REVIEW ARTICLE

ARTIFICIAL INTELLIGENCE IN BRONCHOSCOPY FOR PULMONARY DISEASE: A SYSTEMATIC REVIEW OF RANDOMIZED CONTROLLED TRIALS

Eric Daniel Tenda1*, Almerveldy Azaria Dohong1,2, Nadia Iqbal2, Aziza Harris1, Steffi Cong Andi Nata1, Muhammad Aziz Arridho1,4, Alfino Syahputra1,2, Sahat Halim1,5
1Division of Pulmonology and Critical Medicine, Department of Internal Medicine, Faculty of Medicine, Universitas Indonesia, National General Hospital of Dr. Cipto Mangunkusumo, Indonesia.
2Faculty of Medicine, Palangka Raya University, Palangka Raya, Indonesia
3Faculty of Medicine, Pembangunan Nasional Veteran Jakarta University, Jakarta, Indonesia
4Faculty of Medicine, Syiah Kuala University, Banda Aceh, Indonesia
5Faculty of Medicine, Universitas Sumatera Utara, North Sumatra, Indonesia

ABSTRACT

Background: Bronchoscopy is a core diagnostic and therapeutic procedure for pulmonary diseases, yet performance is highly operator dependent, contributing to variability in airway inspection completeness and procedural efficiency. Artificial intelligence (AI) systems have been developed to support real time airway anatomy recognition, navigation, and performance feedback. This systematic review summarizes randomized controlled trial evidence on the impact of AI guided bronchoscopy on procedural performance.
Methods: This review followed PRISMA guidance and was registered in PROSPERO. We searched PubMed/MEDLINE, ScienceDirect, and SAGE Journals from database inception to January 2026. Two reviewers independently screened studies and extracted data. Risk of bias was assessed using the RoB 2 tool. Given heterogeneity in interventions, populations, and outcome definitions, results were synthesized narratively.
Results: From 1,654 records, four randomized controlled trials met inclusion criteria. All studies were conducted in simulation settings and enrolled participants with varying experience (novice to experienced bronchoscopists, including critical care physicians). Overall, AI guided bronchoscopy favored AI supported groups compared with comparators (no AI, written instruction, directed self learning, or expert instruction), showing improvements in key performance domains such as inspection completeness and structured progression, as well as efficiency related metrics including intersegmental time and or total procedure time.
Conclusion: Simulation based RCT evidence supports AI as a promising approach to improve standardization and procedural performance in bronchoscopy training. Multicenter patient based randomized trials are warranted to confirm benefits on clinically meaningful outcomes.
Keywords: artificial intelligence; bronchoscopy; randomized controlled trial; simulation based training; procedural performance.

ABSTRAK

Latar belakang: Bronkoskopi merupakan prosedur diagnostik dan terapeutik penting pada penyakit paru, tetapi keberhasilannya sangat bergantung pada keterampilan operator sehingga menimbulkan variasi kelengkapan inspeksi dan efisiensi prosedur. Kecerdasan buatan (AI) dikembangkan untuk membantu pengenalan anatomi jalan napas, navigasi, dan umpan balik kinerja secara real time. Tinjauan sistematis ini mengevaluasi bukti uji acak terkontrol terkait dampak bronkoskopi berbasis AI pada kinerja prosedural.
Metode: Tinjauan sistematis mengikuti pedoman PRISMA dan terdaftar di PROSPERO. Pencarian dilakukan pada PubMed/MEDLINE, ScienceDirect, dan SAGE Journals sejak awal basis data hingga Januari 2026. Dua penelaah secara independen melakukan seleksi studi dan ekstraksi data. Risiko bias dinilai menggunakan RoB 2. Sintesis dilakukan secara naratif karena heterogenitas luaran dan desain.
Hasil: Dari 1.654 artikel yang teridentifikasi, empat uji acak terkontrol memenuhi kriteria inklusi. Seluruh studi dilakukan pada setting simulasi dengan peserta yang bervariasi (novis hingga berpengalaman, termasuk dokter perawatan intensif). Secara umum, bronkoskopi dengan panduan AI menunjukkan luaran yang menguntungkan dibanding pembanding (tanpa AI, instruksi tertulis, pembelajaran mandiri terarah, atau instruksi ahli), mencakup peningkatan kelengkapan inspeksi dan progresi terstruktur, serta perbaikan metrik efisiensi seperti waktu antarsegi dan/atau waktu prosedur.
Kesimpulan: Bukti RCT berbasis simulasi mendukung AI sebagai alat yang menjanjikan untuk meningkatkan standardisasi dan kinerja bronkoskopi, terutama pada pelatihan. Uji klinis multicenter berbasis pasien diperlukan untuk menilai dampak terhadap luaran klinis.
Kata kunci: kecerdasan buatan; bronkoskopi; uji acak terkontrol; pelatihan simulasi; kinerja prosedural.

