EISSN 2979-9139

Official Journal of
TURKISH
THORACIC SOCIETY

Thoracic Research & Practice

Formerly Turkish Thoracic Journal

galenos Publishing House

Volume 27 Issue 4 July 2026
thoracrespract.org

Thoracic Research & Practice

Official Journal of the Turkish Thoracic Society

Publisher Contact
Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1 34093 İstanbul, Türkiye
Phone: +90 (530) 177 30 97
E-mail: info@galenos.com.tr **Web:** www.galenos.com.tr
Publisher Certificate Number: 14521

Online Publication Date: July 2026
E-ISSN: 2979-9139
International scientific journal published bimonthly.

Thoracic Research & Practice

Official Journal of the Turkish Thoracic Society

Thoracic Research & Practice
Official Journal of the Turkish Thoracic Society

ABOUT

About the Thoracic Research and Practice

Thoracic Research and Practice is a peer reviewed, open access, online-only journal published by the Turkish Thoracic Society.

Thoracic Research and Practice is a bimonthly journal that is published in English in January, March, May, July, September, and November.

Journal History

Thoracic Research and Practice started its publication life following the merger of two journals which were published under the titles "Turkish Respiratory Journal" and "Toraks Journal" until 2008. From 2008 to 2022, the journal was published under the title "Turkish Thoracic Journal". Archives of the journals were transferred to Thoracic Research and Practice.

Abstracting and Indexing

Thoracic Research and Practice is covered in the following abstracting and indexing databases; PubMed Central, Web of Science - Emerging Sources Citation Index, Scopus, EMBASE, EBSCO, CINAHL, Gale/Cengage Learning, ProQuest, DOAJ, CNKI, TUBITAK ULAKBIM TR Index.

Aims, Scope, and Audience

Thoracic Research and Practice aims to publish studies of the highest scientific and clinical value, and encourages the submission of high-quality research that advances the understanding and treatment of pulmonary diseases.

Thoracic Research and Practice covers a wide range of topics related to adult and pediatric pulmonary diseases, as well as thoracic imaging, environmental and occupational disorders, intensive care, sleep disorders and thoracic surgery, including diagnostic methods, treatment techniques, and prevention strategies. The journal is interested in publishing original research that addresses important clinical questions and advances the understanding and treatment of these conditions. This may include studies on the effectiveness of different treatments, new diagnostic tools or techniques, and novel approaches to preventing or managing pulmonary diseases.

Thoracic Research and Practice publishes clinical and basic research articles, reviews, statements of agreement or disagreement on controversial issues, national and international consensus reports, systematic reviews and meta-analyses, letters to the editor and editorials. Conference proceedings may also be considered for publication.

The target audience of the journal includes healthcare professionals and researchers who are interested in or working in the pulmonary diseases field, and related disciplines.

Open Access Policy

Thoracic Research and Practice is an open access publication.

All published content is available online, free of charge at thoracrespract.org.

Thoracic Research & Practice
Official Journal of the Turkish Thoracic Society

Copyright Policy

A Copyright Agreement and Acknowledgement of Authorship form should be submitted with all manuscripts. By signing this form, authors transfer the copyright of their work to the Turkish Thoracic Society and agree that the article, if accepted for publication by the Thoracic Research and Practice will be licensed under a Creative Commons Attribution-Non Commercial 4.0 International License (CC BY-NC 4.0) which permits third parties to share and adapt the content for non-commercial purposes by giving the appropriate credit to the original work.

When using previously published content, including figures, tables, or any other material in both print and electronic formats, authors must obtain permission from the copyright holder. Legal, financial and criminal liabilities in this regard belong to the author(s).

Authors retain the copyright of their published work in the Thoracic Research and Practice.

Publication Fee Policy

Thoracic Research and Practice is funded by the Turkish Thoracic Society. Authors are not required to pay any fees during the evaluation and publication process.

You can find the current version of the Instructions to Authors at https://thoracrespract.org/instructions-to-authors.

Türk Toraks Derneği
Turkish Thoracic Society

Thoracic Research & Practice

Official Journal of the Turkish Thoracic Society

CONTENTS

Thoracic Research & Practice
Official Journal of the Turkish Thoracic Society

LETTERS TO THE EDITORS

Thorac Res Pract. 2026;27(4):198-199 **DOI:** 10.4274/ThoracResPract.2026.2026.100726

Editorial

Artificial Intelligence and Accountability in Scientific Writing: The Author's Responsibility

Metin Akgün

Department of Pulmonary Medicine, Ağrı İbrahim Çeçen University Faculty of Medicine, Ağrı, Türkiye

Cite this article as: Akgün M. Artificial intelligence and accountability in scientific writing: the author's responsibility. *Thorac Res Pract.* 2026;27(4):198-199

Received: 17.07.2026 ***Revision Requested:*** 20.07.2026

Accepted: 20.07.2026 ***Publication Date:*** 24.07.2026

Generative artificial intelligence (AI) can now draft scientific text, not merely polish it. That single capacity has shifted the debate from convenience to a harder question: when a machine helps produce a manuscript, who answers for it?

The concern is not abstract. Large language models generate output from statistical patterns in language rather than from verified fact, so nonexistent studies, fabricated authors, and distorted claims can enter the literature. The scale is now documented. An audit of more than 2.5 million biomedical papers in the PubMed Central open-access corpus found that, by early 2026, roughly 1 in 277 papers cited a study that does not exist—a more than 12-fold rise since 2023.[1] This is a measurable harm to the reliability of the published record, not a hypothetical risk.

The obvious countermeasure—AI text detectors—is unreliable. These tools are biased against researchers who are not native English speakers, misclassifying much of their original writing as machine-generated; in one evaluation, detectors labeled more than half of essays by non-native writers as AI-produced.[2] Detector output cannot stand as sole evidence. The remedy therefore does not lie in better surveillance but in placing responsibility where it belongs: with the human author.

That is where the real question sits. When the sentences, the arguments, and even the structure of a paper are shaped by AI, is the human still its author? The classic ship-of-Theseus paradox frames the problem: what matters is not how many parts are replaced, but which parts, and by whom.[3] Superficial language editing and the wholesale reconstruction of an argument are ethically different acts. The researcher who poses the question, makes the scientific choices, develops the original claim, and vouches for every result is the author. AI does not assume that intellectual responsibility.

Editors and publishers are converging on this principle. In its January 2026 update, the International Committee of Medical Journal Editors addressed AI in a dedicated section: authors must disclose the tool and its purpose in both the cover letter and the manuscript; they bear full responsibility for the accuracy, originality, citations, and potential bias of any AI-assisted output; AI must not be listed as an author or cited as a primary source; and editors and reviewers must not upload confidential submissions to general-purpose AI tools.[4] The Committee on Publication Ethics states plainly that AI cannot be an author because it cannot take responsibility for the work,[5] and the World Association of Medical Editors offers a comparable framework.[6] Publishers translate these principles into practice—Elsevier, for example, holds that AI is no substitute for human critical thinking, requires authors to verify all sources, and mandates a disclosure statement before the reference list.[7]

Disclosure, however, must do more than name a tool. A practical framework recommends specifying five elements: the content affected, the action taken, the tool and its version, the purpose of use, and how human oversight was exercised.[8] National science agencies are formalizing the same expectation, as in Türkiye's TÜBİTAK guidance on the responsible use of generative AI (version 04, January 2026).[9] A single sentence often suffices: the language of a defined section was

Corresponding author: Metin Akgün, MD, e-mail: akgunm@gmail.com

reviewed with a named tool and version; every suggestion was assessed by the authors; all scientific claims and citations were verified against original sources; and the authors accept full responsibility for the final content.

Used this way, AI is not something to exclude from scholarship but something to use well. Effective use begins with a clearly specified prompt, favors a sectioned rather than all-at-once workflow, matches the right tool to each stage, and—above all—subjects every output to human verification. The discipline is verification, not delegation. Looking ahead, AI-assisted "living" evidence synthesis—systematic reviews updated automatically and published on the web—promises more current evidence, but will make transparent reporting standards, such as the PRISMA extension for living systematic reviews, more necessary rather than less.[10] As speed increases, the need for oversight deepens.

The question, then, is not whether to use AI but how. The technology is already reshaping academic work, from helping non-native English speakers cross the language barrier to assisting statistical analysis, and those who ignore it will be increasingly disadvantaged. But the price of that transformation must not be the loss of original perspective, critical thought, and accountability to standardized output. AI is a powerful instrument used under human supervision, not a substitute for the author. However much a text is shaped by the tool, the person who takes responsibility—and signs it—remains the author.

REFERENCES

1. Topaz M, Roguin N, Gupta P, Zhang Z, Peltonen LM. Fabricated citations: an audit across 2·5 million biomedical papers. *Lancet*. 2026;407(10541):1779-1781. Erratum in: *Lancet*. 2026;408(10551):218. **[Crossref]**
2. Liang W, Yuksekgonul M, Mao Y, Wu E, Zou J. GPT detectors are biased against non-native English writers. *Patterns (N Y)*. 2023;4(7):100779. **[Crossref]**
3. Akgün M. Author or prompter? Scientific writing, identity, and the Theseus paradox. *Philos Ethics Humanit Med*. 2025;20(1):29. **[Crossref]**
4. International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals [Internet]. Updated January 2026. Accessed July 20, 2026. **[Crossref]**
5. Committee on Publication Ethics. Authorship and AI tools: COPE position statement [Internet]. 2023. Accessed July 20, 2026. **[Crossref]**
6. Zielinski C, Winker MA, Aggarwal R, et al. Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. *Colomb Med (Cali)*. 2023;54(3):e1015868. **[Crossref]**
7. Elsevier. Generative AI policies for journals [Internet]. Updated June 2026. Accessed July 20, 2026. **[Crossref]**
8. Centers for Disease Control and Prevention. Considerations for disclosing generative AI use in scientific work [Internet]. Updated May 28, 2026. Accessed July 20, 2026. **[Crossref]**
9. TÜBİTAK. Destek Süreçlerinde Üretken Yapay Zekânın (ÜYZ) Sorumlu ve Güvenilir Kullanımı Rehberi [Guide on the responsible and trustworthy use of generative AI in support processes]. Version 04 [Internet]. Ankara: TÜBİTAK; January 2026. Accessed July 20, 2026. **[Crossref]**
10. Akl EA, Khabsa J, Iannizzi C, et al.; PRISMA-LSR Group. Extension of the PRISMA 2020 statement for living systematic reviews (PRISMA-LSR): checklist and explanation. *BMJ*. 2024;387:e079183. **[Crossref]**

