., Spiro, S. G., Morris, S., Janes, S. M., & Lung-BOOST trial investigators. (2015). Lung cancer diagnosis and staging with endobronchial ultrasound-guided transbronchial needle aspiration compared with conventional approaches: An open-label, pragmatic, randomised controlled trial. *Lancet Respiratory Medicine, 3*(4), 282-9.

NeuralSeg. (n.d.). *neuralseg*. Nubinary. https://nubinary.com/case-studies/neuralseg

Patel, Y. S., Gatti, A. A., Farrokhyar, F., Xie, F., & Hanna, W. C. (2024). Clinical utility of artificial intelligence-augmented endobronchial ultrasound elastography in lymph node staging for lung cancer. *JTCVS Techniques*, *27*, 158–166. https://doi.org/10.1016/j.xjtc.2024.06.024

Patton, M.Q. (2015). Qualitative research & Evaluation Methods: Integrating theory and practice. Fourth edition. SAGE Publications, Inc.

Rosenkranz, S. K., Wang, S., & Hu, W. (2015). Motivating medical students to do research: a mixed methods study using Self-Determination Theory. *BMC Medical Education*, *15*(1). https://doi.org/10.1186/s12909-015-0379-1

Sakamoto, T., Furukawa, T., Lami, K., Pham, H. H. N., Uegami, W., Kuroda, K., Kawai, M., Sakanashi, H., Cooper, L. A. D., Bychkov, A., & Fukuoka, J. (2020). A narrative review of digital pathology and artificial intelligence: focusing on lung cancer. Translational Lung Cancer Research, 9(5), 2255–2276. https://doi.org/10.21037/tlcr-20-591

Sanz-Santos, J., Serra, P., & Rosell, A. (2022). Systematic or targeted sampling during endobronchial ultrasound for mediastinal staging in patients with lung cancer and abnormal mediastinum. *JTCVS Open*, *11*, 346–347. https://doi.org/10.1016/j.xjon.2022.01.030

Tomasi, J., Warren, C., Kolodzey, L., Pinkney, S., Guerguerian, A.-M., Kirsch, R., Hubbert, J., Sperling, C., Sutton, P., Laussen, P., & Trbovich, P. (2018). Convergent parallel mixed-methods study to understand information exchange in paediatric critical care and inform the development of safety-enhancing interventions: a protocol study. *BMJ Open*, *8*(8), e023691. https://doi.org/10.1136/bmjopen-2018-023691

U15 Group of Canadian Research Universities. (2023). *McMaster University - U15 Group of Canadian Research Universities*. U15 Group of Canadian Research Universities. https://u15.ca/members/mcmaster-university/