Correspondence : Eric Daniel Tenda
Division of Pulmonology and Critical Medicine, Department of Internal Medicine, Faculty of Medicine, Universitas Indonesia, National General Hospital of Dr. Cipto Mangunkusumo, Indonesia. Jln. Diponegoro No. 70. Jakarta, Indonesia.
Email: eric.tenda@ui.ac.id

How to cite this article :

Artificial Intelligence in Bronchoscopy for Pulmonary Disease
https://indonesiajournalchest.com/index.php/IJC/issue/view/293

INTRODUCTION

Bronchoscopy remains a central diagnostic and therapeutic procedure in the evaluation and management of pulmonary diseases, particularly lung cancer. Its effectiveness relies on accurate airway navigation, systematic inspection, and correct interpretation of endoscopic findings, all of which are highly operator-dependent. Variability in technical skill and experience contributes to inconsistent diagnostic completeness and procedural efficiency, creating challenges in both clinical practice and training. In recent years, artificial intelligence (AI) has emerged as a promising adjunct in respiratory medicine, offering data-driven approaches to image interpretation, decision support, and procedural guidance across multiple pulmonary applications[1-3]. Within bronchoscopy, AI systems have been developed to assist with real-time bronchial anatomy recognition, navigation, and performance feedback. Randomized controlled trials have demonstrated that AI-guided bronchoscopy can improve key procedural outcomes, including diagnostic completeness, structured progression through bronchial segments, and procedural efficiency, across different levels of operator experience[4-6]. In parallel, broader reviews of AI integration in bronchoscopy and endobronchial ultrasound (EBUS) highlight rapid technological progress but also emphasize that most clinical evidence remains limited by small sample sizes and simulation-based designs[7] . Emerging advances such as AI-assisted robotic bronchoscopy further illustrate the potential of AI to reduce operator dependency and expand access to high-quality bronchoscopic care, although widespread clinical adoption requires robust validation[8] . Therefore, this systematic review focuses exclusively on randomized controlled trials to evaluate the impact of AI-guided bronchoscopy on procedural performance in the diagnosis and treatment of pulmonary diseases.

METHODS

Search strategy and study selection

This systematic review was designed and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A comprehensive electronic search was conducted in PubMed/MEDLINE, ScienceDirect, and SAGE Journals from database inception to January 2026. The search strategy was constructed to capture studies at the intersection of bronchoscopy and artificial intelligence, combining bronchoscopy-related terms (“bronchoscopy,” “bronchoscopic,” “flexible bronchoscopy,” “bronchonavigation,” “airway navigation”) with AI-related terms (“artificial intelligence,” “machine learning,” “deep learning,” “neural network*,” “computer vision,” “AI-guided,” “AI-assisted”). To restrict retrieval to randomized evidence, the strategy was combined with trial-related filters (“randomized,” “randomised,” “randomized controlled trial,” “RCT,” “crossover”). Search syntax was adapted to the requirements of each database. In addition to database searching, reference lists of all included articles were manually screened to identify potentially eligible trials not captured electronically.

Eligibility criteria were defined a priori. Studies were included if they were randomized controlled trials (parallel-group or crossover) evaluating AI-assisted or AI-guided bronchoscopy in adults, in either simulation-based or clinical settings, where AI was integrated into the bronchoscopy procedure or bronchoscopy training (e.g., real-time guidance, navigation support, structured inspection feedback, or AI-guided mastery learning). Eligible comparators included bronchoscopy without AI assistance, conventional simulator training, written instruction, directed self-learning approaches, or real-time human expert instruction. To be eligible, studies had to report at least one quantitative bronchoscopy performance outcome relevant to procedural completeness, navigation/structure, or efficiency. Studies were excluded if they were non-randomized, observational, or descriptive; if AI was applied only for post-hoc image analysis without procedural integration; if they were pediatric-only; or if only abstracts/editorials/letters were available without full trial data. Duplicate reports were handled by retaining the most complete publication.

All retrieved records were exported, and duplicates were removed prior to screening. Two reviewers independently screened titles and abstracts to identify potentially eligible studies. Full texts were then obtained and independently assessed against the eligibility criteria. Discrepancies at any stage were resolved through discussion until consensus was reached. The study selection process was documented using a PRISMA flow diagram, including reasons for exclusion at the full-text stage.

Search strategy and study selection

Data extraction

A standardized extraction form was developed before data collection to ensure consistency and reproducibility. Two reviewers independently extracted data from each included study. Extracted items included bibliographic details (author, year), study characteristics (design, setting), participant characteristics (number of participants, professional background, and bronchoscopy experience level), and detailed descriptions of the AI intervention (system purpose, real-time vs training-focused integration, feedback/navigation components) and comparator conditions. Outcome data were extracted as reported by the original studies, including outcome definitions, measurement methods, and summary statistics. When studies reported multiple metrics for airway coverage or navigation quality, the outcomes most closely aligned with diagnostic completeness and structured progression were prioritized. Where outcomes were reported at multiple time points or attempts, the prespecified primary endpoint in each trial was extracted preferentially; if not explicitly stated, the final reported assessment within the randomized comparison was used. Any disagreements in extracted data were resolved by consensus, with re-checking of the full text to ensure accuracy.