Thorac Res Pract. 2026;27(4):200-206 **DOI:** 10.4274/ThoracResPract.2026.2025-8-6

Original Article

Diagnostic Value of Non-contrast Lung CT in Evaluating Pulmonary Cavitary Lesions in Children and Adolescents: With a Specific Focus in Perforated Hydatid Cyst, Necrotizing Pneumonia, and Lung Abscess

Shokrollah Hafez Eftekhari[1], Anahita Azinfar[2], Masoud Mahdavi Rashed[1], Mahdi Abdollahi Karizno[1]

[1]Department of Radiology, Mashhad University of Medical Sciences Faculty of Medicine, Mashhad, Iran
[2]Mashhad University of Medical Sciences School of Medicine, Mashhad, Iran

Cite this article as: Hafez Eftekhari S, Azinfar A, Mahdavi Rashed M, Abdollahi Karizno M. Diagnostic value of non-contrast lung CT in evaluating pulmonary cavitary lesions in children and adolescents: with a specific focus in perforated hydatid cyst, necrotizing pneumonia, and lung abscess. *Thorac Res Pract.* 2026;27(4):200-206

ABSTRACT

OBJECTIVE: Accurate diagnosis of pulmonary cavitary lesions in children is essential for effective treatment and prevention of complications. While chest X-rays are commonly used, they lack the precision needed for detailed lesion characterization. Non-contrast high-resolution computed tomography (CT) scans offer superior imaging capabilities, particularly for cavitary lesions, though their use in pediatric patients must be balanced against radiation exposure risks. This study evaluates the diagnostic utility of non-contrast CT in distinguishing among hydatid cysts, necrotizing pneumonia, and lung abscesses by incorporating both clinical and imaging data.

MATERIAL AND METHODS: Conducted as a cross-sectional study at Akbar Hospital (Mashhad University of Medical Sciences) from 2021 to 2023, the study included pediatric patients who presented with pulmonary symptoms and underwent non-contrast lung CT. Statistical analyses involved chi-square tests for categorical variables and t-tests or non-parametric equivalents for continuous data. Diagnostic metrics such as sensitivity, specificity, predictive values, and accuracy were calculated.

RESULTS: Key radiological features that identify necrotizing pneumonia with high sensitivity and specificity include irregular lesion borders and fully aerated lesions. Additional indicators, such as pleural effusion heterogeneity and pneumothorax, further enhanced diagnostic accuracy. Hydatid cysts were reliably diagnosed by the presence of floating membranes and associated cystic structures, each demonstrating 100% specificity. In contrast, lung abscesses lacked distinctive imaging features, underscoring the need for further research.

CONCLUSION: Non-contrast CT is a valuable diagnostic tool for pediatric cavitary pulmonary lesions, particularly necrotizing pneumonia and hydatid cysts. The findings support the adoption of standardized imaging protocols to improve diagnostic precision and clinical outcomes.

KEYWORDS: Cavitary pulmonary lesions, non-contrast CT scan, necrotizing pneumonia, hydatid cyst, lung abscesses, children, adolescents

Received: 29.08.2025 ***Revision Requested:*** 24.12.2025 ***Last Revision Received:*** 14.02.2026

Accepted: 05.04.2026 ***Epub:*** 23.07.2026 ***Publication Date:*** 24.07.2026

INTRODUCTION

The accurate diagnosis of pulmonary cavitary lesions in children remains a clinical challenge. Chest radiography is typically the first-line imaging modality; however, its limited sensitivity and inability to adequately characterize complex lung lesions are well recognized. In contrast, computed tomography (CT) provides superior visualization of lung parenchyma and mediastinal structures and plays a key role in the evaluation of complicated pulmonary diseases.[1,2]

Corresponding author: Masoud Mahdavi Rashed, MD, e-mail: mahdavirm@mums.ac.ir

Pulmonary cavities most commonly result from infectious processes in children, including necrotizing pneumonia, lung abscess, and pulmonary hydatid disease. These entities differ significantly in their management and prognosis, making accurate differentiation essential. CT imaging offers high sensitivity for detecting cavitation and provides valuable morphological details such as wall thickness, air-fluid levels, and internal membranes, which aid in distinguishing among these conditions.[3,4]

Pulmonary hydatid disease, caused by Echinococcus granulosus, is particularly prevalent in endemic regions and commonly affects pediatric patients. Necrotizing pneumonia is a severe complication of bacterial infection and is characterized by parenchymal destruction and multiple irregular cavities, whereas lung abscess typically presents as a thick-walled cavity with an air-fluid level. CT findings are crucial for differentiating these entities and guiding appropriate treatment strategies.[5,6]

Despite the increasing use of lung CT in pediatric practice, concerns regarding radiation exposure necessitate careful evaluation of its diagnostic value, particularly when non-contrast protocols are used.[7]

Therefore, this study aims to assess the diagnostic value of non-contrast chest CT for differentiating cavitary pulmonary lesions in children at Akbar Hospital, Mashhad University of Medical Sciences between 2021 and 2023. To our knowledge, this is the first study in Iran focusing specifically on this topic.

MATERIAL AND METHODS

This descriptive-analytical study was conducted on pediatric patients who were referred to Akbar Hospital, Mashhad University of Medical Sciences, between 2021 and 2023 for evaluation of respiratory symptoms and who underwent non-contrast chest CT. The study was approved by the Ethics Committee of Mashhad University of Medical Sciences (IR. MUMS.IRH. REC.1403.115, date: 2024/08/05) and carried out in accordance with ethical principles, with confidentiality ensured and informed consent obtained from parents or legal guardians prior to patient enrollment. During the study period, a total of 2,225 chest CT examinations in pediatric patients were retrospectively reviewed. Among these, 300 patients were identified as having cavitary pulmonary lesions on CT imaging. These patients were evaluated retrospectively to determine their clinical course and final diagnoses. Of the 300 patients with cavitary lung lesions, 125 were children who underwent surgical intervention, which provided definitive pathological confirmation of the diagnosis. All 125 surgically treated patients were included in the final study population and constituted the study cohort.

Main Points

- Non-contrast lung computed tomography (CT) provides valuable diagnostic information for evaluating pulmonary cavitary lesions in children and adolescents.
- Specific CT findings, including irregular lesion borders and complete intralesional aeration, demonstrate high sensitivity and specificity for diagnosing necrotizing pneumonia.
- The presence of floating membranes and associated cystic structures on CT is highly specific for the diagnosis of a perforated pulmonary hydatid cyst.
- Lung abscesses do not show sufficiently distinctive imaging features on non-contrast CT, which indicates the need for further diagnostic evaluation and research.
- Standardized interpretation of CT imaging features can improve diagnostic accuracy and guide appropriate clinical management of pediatric cavitary lung lesions.

Exclusion criteria included patients without cavitary lesions on CT; patients with cavitary lesions who did not undergo surgical intervention; individuals with incomplete clinical or follow-up data (including outpatients); patients unwilling to participate; and cases in which a definitive diagnosis could not be established.

The final diagnoses of the included patients were confirmed by surgical intervention followed by histopathological examination. Among the 125 included patients, 38 were diagnosed with hydatid cysts, 63 with necrotizing pneumonia, and 24 with lung abscesses.

Although necrotizing pneumonia is primarily managed conservatively, surgical intervention was performed in selected cases due to failure of medical treatment or the development of complications. Indications for surgery in patients with necrotizing pneumonia included a lack of clinical or radiological improvement despite appropriate antibiotic therapy, persistent or enlarging cavitary lesions, suspicion of alternative diagnoses, development of complications such as lung abscess formation, bronchopleural fistula, empyema, or massive hemoptysis. Therefore, the necrotizing pneumonia cases included in this study represent a selected subgroup of surgically managed patients and do not reflect the overall population of children with necrotizing pneumonia.

Data collection was performed using a standardized and validated checklist designed to capture demographic information, clinical presentations, and imaging findings. Specific variables related to cavitary lesions were systematically evaluated, including wall thickness; aeration patterns (fully aerated, fluid-air level, or air-bubble pattern); presence of floating membranes or the "water lily sign"; presence or absence of soft tissue within the cavity; and communication between the cavity and the bronchus. Additionally, the presence of pleural effusion, laterality (unilateral or bilateral), and homogeneity or heterogeneity (heterogeneity assessed by densitometry) were assessed.

All CT scans were performed using a Canon Aquilion 16-slice CT scanner, manufactured by Toshiba. The scanning protocols, including kilovoltage and milliampere settings, were adjusted based on each patient's body weight to optimize image quality and minimize radiation exposure.

Statistical Analysis

The reports were interpreted by an experienced radiologist, blinded to the patients' initial clinical diagnoses. Patients were followed up and their final diagnoses were recorded. The final diagnoses for all patients, encompassing necrotizing

pneumonia, hydatid cysts, and lung abscesses, were confirmed through surgical intervention followed by histopathological examination. All collected data were entered into SPSS software for analysis.

For statistical analysis, quantitative variables were summarized using the mean and standard deviation or the median and interquartile range, depending on the data distribution, while qualitative variables were reported as frequencies and percentages. Relationships between qualitative variables were analyzed using the chi-square test, while independent t-tests or their non-parametric equivalents were employed to compare quantitative variables. Additionally, diagnostic performance metrics of CT scans, including sensitivity, specificity, positive predictive value (PPV), negative predictive value, and overall diagnostic accuracy, were calculated, with statistical significance set at a *P* value of less than 0.05.

RESULTS

Evaluation of Wall Thickness in Differentiating Hydatid Cyst, Abscess, and Necrotizing Pneumonia

The descriptive statistical analysis of the study variables revealed significant differences in cavity wall thickness across different disease groups. The mean cavity wall thickness in patients with hydatid cysts was 5.637 mm, the highest among the groups. In patients with necrotizing pneumonia, the mean thickness was 4.45 mm, whereas in lung abscess cases it was the lowest at 3.621 mm. The overall mean thickness was calculated as 4.825 mm, with a range of 2 to 10 mm. This variation in wall thickness may indicate differing levels of tissue involvement and disease progression among these conditions (Table 1).