Study quality assessment

The risk of bias of included randomized trials was assessed independently by two reviewers using the Cochrane Risk of Bias 2 (RoB 2) tool. This tool evaluates bias across five domains: (1) bias arising from the randomization process, (2) bias due to deviations from intended interventions, (3) bias due to missing outcome data, (4) bias in measurement of the outcome, and (5) bias in selection of the reported result. Each domain was judged as low risk, some concerns, or high risk, and an overall risk-of-bias judgment was assigned for each trial. Given that the included studies evaluated AI guidance during bronchoscopy in controlled settings, particular attention was paid to whether outcome assessment relied on objective simulator-derived metrics and whether attrition or incomplete outcome reporting could have influenced results. Any disagreements in domain-level judgments were resolved through discussion to reach consensus. The risk-of-bias assessments were summarized in a dedicated table to facilitate interpretation of study credibility.

Outcomes

The outcomes of interest were defined a priori based on performance and efficiency domains relevant to bronchoscopy. The primary outcomes were diagnostic completeness (reflecting airway coverage or the proportion/number of bronchial segments inspected), structured progression (reflecting systematic navigation and reduced disorganized or repetitive inspection), and total procedure time. Secondary outcomes included intersegmental time (including median intersegmental time where reported), measures of procedural efficiency such as unnecessary segment revisits, and training-related performance outcomes reflecting learning efficiency and attainment of competence. Cognitive workload measures (e.g., NASA-TLX or comparable metrics) were extracted when reported.

Due to heterogeneity across trials in AI system characteristics, operator experience levels, and outcome definitions, results were synthesized narratively. Findings were summarized by study design and AI use case (real-time procedural guidance vs AI-guided training/mastery learning), and the direction and consistency of effects across trials were emphasized. Quantitative pooling was not undertaken unless outcomes were sufficiently homogeneous in definition and measurement; where pooling was not appropriate, effect direction and statistical significance as reported in individual trials were presented descriptively.

Data Synthesis

Given heterogeneity in AI systems, study designs, participant experience levels, and outcome definitions, results were synthesized narratively. Quantitative pooling was not performed because outcomes were not sufficiently comparable across trials. Findings were summarized according to AI application type (real-time procedural guidance vs AI-guided training/mastery learning) and comparator, emphasizing prespecified performance domains including diagnostic completeness, structured progression, procedure time, and intersegmental time, as well as efficiency indicators such as segment revisits.

RESULT

Study Selection

A total of 1,654 records were identified from PubMed, ScienceDirect, SAGE, and SpringerLink. After duplicate removal and application of filters, 279 records remained for screening. Following title/abstract screening, 27 articles were selected for full-text review. Of these, 23 studies were excluded due to ineligibility at full-text assessment, leaving four randomized controlled trials for inclusion in the final qualitative synthesis. No studies were excluded after critical appraisal (0 excluded after JBI appraisal). The study selection process is presented in **Figure 1** (PRISMA flow diagram).

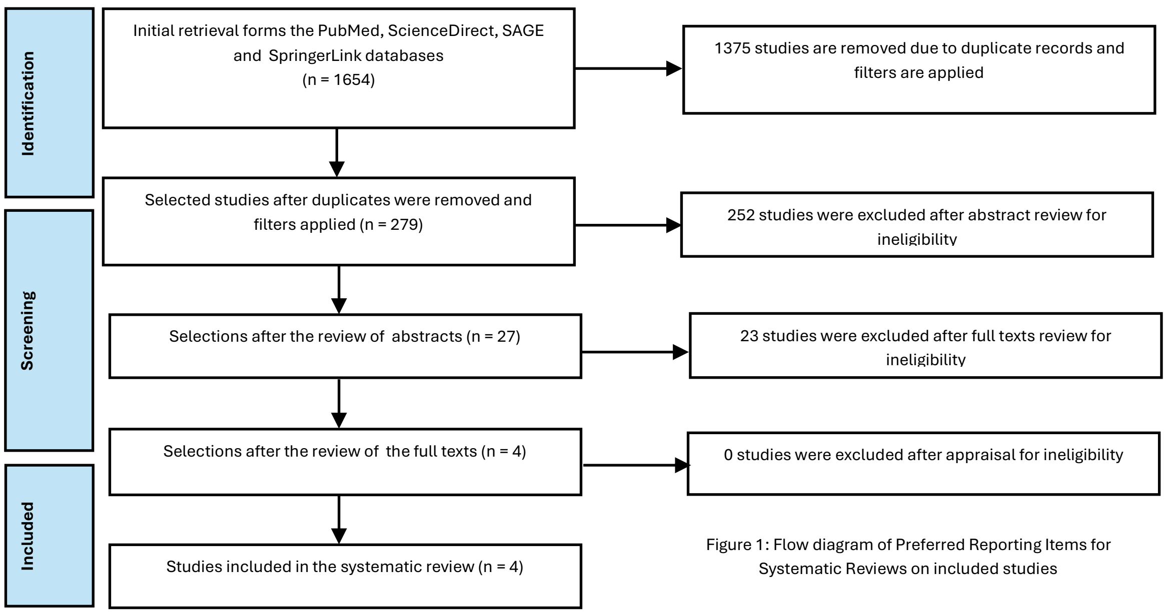

Figure 1: Flow diagram of Preferred Reporting Items for Systematic Reviews on included studies

Study characteristics

The characteristics of included randomized controlled trials are summarized in Table 1. In total, four RCTs were included, published between 2024 and 2025. One study employed a randomized crossover design, while the remaining trials used parallel-group designs. All trials were conducted in simulation-based bronchoscopy settings, reflecting an educational and performance-assessment context rather than patient-based clinical bronchoscopy. Sample sizes ranged from 20 to 101 participants, indicating variation in statistical power and precision across studies.