Receiver operating characteristic (ROC) curve analysis of wall thickness measurements for differentiating hydatid cysts, abscesses, and necrotizing pneumonia showed that the areas under the curve (AUCs) for abscess and necrotizing pneumonia were below 70%, indicating poor diagnostic performance. However, the AUC for hydatid cyst was 83%, suggesting good discriminative ability.

Based on ROC analysis, a wall thickness cut-off of 3.9 mm demonstrated a sensitivity of 90% and a specificity of 60.9%, whereas a cut-off of 4.15 mm yielded a sensitivity of 80% and a specificity of 74%. These findings suggest that increasing the wall thickness threshold enhances specificity but may reduce sensitivity, highlighting the importance of selecting an appropriate cut-off value for accurate differentiation of pulmonary cavitary lesions.

These results suggest that wall thickness can be a useful parameter in distinguishing hydatid cyst from abscess and necrotizing pneumonia, with an optimal threshold in the specified range providing a balance between sensitivity and specificity (Figure 1).

Analysis of Imaging Findings and Lesion Characteristics

In this study, 81.6% of cases presented with a single lesion, while 18.4% exhibited multiple lesions, suggesting that solitary cavitary lung lesions were the predominant pattern. Regarding lesion distribution, 11.2% of cases showed spread to other lung regions, indicating a more severe disease course in these patients. Regarding lesion borders, 61.6% had irregular margins, while 38.4% had regular margins, reflecting heterogeneity in the growth and evolution of these lesions. The location of lesions within the lung varied: 19.2% were centrally located, 31.2% were peripherally distributed, and 49.6% exhibited a mixed (central and peripheral) pattern. This distribution may suggest different pathogenic mechanisms for each condition.

Diagnostic Criteria for Necrotizing Pneumonia

The most reliable imaging indicators for necrotizing pneumonia were:

Irregular pattern, with 92.06% sensitivity and 69.35% specificity, making it the strongest predictor. Fully aerated lesion, with 82.54% sensitivity and 72.58% specificity, providing a well-balanced diagnostic marker. Pleural effusion heterogeneity, with a specificity of 80% and a PPV of 78.58%.

Pneumothorax had a specificity of 90.16% and a PPV of 80%, indicating that, while uncommon, its presence strongly

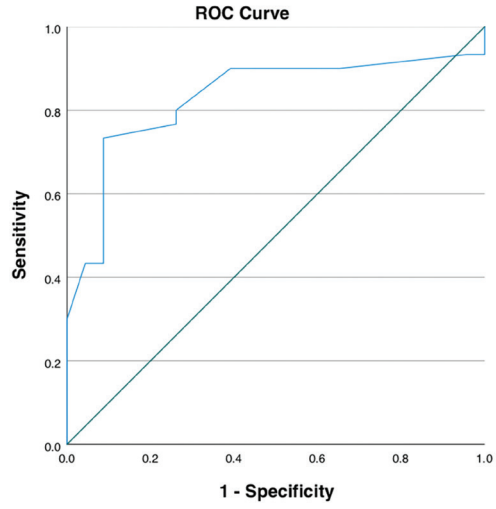

Figure 1. Receiver operating characteristic (ROC) curve analysis for the diagnostic performance of the proposed method. The area under the curve demonstrates high sensitivity and specificity in distinguishing wall thickness

Table 1. Diagnostic performance metrics for necrotizing pneumonia using the proposed imaging criteria

Variable	Group	Number (n)	Mean	Standard deviation	%95 confidence interval
Wall thickness	Hydatid cyst	30	5.637	1.8797	4.935–6.339
	Necrotizing pneumonia	4	4.450	1.6763	1.783–7.117
	Abscess	19	3.621	0.6152	3.325–3.918
	Total	53	4.825	1.7844	4.333–5.316

suggests necrotizing pneumonia. On the other hand, floating membranes and associated cysts had 0% sensitivity, confirming their irrelevance in diagnosing necrotizing pneumonia. The presence of a regular lesion pattern also had very low sensitivity (7.94%), making it an unreliable marker for this condition (Table 2).

Diagnostic Criteria for Hydatid Cyst

The most definitive diagnostic criteria for hydatid cysts included:

Floating membranes, with 100% specificity and PPV, confirming that when a membrane is observed, hydatid cyst is almost certain.

Associated cysts, also with 100% specificity and PPV, reinforcing its diagnostic reliability.

However, consolidation had high sensitivity (86.84%) but very low specificity (9.2%), indicating that while consolidation is commonly seen in hydatid cysts, it is also prevalent in other lung conditions. Similarly, diffuse involvement had only 19.95% specificity, making it an unreliable diagnostic feature for hydatid cysts (Table 3).

Diagnostic Criteria for Pulmonary Abscess

Key diagnostic indicators for lung abscess included:

No qualitative imaging attributes were identified as reliable diagnostic indicators for lung abscesses. Conversely, floating membranes and associated cysts were poor diagnostic features of abscess, with 0% sensitivity, indicating that these characteristics are not helpful in distinguishing abscess from other conditions. Additionally, bronchial communication and pneumothorax demonstrated low sensitivity, further confirming their limited role in the diagnosis of abscesses (Table 4).

Overall Diagnostic Implications

The findings of this study highlight the most effective imaging criteria for distinguishing between necrotizing pneumonia, hydatid cyst, and lung abscess:

For necrotizing pneumonia, irregular, fully aerated lesions, pleural effusion heterogeneity, and pneumothorax were the strongest markers, while associated cysts and floating membranes were irrelevant.

For hydatid cysts, floating membranes and associated cysts had the highest diagnostic accuracy, making them nearly definitive for diagnosis. Conversely, no distinct qualitative imaging characteristics were identified as reliable diagnostic markers for lung abscesses.

Imaging-based Diagnostic Highlights

Lesion margins, aeration patterns, and specific signs such as floating membranes and associated cystic lesions played crucial roles in differentiating among hydatid cysts, necrotizing pneumonia, and lung abscesses.

Table 2. Information on ANOVA analysis for necrotizing pneumonia

Characteristics for predicting necrotizing pneumonia	Sensitivity %	Specificity %	Positive predictive value %	Negative predictive value %	Accuracy %	%95 confidence interval
Solitary	87.30	24.19	53.92	65.22	56	46.84–64.86
Irregular border	92.06	69.35	75.32	89.58	80.80	72.79–87.29
Fully aerated	82.54	72.58	75.36	80.36	77.6	69.28–84.57
Associated cysts	0	77.42	0	43.24	38.40	29.84–47.52
Floating membranes	0	69.35	0	40.07	34.34	26.14–43.42
Heterogeneous effusion	48.89	80.0	78.57	51.06	61.33	49.38–72.36
Consolidation	96.83	17.74	54.46	84.62	57.60	48.44–66.39
Pneumothorax	38.1	90.16	80	58.51	63.71	54.6–72.15

Table 3. Comparison of sensitivity, specificity, and accuracy values across the study

Characteristics for predicting hydatid cyst	Sensitivity %	Specificity %	Positive predictive value %	Negative predictive value %	Accuracy %	%95 confidence interval
Solitary	79.17	17.82	18.63	78.26	29.60	21.77–38.42
Diffuse	18.42	91.95	50.00	72.07	69.60	60.74–77.51
Regular border	63.16	72.41	50.00	81.82	69.60	60.74–77.51
Associated cysts	36.84	100	100	78.38	80.80	72.79–87.29
Floating membranes	50.00	100	100	82.08	84.80	77.29–90.59
Homogeneous effusion	83.33	47.06	42.55	85.71	58.67	46.70–69.92
Consolidation	86.84	9.20	29.46	61.54	32.80	24.67–41.77
Pneumothorax	16.22	72.41	20	67.02	55.65	46.45–64.56

ANOVA: analysis of Variance

Irregular lesion margins were reliable features of necrotizing pneumonia (Figure 2A), whereas smooth, regular margins were seen in hydatid cysts (Figure 2B) and lung abscesses (Figure 2C).

Complete aeration of the cavity was a dependable characteristic in necrotizing pneumonia (Figure 3A), although air-fluid levels and air bubbles were not effective qualitative indicators for distinguishing among these three diseases (Figure 3B and 3C).

Heterogeneous pleural effusion patterns were most notable in necrotizing pneumonia (Figure 4A), whereas homogeneous effusions were seen in hydatid cysts (Figure 4B) and abscesses (Figure 4C).

The presence of floating membranes and associated cystic lesions was highly specific for hydatid cysts and was not observed in other conditions. Floating membranes were visualized within the cystic cavity (Figure 5A), and associated cystic lesions with fluid-density areas were observed in different lobes (Figure 5B). These signs provide high diagnostic specificity for hydatid disease.

Additional Findings on Floating Membrane Presence

A notable finding in this study was the association between the floating membrane and disease type. The results of the chi-square test (P = 0.000) demonstrated a highly significant

Table 4. Statistical evaluation of diagnostic performance indices

Characteristics for predicting abscess	Sensitivity %	Specificity %	Positive predictive value %	Negative predictive value %	Accuracy %	%95 confidence interval
Diffuse	0	86.14	0	78.38	69.6	60.74–77.51
Regular border	79.17	71.39	39.58	93.51	72.8	64.12–80.37
Air fluid level	75	71.29	38.30	92.31	72	63.27–79.66
Air bubble	8.33	93.07	22.22	81.03	76.80	68.41–83.88
Floating membranes	0	81.19	0	77.36	65.6	56.58–73.86
Associated cysts	0	86.14	0	78.38	69.6	60.74–77.51
Bronchial communication	0	89.11	0	78.95	72	63.27–79.66
Homogeneous effusion	66.67	37.68	8.51	92.86	40	28.85–51.96
Pneumothorax	0	70	0	74.47	56.45	47.26–65.33

ANOVA: analysis of Variance

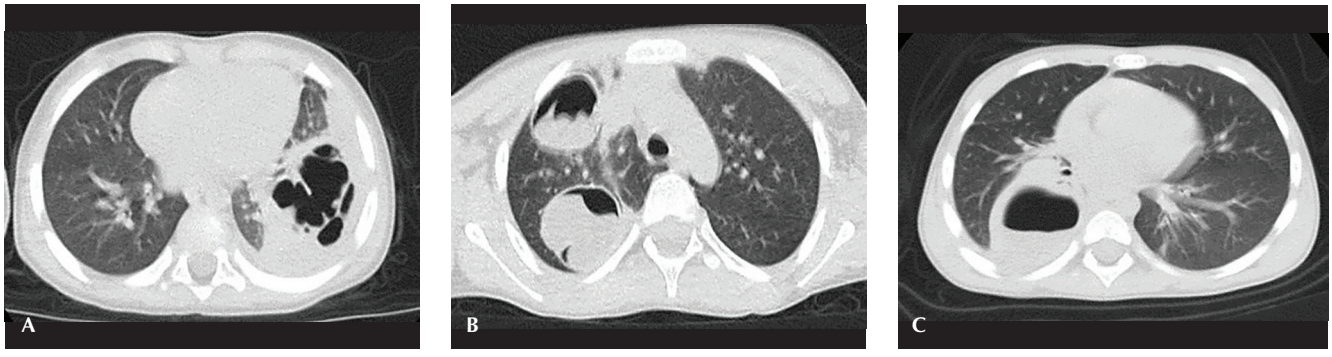

Figure 2. Cavity margin characteristics. (A) Irregular margin in a patient with necrotizing pneumonia. (B) Smooth and well-defined margin in a patient with hydatid cyst. (C) Smooth and defined margin in a patient with lung abscess

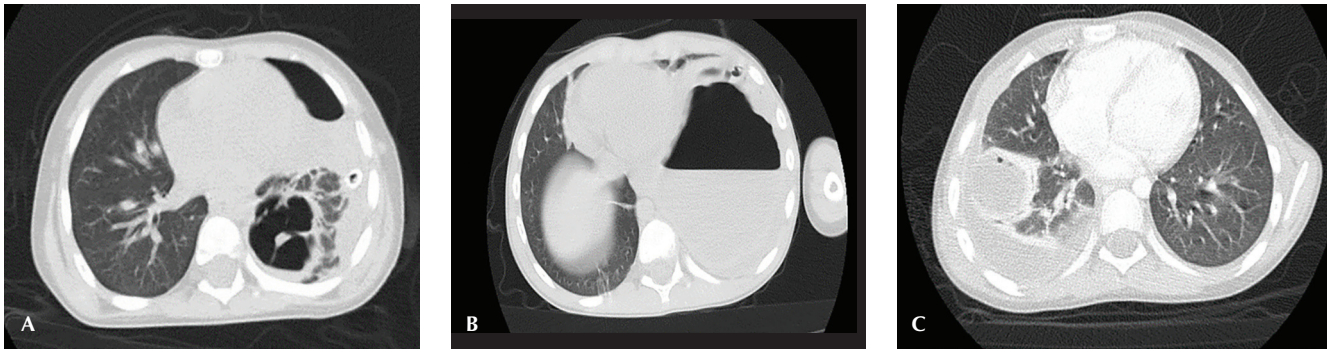

Figure 3. Air patterns within cavitary lesions. (A) Completely aerated cavity in necrotizing pneumonia. (B) Air-fluid level in a patient with diagnosis of hydatid cyst. (C) Presence of air bubbles in a patient with diagnosis of lung abscess

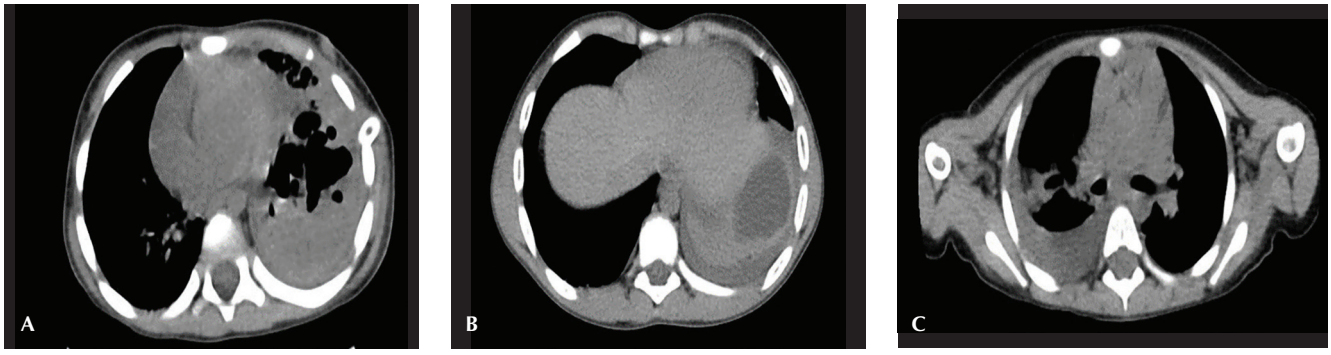

Figure 4. Pleural effusion characteristics. (A) Heterogeneous pleural effusion in necrotizing pneumonia. (B) Homogeneous pleural effusion in hydatid cyst. (C) Homogeneous pleural effusion in lung abscess

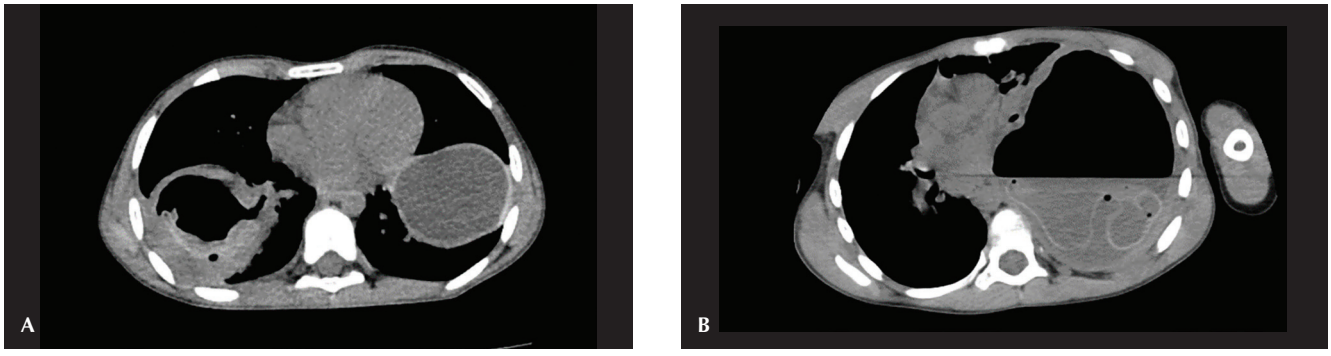

Figure 5. Hydatid-specific features. (A) Floating membranes within the cavity, highly suggestive of hydatid cyst. (B) A cavitary lesion in the right lung and an associated cystic lesion with fluid density in the left lung, suggestive of a hydatid cyst

relationship, confirming that the presence of a floating membrane is a strong indicator of a hydatid cyst, whereas none of the patients with necrotizing pneumonia or abscess exhibited membrane structures.

DISCUSSION

Summary of Key Findings

This study evaluated the diagnostic value of non-contrast chest CT in differentiating pediatric cavitary pulmonary lesions, focusing on hydatid cysts, necrotizing pneumonia, and lung abscesses. Our findings demonstrate that several CT features—including cavity wall characteristics, internal aeration patterns, the presence of floating membranes or associated cysts, pleural effusion, and pneumothorax—show distinct distributions among these entities. Non-contrast CT provided sufficient diagnostic information to reliably differentiate hydatid cysts from necrotizing pneumonia, whereas lung abscesses demonstrated more overlapping and non-specific imaging features. These results support the clinical utility of non-contrast CT as an effective first-line imaging modality in children with cavitary lung lesions.

Diagnostic Implications of Computed Tomography Findings

Accurate differentiation of cavitary lung lesions remains a diagnostic challenge due to overlapping clinical and radiological presentations. In our cohort, necrotizing pneumonia was more commonly associated with irregular lesion borders, fully aerated cavities, pleural effusion heterogeneity, and pneumothorax. In contrast, hydatid cysts demonstrated more specific imaging characteristics, including floating membranes and associated cystic lesions, which are well-recognized diagnostic indicators. Lung abscesses, however, lacked distinctive qualitative CT features, underscoring the need for further studies to establish more reliable imaging criteria for this subgroup. These findings highlight the value of pattern recognition on non-contrast CT in guiding accurate diagnosis and timely clinical management.

Role of Non-contrast Computed Tomography in Pediatric Imaging

Radiation exposure is a major concern in pediatric imaging, and minimizing dose without compromising diagnostic accuracy is a key priority. In this context, our study emphasizes that non-contrast CT can provide adequate diagnostic detail for evaluating cavitary lung lesions in children, potentially avoiding unnecessary contrast administration or repeat imaging. Dose optimization strategies—such as adjusting tube current based on body weight and using appropriate reconstruction techniques—allow CT to be performed safely while maintaining diagnostic confidence. Importantly, our findings do not argue against CT use, but rather support optimized and judicious CT utilization in pediatric patients.

Comparison with Other Imaging Modalities and Previous Studies

Alternative imaging modalities, such as lung ultrasound and magnetic resonance imaging (MRI), have been proposed as complementary tools in selected pediatric cases. Ultrasound

may be useful for assessing pleural complications, particularly in necrotizing pneumonia, while MRI offers superior soft-tissue contrast in complex lesions. However, both modalities have limitations in routine evaluation of air-containing lung parenchyma and are less practical in acute clinical settings. Therefore, CT—particularly non-contrast CT—remains the most informative and accessible modality for the initial evaluation of cavitary pulmonary lesions.

Our findings are consistent with previous studies, including Naggar et al.[3], who emphasized the importance of recognizing CT patterns in narrowing differential diagnoses of lung cavitation. Carrard et al.[8] reported that CT was superior to ultrasound in detecting parenchymal necrosis and cavitation, while Erdem and Erdem[9] demonstrated the superiority of CT over chest radiography in characterizing hydatid cysts. Collectively, these studies support the central role of CT in pediatric cavitary lung disease, aligning with our results.