Across trials, participant profiles differed substantially in baseline bronchoscopy experience. Two studies focused on novice bronchoscopists undergoing structured bronchoscopy training, one trial enrolled a mixed cohort spanning novice, intermediate, and experienced bronchoscopists, and one study evaluated critical-care physicians with variable bronchoscopy exposure. Despite these differences, all trials examined AI systems integrated into bronchoscopy performance or training with the aim of improving procedural standardization. The AI interventions generally provided real-time guidance and/or feedback to support airway navigation, bronchial anatomy recognition, and structured progression through bronchial segments. Comparators varied across studies and included bronchoscopy without AI guidance, conventional training with written instruction, directed self-regulated learning, and real-time human expert instruction.

Outcome selection was broadly aligned across trials around procedural performance and efficiency metrics. The most frequently assessed outcomes were diagnostic completeness, structured progression, and procedure time, with some trials additionally reporting intersegmental time and efficiency indicators such as segment revisits. This heterogeneity in study populations, intervention implementation, and outcome definitions supported the use of narrative synthesis rather than quantitative pooling.

Author (year)	**Study design**	**Setting**	**Sample size**	**Participant type / experience level**	**AI intervention (short description)**	**Comparator**	**Primary outcomes assessed**
Cold et al., 2025 (randomized crossover trial)	Randomized crossover RCT	Simulation (high-fidelity bronchoscopy simulator)	101	Mixed experience: novice, intermediate, and experienced bronchoscopists	AI-guided bronchoscopy system providing real-time anatomical recognition, structured airway navigation prompts, and performance feedback during bronchoscope advancement	Standard simulator-based bronchoscopy without AI guidance	Diagnostic completeness, structured progression, intersegmental time, total procedure time

Cold et al., 2024	Parallel-group RCT	Simulation (bronchoscopy simulator)	20	Novice bronchoscopists with limited prior bronchoscopy experience	Real-time AI feedback system guiding systematic airway inspection and segmental progression during training	Conventional training using written instructional material without AI assistance	Diagnostic completeness, structured progression score, total procedure time
Agbontaen et al., 2025	Parallel-group RCT	Simulation (bronchoscopy simulator)	40	Critical-care physicians with varying levels of bronchoscopy experience	AI-guided bronchoscopy providing continuous navigational assistance and feedback to optimize airway inspection and minimize unnecessary segment revisits	Real-time instruction and feedback from a human bronchoscopy expert	Procedure time, intersegmental time, number of bronchial segment revisits
Cold et al., 2025 (mastery learning RCT)	Parallel-group RCT	Simulation (bronchoscopy simulator)	24	Novice bronchoscopists undergoing structured training	AI-guided mastery learning framework with automated assessment and feedback, requiring achievement of predefined performance criteria	Directed self-regulated learning without AI guidance	Procedure time to task completion, diagnostic completeness

Table.1
Study characteristics

Summary of outcomes

Outcome findings are summarized in Table 2. Across the included RCTs, AI-guided bronchoscopy consistently favored AI assistance for procedural performance and efficiency. In the randomized crossover trial, AI guidance resulted in higher diagnostic completeness and better structured progression, with shorter intersegmental time, while total procedure time was reported as slightly longer, interpreted as reflecting more thorough airway inspection ($p < 0.001$ for completeness, progression, and intersegmental time). Among novice operators, AI-assisted training improved diagnostic completeness and structured progression and reduced procedure time compared with written instruction ($p < 0.01$). When compared directly with human expert instruction, AI guidance demonstrated superior efficiency metrics, including shorter procedure time, reduced intersegmental time, and fewer segment revisits ($p < 0.05$). In an AI-guided mastery learning framework, trainees achieved comparable diagnostic completeness with significantly faster task completion ($p < 0.001$).

Outcome findings across the included randomized controlled trials are summarized in Table 2. Across studies, reported endpoints primarily included diagnostic completeness (DC), structured progression (SP), procedure time (PT), and intersegmental time (including mean/median intersegmental time, depending on the trial), with some trials additionally reporting segment revisits and learner-related measures.

In the novice parallel-group RCT comparing AI-based real-time feedback with written instruction, the AI group demonstrated higher performance on the final assessment: DC was higher (median 18 vs 14.5 segments inspected; $p < 0.001$) and SP was higher (median 16.5 vs 3 progressions; $p < 0.001$). The AI group also had a shorter PT (median 217.5 s vs 431.5 s; $p = 0.002$). In addition, training-related measures favored the AI group, including longer training exposure (difference 49 minutes; $p = 0.029$) and higher intrinsic motivation (IMI total score difference 11; $p = 0.001$).

In the randomized crossover trial including operators with mixed experience levels, AI guidance was associated with higher airway inspection and navigation structure compared with procedures performed without AI by the same participants. DC increased by a mean of +6.0 inspected segments ($p < 0.001$) and SP increased by +5.2 progressions ($p < 0.001$). Time-based metrics differed by outcome: mean intersegmental time decreased by −10.3 s ($p < 0.001$) while total PT increased by +72 s ($p < 0.001$). Subgroup analyses reported statistically significant improvements in DC and SP across experience strata, with the largest magnitude among novice participants.

In the parallel-group RCT comparing AI-guided bronchoscopy with real-time human expert instruction among critical-care physicians, both arms demonstrated improvements from pre- to post-training across multiple metrics. In the final post-training comparison, efficiency outcomes favored AI guidance: intersegmental time was lower in the AI group (16.5 vs 23.9; $p = 0.049$) and PT was shorter (264 vs 338; $p = 0.039$). Between-group differences were not statistically significant for DC (16 vs 15.5; $p = 0.641$) or SP (6 vs 5.5; $p = 0.284$). Segment revisits were numerically lower with AI but not statistically different (7 vs 11.5; $p = 0.245$). NASA-TLX workload scores did not differ significantly between groups.

In the mastery learning RCT comparing AI-guided mastery learning with directed self-regulated learning, the AI-guided mastery learning group reached completion faster (PT 107 s vs 180 s; $p < 0.001$). Diagnostic completeness was reported as comparable between groups at the assessment endpoint.

Author (year)	Outcome domain	Outcome metric	AI group result	Comparator group result	Direction of effect	Statistical significance	Key interpretation
Cold et al., 2025 (randomized crossover trial)	Procedural performance	Diagnostic completeness, structured progression, intersegmental time, procedure time	Higher diagnostic completeness and structured progression; shorter intersegmental time; slightly longer total procedure time	Lower completeness and structured progression; longer intersegmental time; shorter total procedure time	Favors AI	p < 0.001 for completeness, progression, and intersegmental time	AI guidance improved systematic airway inspection and efficiency, with longer procedure time reflecting more thorough examination
Cold et al., 2024	Training performance	Diagnostic completeness, structured progression, procedure time	Significantly higher completeness and structured progression; shorter procedure time	Lower completeness and progression; longer procedure time	Favors AI	p < 0.01 for all primary outcomes	AI-based feedback enhanced learning efficiency and accelerated skill acquisition among novices
Agbontaen et al., 2025	Procedural efficiency	Procedure time, intersegmental time, segment revisits	Shorter procedure time; reduced intersegmental time; fewer segment revisits	Longer procedure time; more revisits	Favors AI	p < 0.05 for efficiency outcomes	AI guidance matched or exceeded human expert instruction in efficiency-related metrics
Cold et al., 2025 (mastery learning RCT)	Training efficiency	Procedure time, diagnostic completeness	Faster task completion with comparable diagnostic completeness	Slower task completion with comparable diagnostic completeness	Favors AI	p < 0.001 for procedure time	AI-guided mastery learning achieved equivalent competence in significantly less time

Table.2 Summary of outcomes of included randomized controlled trials of AI-guided bronchoscopy

Quality assessment

Methodological quality was assessed using the Cochrane Risk of Bias 2 (RoB 2) tool, and domain-level judgments are summarized in Figure 2. Overall, two trials were judged to have low risk of bias across all domains (Cold et al., 2025 randomized crossover trial; Cold et al., 2025 mastery learning RCT). The remaining two trials (Cold et al., 2024; Agbontaen et al., 2025) were rated as having some concerns overall, driven solely by Domain 2 (deviations from intended interventions). In contrast, Domain 1 (randomization process), Domain 3 (missing outcome data), Domain 4 (measurement of outcomes), and Domain 5 (selection of the reported result) were consistently judged as low risk across all included studies.

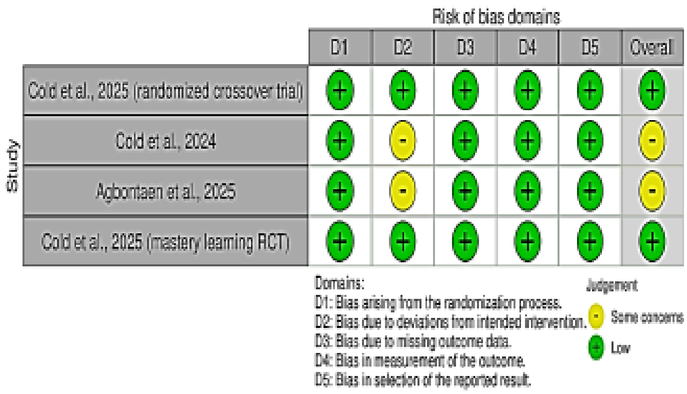

Figure. 2 Risk of bias assessment of included randomized controlled trials using the RoB 2 tool

Domains: D1, bias arising from the randomization process; D2, bias due to deviations from intended interventions; D3, bias due to missing outcome data; D4, bias in measurement of the outcome; D5, bias in selection of the reported result. Green indicates low risk of bias and yellow indicates some concerns.