Clinical Impact and Future Directions

By identifying reliable CT-based imaging markers, this study contributes to improved clinical decision-making and to more targeted patient management. Integrating radiological findings with clinical and pathological data can reduce diagnostic uncertainty and help avoid unnecessary contrast-enhanced or repeated CT examinations. Future research should focus on multicenter validation of these imaging criteria, the selective role of MRI in problem-solving scenarios, and the development of standardized diagnostic algorithms for pediatric cavitary lung lesions.

Strengths and Study Limitations

The strengths of this study include a well-defined pediatric cohort with surgically and histopathologically confirmed diagnoses, a blinded interpretation of CT images by an experienced radiologist, and an exclusive evaluation of non-contrast CT, reflecting real-world pediatric imaging practice.

Several limitations should be acknowledged. The retrospective design may introduce selection bias and result in incomplete data. The study was conducted at a single tertiary referral center, which may limit the generalizability of the findings. Additionally, only surgically managed cases were included, representing a subset of patients with more severe disease. Chest radiography findings were not systematically analyzed because variability in acquisition and reporting prevented direct comparison with CT. Finally, the relatively limited sample size may have reduced statistical power for subgroup analyses.

CONCLUSION

This study provides critical insights into the radiological evaluation of pediatric cavitary lung lesions, highlighting key diagnostic features that can improve clinical decision-making. By addressing current diagnostic challenges and proposing evidence-based solutions, such as radiation dose reduction, standardized diagnostic criteria, and alternative imaging techniques, this research contributes to enhancing pediatric imaging practices and optimizing patient care. Future studies should validate these findings through multicenter trials and exploring novel imaging techniques to ensure the best possible outcomes for pediatric patients.

Ethics

Ethics Committee Approval: The study was approved by the Ethics Committee of Mashhad University of Medical Sciences (IR. MUMS.IRH.REC.1403.115, date: 2024/08/05).

Informed Consent: Informed consent obtained from parents or legal guardians prior to patient enrollment.

Footnotes

Authorship Contributions

Concept: S.H.E., M.M.R., Design: S.H.E., M.M.R., Data Collection or Processing: A.A., M.A.K., Analysis or Interpretation: A.A., Literature Search: M.M.R., Writing: S.H.E., M.A.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Copley SJ, Coren M, Nicholson AG, Rubens MB, Bush A, Hansell DM. Diagnostic accuracy of thin-section CT and chest radiography of pediatric interstitial lung disease. *AJR Am J Roentgenol*. 2000;174(2):549-554. **[Crossref]**
2. Lynch DA, Brasch RC, Hardy KA, Webb WR. Pediatric pulmonary disease: assessment with high-resolution ultrafast CT. *Radiology*. 1990;176(1):243-248. **[Crossref]**
3. Naggar A, Laasri K, Berrada K, et al. Differential diagnoses of cavitary lung lesions on computed tomography: a pictorial essay. *Egypt J Radiol Nucl Med*. 2023;54(1):149. **[Crossref]**
4. Lee KC, Kang EY, Yong HS, et al. A stepwise diagnostic approach to cystic lung diseases for radiologists. *Korean J Radiol*. 2019;20(9):1368-1380. **[Crossref]**
5. Sayedi SJ, Zandi H, Bagheri S. Characteristics of pediatric pulmonary hydatid cyst in Mashhad, North East of Iran. *Rev Clin Med*. 2022;9(4):196-199. **[Crossref]**
6. Hamouri S, Odat H, Syaj S, Hecker E, Alrabadi N. Rupture of pulmonary hydatid cyst in pediatrics: a cross-sectional study. *Ann Med Surg (Lond)*. 2021;62:31-36. **[Crossref]**
7. Owens C. Pearls and pitfalls in HRCT in children. *Paediatr Respir Rev*. 2006;7(Suppl 1):S44-S49. **[Crossref]**
8. Carrard J, Bacher S, Rochat-Guignard I, et al. Necrotizing pneumonia in children: chest computed tomography vs. lung ultrasound. *Front Pediatr*. 2022;10:898402. Erratum in: *Front Pediatr*. 2023;11:1178223. **[Crossref]**
9. Erdem CZ, Erdem LO. Radiological characteristics of pulmonary hydatid disease in children: less common radiological appearances. *Eur J Radiol*. 2003;45(2):123-128. **[Crossref]**

Thorac Res Pract. 2026;27(4):207-214 **DOI:** 10.4274/ThoracResPract.2025.2025-7-11

Original Article

Correlation of COVID-19 Mortality with Socioeconomic Status, Air Quality, and Housing Density

Elif Reyhan Şahin[1], Ömür Güngör[2], Ahmet Uğur Demir[3], Ali Fuat Kalyoncu[4]

[1]Clinic of Occupational Medicine, Erzurum City Hospital, Erzurum, Türkiye
[2]Clinic of Occupational Medicine, University of Health Sciences Türkiye, Adana City Training and Research Hospital, Adana, Türkiye
[3]Department of Chest Diseases, Hacettepe University Faculty of Medicine, Ankara, Türkiye
[4]Department of Chest Diseases, Division of Allergy and Clinical Immunology, Hacettepe University Faculty of Medicine, Ankara, Türkiye

Cite this article as: Şahin ER, Güngör Ö, Demir AU, Kalyoncu AF. Correlation of COVID-19 mortality with socioeconomic status, air quality, and housing density. *Thorac Res Pract.* 2026;27(4):207-214

ABSTRACT

OBJECTIVE: Studies have reported associations between coronavirus disease-2019 (COVID-19) mortality and socioeconomic status (SES) and air pollution, but few have investigated the association with housing density. This ecological study aims to demonstrate the relationship between the COVID-19 mortality rate and these variables in the districts of Ankara.

MATERIAL AND METHODS: COVID-19 mortality rate calculations were based on data from the Ankara Metropolitan Municipality Cemetery Directorate. Air pollution data for the five years preceding the COVID-19 pandemic were obtained from the North Anatolia Clean Air Regional Directorate. The SES values for each district were divided into five categories. Housing density was calculated by dividing the number of households by the area of the district in square kilometers.

RESULTS: Levels of NO_X, O_3, $PM_{2.5}$, and NO (r = 0.78, 0.84, 0.74, and 0.80, respectively) and district-level SES (r = 0.43) were found to be moderately correlated with COVID-19-related deaths. COVID-19-related deaths were positively correlated with housing density (r = 0.68, *P* = 0.001).

CONCLUSION: These variables interact closely and are correlated with COVID-19 mortality across districts of Ankara.

KEYWORDS: Air pollution, socioeconomic status, housing density, COVID-19 mortality, environmental exposure

Received: 05.08.2025 ***Revision Requested:*** 22.09.2025 ***Last Revision Received:*** 24.09.2025

Accepted: 23.10.2025 ***Epub:*** 26.12.2025 ***Publication Date:*** 24.07.2026

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 is mainly transmitted through close person-to-person contact and by aerosol droplets smaller than 5 μm in diameter. Although nations and governments adopted drastic prevention measures, as of July 23, 2025, the coronavirus disease-2019 (COVID-19) pandemic has caused 7,098,440 deaths and 778,407,760 positive cases.[1,2] Global reports link social factors, in addition to medical ones, to COVID-19 mortality. To our knowledge, this is the first study in Ankara to examine the combined impact of housing density, socioeconomic status (SES), and air pollution indicators on COVID-19 mortality.

Several ecological studies have reported that outdoor air pollution is one of the most significant factors affecting COVID-19 mortality by increasing susceptibility to respiratory tract infections through long-term exposure to air pollution, which induces immunotoxicity, inflammation, and free radical production.[3-6] Similar to exposure to air pollution, socioeconomic inequality affects COVID-19 mortality and should be considered in detail.[6-8] Only a few studies have investigated the effects of housing density on COVID-19 mortality and demonstrated higher mortality in high-density districts.[9,10] This ecological study aims to examine the correlation between COVID-19 mortality and housing density, SES, and air pollution indicators in the districts of Ankara.

Corresponding author: Elif Reyhan Şahin, MD, e-mail: elifreyhan2020@gmail.com

MATERIAL AND METHODS

The study was conducted in accordance with the principles of the Declaration of Helsinki. As this is an ecological study utilizing publicly available aggregated data without any individual-level identifiers, ethical approval was not required in line with national and international research ethics guidelines.

Mortality Data: The COVID-19 mortality data were obtained from the Ankara Metropolitan Municipality Cemetery Directorate because the district-wise numbers of COVID-19-related deaths were not announced by the Turkish Ministry of Health for Ankara. Mortality data were collected from three large central cemeteries where most burials occurred between March 10, 2020, and May 31, 2022. In Türkiye, the code "Natural Death: Contagious Disease" has been added to the ICD-10 classification to determine COVID-19-related deaths.

In our country, the death notification system obligates sending a copy of the death certificate issued by the physician to the cemetery and holds the information on the residential area, age, cause, and date of death.[11] For this study, the number of deaths with the code "Natural Death: Contagious Disease" between March 10, 2020, and May 31, 2022 was obtained from each district's cemetery database and processed accordingly.

Calculation of Housing Density: The population living in the districts of Ankara was determined using the address-based population registration system of TurkStat. First, the number of households was calculated by dividing the population of each district by the average household size, using data from TurkStat's website. Housing was calculated as the number of houses per square kilometer of district area.[12] The housing density of each district is depicted in Figure 1A.

Socioeconomic Status Data: In 2022, the General Directorate of Development Agencies of the Ministry of Industry and Technology published the "Socioeconomic Development of Districts Ranking Study (*Sosyo-Ekonomik Gelişmişlik Endeksi*, SEGE-2022)", which contains information on differences in development among districts in Ankara and represents all dimensions of socioeconomic development.[13] The SEGE index for each district is based on numerical data stratified into five categories, as shown in Figure 1B.