Protocol and Registration

This systematic review protocol was registered in PROSPERO (International Prospective Register of Systematic Reviews) under the registration number CRD420261283454 (registered 12 January 2026).

DISCUSSION

Principal Findings

This systematic review synthesizes randomized controlled trial evidence on AI supported bronchoscopy and shows a consistent direction of benefit in simulation based settings, particularly for bronchoscopy training. Across trials, AI guided systems improved procedural execution domains that are central to bronchoscopy competence, including systematic airway survey behavior, navigational structure, and efficiency related performance.[1–3] The additional bronchoscopy focused review material provided also supports that AI applications in bronchoscopy and EBUS are expanding rapidly, with training and procedural support representing one of the most immediately actionable pathways for adoption.[4]

A granular analysis of the randomized crossover trial by Cold et al. (2025) underscores that these benefits are not confined to the novice learner. In a cohort of 101 participants ranging from complete novices to expert bronchoscopists performing on a high-fidelity phantom, AI guidance resulted in a mean increase of 6.0 inspected segments and 5.2 structured progressions compared to standard bronchoscopy. Crucially, while the magnitude of improvement was most pronounced in novices, statistically significant gains were observed across all experience levels. This challenges the prevailing assumption of a performance "ceiling" in expert bronchoscopy and suggests that human operators, regardless of tenure, remain susceptible to navigational drift and omission errors that AI augmentation effectively mitigates.[2]

Furthermore, the data reveal a complex relationship between AI guidance and procedural time. In some contexts, total procedure time increased with AI assistance, a finding attributed to the system enforcing a more thorough and rigorous inspection of the bronchial tree than the operator would naturally perform. However, measures of efficiency, specifically the mean intersegmental time, consistently favored the AI-guided groups. This paradox (longer total time but faster interval navigation) indicates that AI reduces the "hunting" behavior and disorientation associated with identifying specific segmental bronchi, thereby transforming the procedure from a rapid but potentially incomplete survey into a deliberate, comprehensive examination.[3]

Beyond the simulation laboratory, the clinical evidence base is rapidly maturing. While the RELIANT trial (2025) indicated statistical non-inferiority between robotic-assisted bronchoscopy and electromagnetic navigation bronchoscopy regarding diagnostic yield in expert hands, emerging data presented at the European Respiratory Society Congress 2025 suggests a different narrative when advanced imaging is integrated. The randomized trial by Steinack et al. demonstrated that robotic-assisted bronchoscopy combined with cone-beam computed tomography achieved a diagnostic yield of approximately 85-89%, significantly outperforming conventional fluoroscopic techniques. This discrepancy highlights that the "principal finding" of the current era is not merely the superiority of one robotic platform over another, but the synergistic efficacy of AI-driven navigation coupled with real-time, AI-enhanced tomographic confirmation. Thus, AI is transitioning from a theoretical educational adjunct to a critical component of the high-precision diagnostic ecosystem, essential for bridging the gap between operator capability and the increasing complexity of peripheral pulmonary nodule management.[5,6]

How AI may improve bronchoscopy performance

A unifying explanation across the RCTs is that AI delivers immediate, standardized, and objective feedback during airway navigation and inspection, which may reduce variability in how operators progress through segments and how consistently they maintain a structured survey.[1–3] This training focused mechanism aligns with broader respiratory AI reviews that emphasize AI's strengths in standardization, pattern recognition, and workflow support, provided that the tool is usable and integrates into practice without increasing burden.[7–9]

AI systems that utilize deep learning algorithms, specifically convolutional neural networks and transformers, are now capable of performing semantic segmentation of the airway in real-time. These models differentiate between the carina, mainstem bronchi, and lobar segments with accuracy often exceeding that of human experts. For instance, the Multiscale Attention Residual Network proposed by Sun et al. (2024) utilizes a multiscale convolutional block attention module to extract spatial and channel features, enabling precise focus on lesion regions and anatomical landmarks.1 By overlaying this information onto the clinical video feed (essentially creating an augmented reality environment) the AI offloads the navigational processing from the physician. This allows the operator to reallocate cognitive resources toward fine motor control, mucosal inspection, and clinical decision-making.[3,10]

The improvement in mean intersegmental time observed in the Agbontaen et al. (2025) trial provides further insight into the mechanism of improvement. Mean intersegmental time is a surrogate for navigational efficiency; a lower mean intersegmental time implies less time spent orienting or correcting wrong turns between segments. The AI system provides a continuous "you are here" reference relative to the bronchial tree, minimizing the need for withdrawal and re-orientation. Additionally, newer AI models are incorporating "wall collision avoidance" and centerline constraints, predicting the optimal path to a target and visually guiding the operator to keep the scope centered in the lumen.This creates a "virtual rail" for the bronchoscopist, reducing mucosal trauma and improving the stability of the platform during critical maneuvers like biopsy. Thus, AI improves performance not merely by "grading" the operator post-hoc, but by actively augmenting their sensory perception and spatial awareness during the procedure, effectively functioning as a digital co-pilot that is immune to fatigue, distraction, or the degradation of spatial memory.[3,11,12]

Educational implications and competency based training

The current RCT evidence base is most mature in bronchoscopy education rather than patient outcome evaluation.[1–3] This is consistent with implementation oriented literature in health care AI, where early real world gains often occur in structured, measurable tasks and controlled deployment contexts such as training pipelines.[13] Within bronchoscopy training, AI support may be particularly well suited for competency based frameworks, because repeated objective feedback can be mapped to proficiency targets and reduce dependence on variable instructor availability.[2,7,9] This direction is also coherent with respiratory AI discussions highlighting that scalability and consistency are key practical advantages when compared with human dependent feedback.[7–9]

AI guidance compared with expert instruction

Direct comparison between AI guidance and human expert instruction is an important step because it positions AI as an instructional modality rather than only an add-on tool. The RCT comparing AI guided bronchoscopy training with expert instruction suggests that AI can deliver at least comparable training benefit for selected performance domains, especially those that are structured and repeatable.[3] In practical terms, this supports a division of roles model, where AI reinforces systematic survey behaviors and efficiency mechanics, while expert faculty focus on higher level judgement, troubleshooting, and clinical decision making.[3,4]

Positioning within advanced bronchoscopy technology

Bronchoscopy practice is evolving alongside navigation bronchoscopy and robotic assisted bronchoscopy systems, with evidence syntheses focusing on diagnostic yield and safety for peripheral pulmonary lesions. These clinically oriented outcomes are different from the simulator based performance endpoints in the included RCTs, but they provide context for why procedural standardization and operator reliability remain high value targets. As procedures become more technologically layered, the potential role of AI guidance as a complementary support layer deserves direct evaluation in clinical trials, particularly to determine incremental benefit when combined with navigation or robotic platforms.[14–16]

Toward assisted navigation and robotics

The AI co-pilot bronchoscope robot work illustrates the field's trajectory from feedback systems toward more active assistance and shared control navigation, which could change both training paradigms and procedural delivery models. However, increasing autonomy heightens the need for rigorous safety evaluation, human oversight design, and reporting of failure modes, particularly when moving beyond simulation.[17]

Translation and implementation considerations

The broader AI literature in respiratory care and in health care implementation repeatedly identifies a translation gap between promising performance in controlled settings and durable benefit in routine practice.[7–9,13]. Scoping evidence in acute respiratory failure decision support further emphasizes that clinically validated tools remain limited and that reporting of evaluation quality, human factors, adherence, and safety related considerations is inconsistent, which can slow deployment even when models appear promising.[18] These implementation concerns are directly relevant to bronchoscopy AI, especially if guidance tools are expanded from training into live patient procedures or used in high stakes competency assessment pathways.[13,18]

Relationship to respiratory AI in adjacent domains

Evidence from lung imaging and AI in ARDS shows how respiratory AI development often progresses from feasibility and prediction tasks toward clinically integrated workflows, reinforcing that validation, usability, and appropriate outcome selection are essential for translation.[19] For bronchoscopy AI, this supports prioritizing patient based trials that connect improved procedural execution with clinically meaningful endpoints rather than relying only on surrogate simulator metrics.[14–16]

Limitations of the current evidence base

Several limitations constrain interpretation. First, the RCT evidence is predominantly simulation based, which strengthens internal control but limits generalizability to clinical bronchoscopy where anatomy variability, secretions, bleeding, sedation, and time pressure may influence performance and tool interaction. Second, heterogeneity exists across intervention designs, participant experience levels, and outcome definitions, which complicates cross study comparison and limits quantitative pooling. Third, blinding is difficult in trials of guidance tools and behavior may be influenced by awareness of the intervention, even when outcomes are objectively captured by simulators.[2,7,9]

Future directions

Next step research should focus on pragmatic clinical RCTs evaluating whether AI guided bronchoscopy improves outcomes that matter in practice, including diagnostic yield, complication rates, time to diagnosis, and workflow efficiency, and should also evaluate performance across diverse operator skill levels and clinical contexts.[4,14–16] Studies should additionally report implementation and human factors measures, including usability, adherence to guidance, and error handling, consistent with recommendations emerging from respiratory decision support evaluation literature.[13,18] Finally, comparative studies should clarify incremental value when AI guidance is layered onto navigation or robotic bronchoscopy platforms.[14–17]

CONCLUSION

In this systematic review of randomized controlled trials, AI guided bronchoscopy consistently improved simulator based bronchoscopy performance, particularly in training contexts. AI support was associated with improved structured airway navigation, more complete airway inspection, and favorable efficiency related metrics compared with conventional instruction, self regulated learning, or expert tutoring. Although the current evidence is largely derived from simulation settings and heterogeneity in outcome definitions limits quantitative pooling, these findings support AI as a promising approach to standardize bronchoscopy training and reduce performance variability. Future work should prioritize multicenter clinical randomized trials using harmonized outcome measures and implementation aware evaluation to determine whether these performance gains translate into improved diagnostic yield, safety, and workflow efficiency in routine bronchoscopy practice.[2,7,9]

REFERENCES

1. Cold KM, Xie S, Nielsen AO, Clementsen PF, Konge L. Artificial Intelligence Improves Novices' Bronchoscopy Performance. CHEST. 2024 Feb;165(2):405–13.
2. Cold KM, Agbontaen K, Nielsen AO, Andersen CS, Singh S, Konge L. Artificial intelligence improves bronchoscopy performance: a randomised crossover trial. ERJ Open Res. 2025 Jan;11(1):00395–2024.
3. Agbontaen KO, Cold KM, Woods D, Grover V, Aboumarie HS, Kaul S, et al. Artificial Intelligence-Guided Bronchoscopy is Superior to Human Expert Instruction for the Performance of Critical-Care Physicians: A Randomized Controlled Trial. Crit Care Med. 2025 May;53(5):e1105–15.