Measurements of Air Pollution and National Air Quality Monitoring Network: Ankara is the capital city of Türkiye and has 25 districts, 1,425 neighborhoods, and 18 fixed air quality monitoring stations located in only 9 districts. Neighborhoods covered by these stations and the pollutants measured at each station were obtained from the "Ankara Province Clean Air Action Plan" published by the Ministry of Environment, Urbanization and Climate Change, Ankara Governorship, and Ankara Metropolitan Municipality.[14] The data obtained from the stations, as hourly averages, are transferred to the Data Operation Center of the Laboratory, Measurement and Monitoring Department, which is affiliated with the Ministry of Environment, Urbanization and Climate Change, General Directorate of Environmental Impact Assessment, Permit, and Inspection. At this center, data verification is performed, considering the devices' calibration and alarm information. Accordingly, monthly and annual reports are prepared using verified data, and the raw data obtained from the monitoring network are published simultaneously on www.havaizleme.gov.tr. After data validation is completed, the verified data are transferred to the website at the end of each month. Meteorological variables such as air temperature, humidity, and air pressure are recorded and processed simultaneously.[15,16]

Air pollution data for the five years preceding the COVID-19 pandemic, excluding data from the spring–summer season when reductions in air pollution were expected, were obtained from the North Anatolia Clean Air Regional Directorate. Air pollutants measured at all stations are shown in Table 1. Figures 1C-F show pre-pandemic levels of $PM_{2.5}$, NO_X, NO, and O_3 in the districts. The arithmetic mean of the measured air quality parameters was calculated for each pollutant. Measurements of meteorological variables, namely temperature, humidity, and air pressure, corresponding to the pollutant data, have also been considered in this study.

Statistical Analysis

IBM SPSS Statistics v. 23.0 for Windows was used for statistical analysis. District-level correlations between the COVID-19 mortality rate and air pollutants, housing density, and SES score were assessed using the Pearson correlation test. Cross-tabulation analysis was used to assess the relationships among multiple variables. A *P* value smaller than 0.05 was considered statistically significant.

Main Points

- This is the first ecological study in Ankara to jointly evaluate the impacts of air pollution, housing density, and socioeconomic disparities on coronavirus disease-2019 (COVID-19) mortality.
- COVID-19-related deaths in Ankara were significantly correlated with air pollution indicators, particularly long-term exposure to $PM_{2.5}$, NO_X, NO, and O_3.
- Higher housing density was strongly associated with increased COVID-19 mortality across Ankara's districts.
- Contrary to prior international findings, districts with higher socioeconomic status (SES) in Ankara had higher COVID-19 mortality rates.
- COVID-19 mortality was not significantly affected by meteorological variables, but was associated with age and gender, with elderly males accounted for the majority of deaths. This aligns with global evidence emphasizing age and sex as key risk factors for COVID-19 mortality.

RESULTS

Burials from four districts —Kalecik, Güdül, Pursaklar, and Evren— located in the three central cemeteries were not taken into consideration due to their location. More than 75% of the patients who died due to COVID-19 were older than 65 years of age. In addition, approximately 60% of elderly individuals were male, accounting for nearly the same proportion of total COVID-19-related deaths.

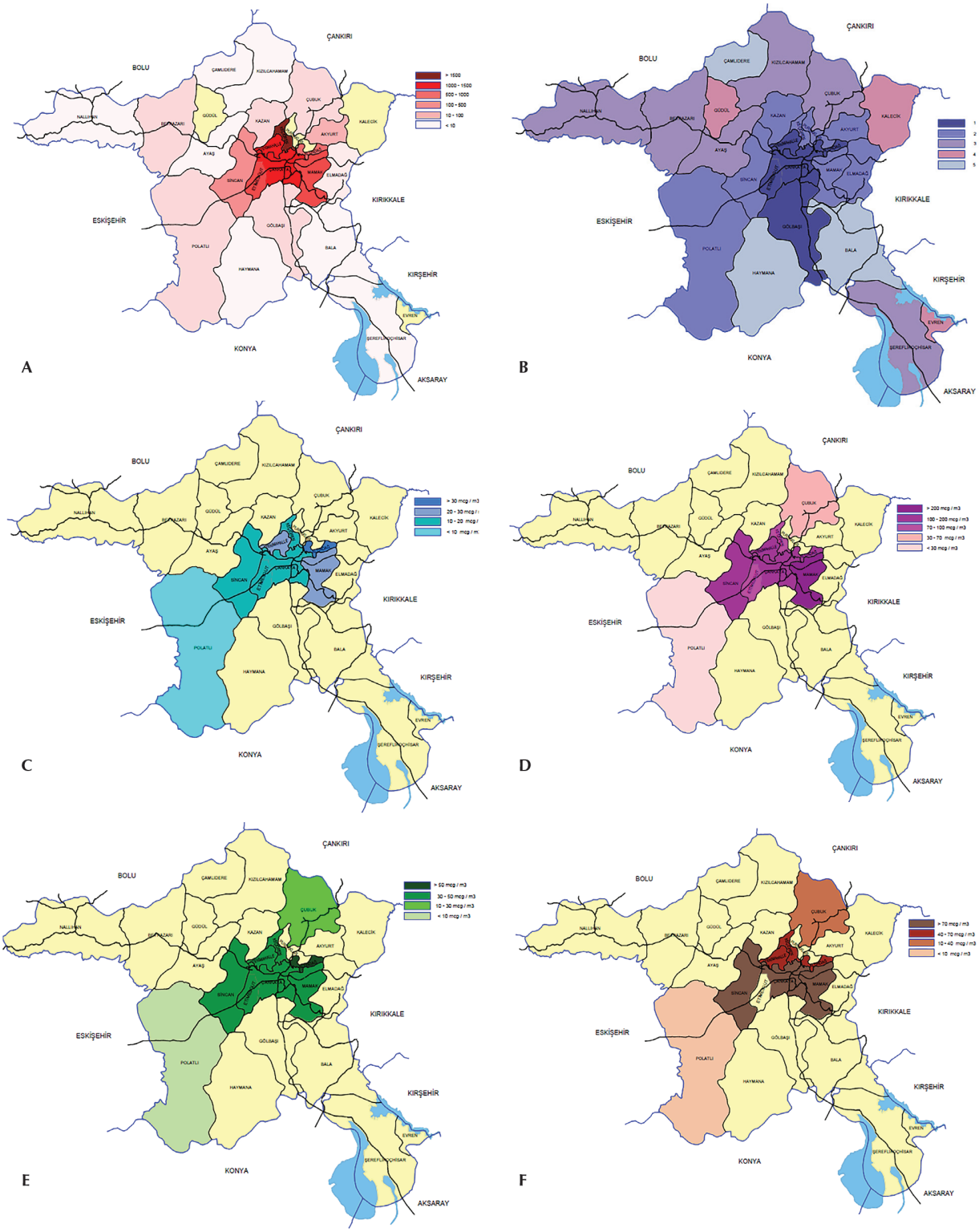

Figure 1. Spatial distribution of key variables analyzed in the study

A) Housing density (households/km^2) by district in Ankara, calculated using TurkStat address-based population registration system data (2022)

B) Socioeconomic status index (*Sosyo-Ekonomik Gelişmişlik Endeksi*, SEGE-2022), stratified into 5 categories by the General Directorate of Development Agencies, Ministry of Industry and Technology

C) Mean pre-pandemic levels of $PM_{2.5}$ (μg/m^3) across districts, based on validated data from the National Air Quality Monitoring Network (2015-2019)

D) Mean pre-pandemic levels of NO_X (μg/m^3) across districts, measured by fixed monitoring stations in 9 districts and averaged over 2015-2019

E) Mean pre-pandemic levels of NO (μg/m^3) across districts, obtained from the same monitoring network (2015-2019)

F) Mean pre-pandemic levels of O_3 (μg/m^3) across districts, reported by the Ministry of Environment, Urbanization and Climate Change

All variables shown were subsequently analyzed in relation to district-level COVID-19 mortality (March 2020-May 2022)

COVID-19: Coronavirus disease-2019

The percentage of the population and the COVID-19 mortality rates for each district of Ankara are presented in Table 2. The population is mainly concentrated in seven districts where air quality measurements are conducted, while the remaining districts hold less than 5% of the population, resulting in a heterogeneous distribution.

Table 3 demonstrates the correlations between COVID-19-related deaths and the variables considered in this study. Total COVID-19-related deaths were positively correlated with housing density and SES indices (r = 0.68, P = 0.001, and r = 0.48, P = 0.03, respectively). In Table 4, the correlations between the variables and COVID-19-related deaths among people aged 65 years and older are presented. Regarding pollutants, levels of $PM_{2.5}$, NO_X, O_3, and NO were correlated with total COVID-19-related deaths; in contrast, for deaths among the elderly only O_3 levels were correlated (r = 0.77, P = 0.02) (Tables 3 and 4). Total COVID-19-related deaths were not correlated with meteorological variables. Although the average household size was not correlated with the total COVID-19 death rate, it was correlated with COVID-19–related deaths among the elderly (Tables 3 and 4). Additionally, Table 5 illustrates the relationship between the study variables and COVID-19 mortality rates in males and females older than 65 years.