4. Winiarski S, Radziszewski M, Wiśniewski M, Cisek J, Wąsowski D, Plewczyński D, et al. Integrating Artificial Intelligence in Bronchoscopy and Endobronchial Ultrasound (EBUS) for Lung Cancer Diagnosis and Staging: A Comprehensive Review. Cancers. 2025 Aug 29;17(17):2835.
5. Paez R, Lentz RJ, Duke JD, Siemann JK, Salmon C, Dahlberg GJ, et al. Robotic versus Electromagnetic Bronchoscopy for Peripheral Pulmonary Lesions: A Randomized Trial (RELIANT). Am J Respir Crit Care Med. 2025 Sept;211(9):1644–51.
6. Steinack C, Baumgartner P, Monsch GM, Schneiter D, Oh S, Theano Samara E, et al. Robotic-Assisted Bronchoscopy with Integrated Cone-Beam CT vs. Conventional Bronchoscopy for Diagnosing Peripheral Pulmonary Lesions: An Open-Label Randomized Controlled Trial. In: Interventional pulmonology [Internet]. European Respiratory Society; 2025 [cited 2026 Jan 16]. p. OA1178. Available from: https://publications.ersnet.org/lookup/doi/10.1183/13993003.congress-2025.OA1178
7. Al-Anazi S, Al-Omari A, Alanazi S, Marar A, Asad M, Alawaji F, et al. Artificial intelligence in respiratory care: Current scenario and future perspective. Ann Thorac Med. 2024 Apr;19(2):117–30.
8. Karthika M, Sreedharan JK, Shevade M, Mathew CS, Ray S. Artificial intelligence in respiratory care. Front Digit Health. 2024 Dec 23;6:1502434.
9. Kalaiyarasan K, Sridhar R. Artificial Intelligence in Respiratory Medicine: The Journey So Far – A review. J Assoc Pulmonologist Tamil Nadu. 2023 May;6(2):53–68.
10. Sun W, Yan P, Li M, Li X, Jiang Y, Luo H, et al. An accurate prediction for respiratory diseases using deep learning on bronchoscopy diagnosis images. J Adv Res. 2025 Oct;76:423–38.
11. Cold KM, Vamadevan A, Laursen CB, Bjerrum F, Singh S, Konge L. Artificial intelligence in bronchoscopy: a systematic review. Eur Respir Rev. 2025 Apr;34(176):240274.
12. Shu H, Soberanis-Mukul RD, Xu J, Ding H, Ringel M, Shen M, et al. BronchOpt : Vision-Based Pose Optimization with Fine-Tuned Foundation Models for Accurate Bronchoscopy Navigation [Internet]. arXiv; 2025 [cited 2026 Jan 16]. Available from: https://arxiv.org/abs/2511.09443
13. Sharma M, Savage C, Nair M, Larsson I, Svedberg P, Nygren JM. Artificial Intelligence Applications in Health Care Practice: Scoping Review. J Med Internet Res. 2022 Oct 5;24(10):e40238.
14. Kops SEP, Heus P, Korevaar DA, Damen JAA, Idema DL, Verhoeven RLJ, et al. Diagnostic yield and safety of navigation bronchoscopy: A systematic review and meta-analysis. Lung Cancer. 2023 June;180:107196.
15. Ali MS, Ghori UK, Wayne MT, Shostak E, De Cardenas J. Diagnostic Performance and Safety Profile of Robotic-assisted Bronchoscopy: A Systematic Review and Meta-Analysis. Ann Am Thorac Soc. 2023 Dec;20(12):1801–12.
16. Zhang C, Xie F, Li R, Cui N, Herth FJF, Sun J. Robotic-assisted bronchoscopy for the diagnosis of peripheral pulmonary lesions: A systematic review and meta-analysis. Thorac Cancer. 2024 Mar;15(7):505–12.
17. Zhang J, Liu L, Xiang P, Fang Q, Nie X, Ma H, et al. AI co-pilot bronchoscope robot. Nat Commun. 2024 Jan 4;15(1):241.
18. Gupta P, Pearce AK, Pham T, Miller M, Brunetti K, Heskett K, et al. Artificial intelligence-driven decision support for patients with acute respiratory failure: a scoping review. Intensive Care Med Exp. 2025 Aug 8;13(1):83.
19. Chiumello D, Coppola S, Catozzi G, Danzo F, Santus P, Radovanovic D. Lung