Table 1. List of air quality monitoring stations and measured pollutants in each district

Name of the district	Measured parameter
Çankaya	
Bahçelievler	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, CO, SO_2
Çankaya	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, CO, SO_2, O_3
Sıhhiye	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, CO, SO_2, O_3
Yaşamkent	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, CO, SO_2, O_3
Yenimahalle	
Batıkent	PM_{10}, NO_X, NO_2, NO, SO_2, O_3
Demetevler	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, SO_2
Ostim	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, CO, SO_2, O_3
Çubuk	NO_X, NO_2, NO, SO_2, O_3
Etimesgut	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, CO, SO_2
Mamak	PM_{10}, NO_X, NO_2, NO, CO, SO_2, O_3
Kayaş	PM_{10}, SO_2
Keçiören	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, SO_2, O_3
Etlik	PM_{10}, NO_X, NO_2, NO, CO, SO_2
Polatlı	PM_{10}, NO_X, NO_2, NO, CO, SO_2, O_3
Sincan	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, SO_2
Törekent	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, SO_2, CO
Altındağ	
Siteler	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, CO, SO_2, O_3
Ulus	PM_{10}, $PM_{2.5}$, NO_X, NO_2, NO, CO, SO_2, O_3

Table 2. COVID-19 mortality rates of districts

District	Population (% in the province)	COVID-19 mortality rate (per 10,000)	COVID-19 mortality rate of elderly (per 10,000)	COVID-19 mortality rate of males (per 10,000)	COVID-19 mortality rate of females (per 10,000)
With air quality measurement station					
Çankaya	16.52	13.03	19.10	16.57	9.77
Keçiören	16.41	12.62	31.71	14.96	10.37
Yenimahalle	12.25	14.09	31.15	18.11	10.29
Mamak	11.87	14.78	38.79	20.94	8.67
Etimesgut	10.55	6.89	55.98	9.32	4.47
Sincan	9.77	12.18	62.84	13.72	10.57

Table 2. Continued

District	Population (% in the province)	COVID-19 mortality rate (per 10,000)	COVID-19 mortality rate of elderly (per 10,000)	COVID-19 mortality rate of males (per 10,000)	COVID-19 mortality rate of females (per 10,000)
Altındağ	7.09	11.97	18.52	11.95	11.98
Polatlı	2.22	4.7	6.95	5.33	4.07
Çubuk	1.59	9.52	14.49	10.30	8.74
Without air quality measurement station					
Gölbaşı	2.49	1.89	6.22	1.39	0.97
Kahramankazan	1.01	3.62	6.36	4.72	2.47
Beypazarı	0.84	3.72	2.71	3.31	4.11
Elmadağ	0.77	7.68	9.90	6.74	8.63
Akyurt	0.67	6.99	18.92	7.10	6.88
Şereflikoçhisar	0.58	5.07	3.10	4.16	5.99
Haymana	0.47	6.96	4.40	7.02	6.88
Kızılcahamam	0.47	6.3	3.21	2.91	9.83
Nallıhan	0.47	2.59	0.88	2.98	2.21
Bala	0.40	5.66	3.86	5.55	5.78
Ayaş	0.23	9.16	4.13	9.95	8.24
Çamlıdere	0.15	8.38	3.67	6.78	10.18

COVID-19: Coronavirus disease-2019

Table 3. Correlations of COVID-19 mortality rate with the variables

	n	%	Correlation coefficient	*P*
NO	5975	93	**0.80**	**0.01***
NO_2	5975	93	0.43	0.25
NO_X	5975	93	**0.78**	**0.01***
PM_{10}	6078	95	0.53	0.18
$PM_{2.5}$	5888	**92**	**0.74**	**0.04***
SO_2	6105	96	0.58	0.14
O_3	3913	**61**	**0.84**	**0.01***
CO	2750	43	0.53	0.23
Temperature	6165	97	0.11	0.77
Humidity (%)	6165	97	0.14	0.72
Pressure	6165	97	-0.31	0.40
SES index	6384	100	**0.48**	**0.03***
Average household size	6384	100	0.17	0.45
Housing density	6384	100	**0.68**	**0.001***

*$P < 0.05$ was considered statistically significant, COVID-19: Coronavirus disease-2019, SES: Socioeconomic status

Table 4. Correlation of air pollutant types with the COVID-19 mortality rate in the elderly (>65 years old)

	n	%	Correlation coefficient	*P*
NO	4584	72	0.32	0.38
NO_2	4584	72	0.23	0.54
NO_X	4584	72	0.39	0.29
PM_{10}	4610	72	0.30	0.46
$PM_{2.5}$	4519	71	0.11	0.78
SO_2	4691	73	0.06	0.87
O_3	3013	**47**	**0.77**	**0.02***
CO	2132	33	-0.19	0.68
Temperature	4735	74	0.15	0.68
Humidity (%)	4735	74	-0.24	0.53
Pressure	4735	74	0.09	0.81
SES index	4880	**76**	0.35	0.10
Average household size	4880	76	**0.50**	**0.02***
Housing density	4880	76	**0.49**	**0.02***

*$P < 0.05$ was considered statistically significant, COVID-19: Coronavirus disease-2019, SES: Socioeconomic status

Table 5. Correlation of air pollutants with COVID-19 mortality rates in male and female patients (>65 years old)

	Male			Female		
	n	Correlation coefficient	*P*	n	Correlation coefficient	*P*
NO	3654	**0.70**	**0.03***	2321	**0.76**	**0.02***
NO_2	3654	0.39	0.30	2321	0.38	0.30
NO_X	3654	**0.84**	**0.004***	2321	0.43	0.24
PM_{10}	3731	0.51	0.19	2347	0.41	0.31
$PM_{2.5}$	3607	0.61	0.10	2281	**0.76**	**0.03***
SO_2	3744	0.48	0.22	2361	0.56	0.14
O_3	2397	**0.82**	**0.01***	1516	0.60	0.11
CO	1658	0.40	0.36	1092	0.61	0.14
Temperature	3778	-0.09	0.80	2387	0.43	0.24
Humidity (%)	3778	0.28	0.45	2387	-0.12	0.74
Pressure	3778	-0.28	0.45	2387	-0.31	0.41
SES index	3891	0.55	0.01	2493	0.24	0.29
Average household size	3891	0.27	0.23	2493	-0.06	0.76
Housing density	3891	**0.72**	**0.001***	2493	**0.46**	**0.03***

*$P < 0.05$ was considered statistically significant, COVID-19: Coronavirus disease-2019, SES: Socioeconomic status

DISCUSSION

Ankara, located in central Anatolia, is the second-most populous city in Türkiye, with a population of 5,747,325, according to Turkish Statistical Institute 2021 data. By surface area (km^2), it is the third-largest city in Türkiye, after Konya and Sivas. The province has a population density of 224 people per km^2, and the district with the highest density is Keçiören, with 5,930 people per km^2.[12] Approximately three-quarters of the population are employed in the service sector, which includes civil service, transportation, communication, and trade; one-quarter are employed in industry, and 2% are employed in agriculture. The industry is concentrated mainly in the textile, furniture, foundry, defense, and construction, which are located in 12 organized industrial zones covering a total area of more than 60 million square meters.[17] The traffic burden is considerable, particularly in central districts. Air pollution in Ankara reached dangerous levels in the early 1980s. However, currently, moderate air pollution exists due to the widespread use of natural gas instead of low-quality coal.[14]

Harmful effects of air pollutants on the respiratory tract have been investigated extensively, and many studies have revealed a correlation between air pollution and COVID-19-related deaths.[4-6,18-20] In a meta-analysis, Zang et al.[21] reported that both short- and long-term exposure to air pollutants were correlated with COVID-19 incidence and mortality. To reduce the adverse effects of outdoor air pollution during pandemics, epidemiological research should be conducted, and appropriate precautions should be implemented promptly.[21] In our study, we observed a significant correlation between COVID-19-related deaths and particularly high concentrations of NO_2, O_3, and $PM_{2.5}$. However, a major limitation of this study was the insufficient number of air-quality monitoring stations, which were located in only nine densely populated districts. This limitation prevented our research from providing a comprehensive assessment of air quality across the city, thereby reducing the representativeness of the data. Future research would benefit from an expanded network of monitoring stations covering a broader range of districts within the city. By drawing attention to the inadequacy of monitoring stations in the capital, we aim to encourage the expansion and strategic placement of additional stations, contribute to public health efforts, and support the development of more effective policies to address air pollution and its potential impact on COVID-19 outcomes. These improvements would allow a more accurate assessment of air pollution levels, better-informed decision-making, and targeted interventions to mitigate health risks, ultimately encouraging authorities to allocate resources to strengthen the city's air quality monitoring infrastructure.

The importance of social determinants of health has become more apparent during the COVID-19 pandemic. Many studies have indicated that COVID-19 is more severe and that the number of deaths is higher in regions where the inhabitants have low SES.[22-25] A systematic review and meta-analysis found a strong correlation between SES and COVID-19 outcomes.[26]

In this study, we observed a higher COVID-19 mortality rate in districts with high SES, which contrasts with findings from other studies.[27-29] This discrepancy is attributable to factors such as population crowding and an older population in those districts, as supported by data from TurkStat and the SEGE 2022 study.[12-14] These unique characteristics of the population in high-SES districts contributed to higher vulnerability to COVID-19 and ultimately higher mortality rates. Additionally, alternative explanations should be considered. Residents of socioeconomically disadvantaged districts often face obstacles in accessing healthcare services and may experience delays in diagnosis or treatment.[30,31] Limited diagnostic resources in hospitals serving disadvantaged areas and well-documented underreporting of deaths due to documentation problems can

also contribute to artificially low mortality figures.[32,33] Thus, the inverse association observed in our study may partly reflect disparities in access to care and reporting practices rather than true differences in mortality burden.

It is also possible that deaths in central, densely populated districts were more consistently documented due to easier hospital access, whereas deaths in rural or peripheral districts might have occurred at home and remained undocumented. This potential reporting bias should be acknowledged when interpreting the results.

Furthermore, sourcing data on COVID-19-related deaths exclusively from three central cemeteries is a significant limitation of our study. The limited scope of data collection from central cemeteries prevents a comprehensive understanding of COVID-19 mortality rates in smaller districts. Access to information on COVID-19-related deaths in smaller districts would provide a more accurate and complete assessment of the impact of the virus across different areas of the city. It is essential to acknowledge these limitations and consider them when interpreting the results of the study. Our future research efforts could aim to address these limitations by expanding data collection to include smaller districts and implementing more comprehensive sampling strategies to capture a broader representation of the population.

Housing density reflects the number of households per km^2. The effect of housing density on the COVID-19 mortality rate has also been reported in various studies, supporting our findings;[34-36] yet Wang et al.[37] observed a dose-response relationship between these parameters. In our study, however, the association was positive but not dose-dependent. This discrepancy may be explained by differences in study design (ecological vs. longitudinal), population structure, and the level of detail in the available housing and environmental data. Nevertheless, our results are consistent with the broader conclusion of Wang et al.[37] that high housing density increases the risk of COVID-19 mortality. These studies demonstrate that areas with higher housing density are associated with an increased risk of COVID-19 transmission and mortality. The close proximity and shared spaces in densely populated housing contribute to higher infection rates and poorer health outcomes. Our study contributes to the evidence that housing density is a significant factor in understanding and addressing COVID-19 mortality rates. Further investigations into the influence of housing density on COVID-19 outcomes would provide valuable guidance for public health interventions and policymaking.

Study Limitations

Despite the aforementioned limitations, this study stands out as the first to demonstrate this relationship specifically in the densely populated districts of Ankara. In the near future, as the Ministry of Health provides updated statistics on the distribution of COVID-19-related deaths, it will become possible to conduct further studies encompassing all the provinces of Ankara. These future studies could yield additional valuable insights into the impact of social disparities, air pollution, and housing density on COVID-19 outcomes.

CONCLUSION

This ecological study identified notable associations between air pollution, housing density, and COVID-19-related mortality across the districts of Ankara. Higher mortality rates were observed in areas with greater SES, likely reflecting demographic characteristics such as older age and higher population density. These findings call for improved air quality monitoring and public health strategies that consider both environmental and social determinants of health. While limited by data availability and uneven distribution of monitoring stations, this study provides an important foundation for future research aimed at guiding more effective and equitable health policies.

Ethics

Ethics Committee Approval: Not applicable.

Informed Consent: Informed consent was not required fort his ecological study.

Acknowledgments

The authors express their gratitude to Neşe Aydın (Middle East Technical University, Department of City and Regional Planning) for the calculations of the housing density and Vecdet Ünal (Topographical Engineer) for mapping.

Footnotes

Authorship Contributions

Surgical and Medical Practices: E.R.Ş., Ö.G., A.U.D., A.F.K., Concept: E.R.Ş., A.U.D., A.F.K., Design: E.R.Ş., A.U.D., A.F.K., Data Collection or Processing: E.R.Ş., Ö.G., A.U.D., A.F.K., Analysis or Interpretation: Ö.G., A.U.D., A.F.K., Literature Search: E.R.Ş., Ö.G., A.U.D., A.F.K., Writing: E.R.Ş., Ö.G., A.U.D., A.F.K.

Conflict of Interest: Ahmet Uğur Demir is a biostatistical consultant for Thoracic Research and Practice. However, he was not involved in any stage of the editorial evaluation or decision-making process for this manuscript. The other authors has no disclosures.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. World Health Organization. WHO COVID-19 dashboard. Accessed July 23, 2025. **[Crossref]**
2. World Health Organization. WHO COVID-19 dashboard. COVID-19 Cases, World. Accessed July 23, 2025. **[Crossref]**
3. Bourdrel T, Annesi-Maesano I, Alahmad B, Maesano CN, Bind MA. The impact of outdoor air pollution on COVID-19: a review of evidence from *in vitro*, animal, and human studies. *Eur Respir Rev*. 2021;30(159):200242. **[Crossref]**
4. Karan A, Ali K, Teelucksingh S, Sakhamuri S. The impact of air pollution on the incidence and mortality of COVID-19. *Glob Health Res Policy*. 2020;5:39. **[Crossref]**
5. Chen Z, Huang BZ, Sidell MA, et al. Near-roadway air pollution associated with COVID-19 severity and mortality - Multiethnic cohort study in Southern California. *Environ Int*. 2021;157:106862. **[Crossref]**

6. Aykaç N, Etiler N. COVID-19 mortality in Istanbul in association with air pollution and socioeconomic status: an ecological study. *Environ Sci Pollut Res Int*. 2021;29(9):13700-13708. **[Crossref]**
7. Clouston SAP, Nataleb G, Link B. Socioeconomic inequalities in the spread of coronavirus-19 in the United States: a examination of the emergence of social inequalities. *Soc Sci Med*. 2020;268:113554. **[Crossref]**
8. Khalatbari-Soltani S, Cumming RC, Delpierre C, Kelly-Irving M. Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *J Epidemiol Community Health*. 2020;74(8):620-623. **[Crossref]**
9. Larasati A, Handayani W, Febriarta E. Building density and its implications to COVID-19 health risk management: an example from Yogyakarta, Indonesia. *IOP Conf Ser Earth Environ Sci*. 2022;1039(1):012019. **[Crossref]**
10. Hong B, Bonczak BJ, Gupta A, et al. Exposure density and neighborhood disparities in COVID-19 infection risk. *Proc Natl Acad Sci U S A*. 2021;118(13):e2021258118. **[Crossref]**
11. Ministry of Health. Death notification system circular. Accessed July 23, 2025. **[Crossref]**
12. Turkish Statistical Institute. Address-based population registration system. Accessed July 23, 2025. **[Crossref]**
13. General Directorate of Development Agencies. Socioeconomic Development of Districts Ranking Study. İlçe SEGE-2022. Ministry of Industry and Technology. Accessed December 21, 2024. **[Crossref]**
14. City of Ankara Clean Air Action Plan. THEP 2020-2024. Accessed July 23, 2025. **[Crossref]**
15. North Central Anatolia Clean Air Regional Directorate. Results of the Air Quality Measurements. Kuzey İç Anadolu Temiz Hava Merkezi Müdürlüğü. Accessed July 23, 2025. **[Crossref]**
16. SİM (Sürekli İzleme Merkezi) | T.C. Çevre, Şehircilik ve İklim Değişikliği Bakanlığı [Internet]. Accessed July 23, 2025. **[Crossref]**
17. Ankara Sanayi Odası. Accessed July 23, 2025. **[Crossref]**
18. Bianconi V, Bronzo P, Banach M, Sahebkar A, Mannarino MR, Pirro M. Particulate matter pollution and the COVID-19 outbreak: results from Italian regions and provinces. *Arch Med Sci*. 2020;16(5):985-992. **[Crossref]**
19. Kim H, Samet JM, Bell ML. Association between short-term exposure to air pollution and COVID-19 mortality: a population-based case-crossover study using individual-level mortality registry confirmed by medical examiners. *Environ Health Perspect*. 2022;130(11):117006. **[Crossref]**
20. Beloconi A, Vounatsou P. Long-term air pollution exposure and COVID-19 case-severity: an analysis of individual-level data from Switzerland. *Environ Res*. 2023;216:114481. **[Crossref]**
21. Zang ST, Luan J, Li L, et al. Ambient air pollution and COVID-19 risk: evidence from 35 observational studies. *Environ Res*. 2022;204(Pt B):112065. **[Crossref]**
22. Quan D, Luna Wong L, Shallal A, et al. Impact of race and socioeconomic status on outcomes in patients hospitalized with COVID-19. *J Gen Intern Med*. 2021;36(5):1302-1309. **[Crossref]**
23. Politi J, Martín-Sánchez M, Mercuriali L, et al. Epidemiological characteristics and outcomes of COVID-19 cases: mortality inequalities by socio-economic status, Barcelona, Spain, 24 February to 4 May 2020. *Euro Surveill*. 2021;26(20):2001138. **[Crossref]**
24. Chen JT, Krieger N. Revealing the unequal burden of COVID-19 by income, race/ethnicity, and household crowding: US county versus zip code analyses. *J Public Health Manag Pract*. 2021;27(Suppl 1):43-56. **[Crossref]**
25. Niedzwiedz CL, O'Donnell CA, Jani BD, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. *BMC Med*. 2020;18(1):160. **[Crossref]**
26. Magesh S, John D, Li WT, et al. Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a systematic-review and meta-analysis. *JAMA Netw Open*. 2021;4(11):e2134147. **[Crossref]** Erratum in: *JAMA Netw Open*. 2021;4(12):e2144237. Erratum in: *JAMA Netw Open*. 2022;5(2):e222170. **[Crossref]**
27. Samuel LJ, Gaskin DJ, Trujillo AJ, Szanton SL, Samuel A, Slade E. Race, ethnicity, poverty and the social determinants of the coronavirus divide: U.S. county-level disparities and risk factors. *BMC Public Health*. 2021;21(1):1250. **[Crossref]**
28. Huang Q, Jackson S, Derakhshan S, et al. Urban-rural differences in COVID-19 exposures and outcomes in the South: a preliminary analysis of South Carolina. *PLoS One*. 2021;16(2):e0246548. **[Crossref]**
29. Baidal JW, Wang AY, Zumwalt K, et al. Social determinants of health and COVID-19 among patients in New York City. *Res Sq*. 2020:rs.3.rs70959. **[Crossref]**
30. Loignon C, Hudon C, Goulet É, et al. Perceived barriers to healthcare for persons living in poverty in Quebec, Canada: the EQUIhealThY project. *Int J Equity Health*. 2015;14:4. **[Crossref]**
31. George S, Daniels K, Fioratou E. A qualitative study into the perceived barriers of accessing healthcare among a vulnerable population involved with a community centre in Romania. *Int J Equity Health*. 2018;17(1):41. **[Crossref]**
32. Carmichael C, Schiffler T, Smith L, et al. Barriers and facilitators to health care access for people experiencing homelessness in four European countries: an exploratory qualitative study. *Int J Equity Health*. 2023;22(1):206. **[Crossref]**
33. Sevik I, Ciceklioglu M. Healthcare access worsened for women in precarious housing during the COVID-19 pandemic: a qualitative study. *Inquiry*. 2024;61:469580241246478. **[Crossref]**
34. Daras K, Alexiou A, Rose TC, Buchan I, Taylor-Robinson D, Barr B. How does vulnerability to COVID-19 vary between communities in England? Developing a small area vulnerability index (SAVI). *J Epidemiol Community Health*. 2021;75(8):729-734. **[Crossref]**
35. Varshney K, Glodjo T, Adalbert J. Overcrowded housing increases risk for COVID-19 mortality: an ecological study. *BMC Res Notes*. 2022;15(1):126. **[Crossref]**
36. Karmakar M, Lantz PM, Tipirneni R. Association of social and demographic factors with COVID-19 incidence and death rates in the US. *JAMA Netw Open*. 2021;4(1):e2036462. **[Crossref]**
37. Wang L, Calzavara A, Baral S, et al. Differential patterns by area-level social determinants of health in coronavirus disease 2019 (COVID-19)-related mortality and non-COVID-19 mortality: a population-based study of 11.8 million people in Ontario, Canada. *Clin Infect Dis*. 2023;76(6):1110-1120. **[Crossref]**

Thorac Res Pract. 2026;27(4):215-221 **DOI:** 10.4274/ThoracResPract.2025.2025-10-10

Original Article

Emphysema or Bronchiectasis with Pulmonary Nodules: Impact on The Risk of Malignancy

Ayşegül Gencer, Ersan Atahan, Bilun Gemicioğlu, Şermin Börekçi, Buket Çalışkaner Öztürk

Department of Pulmonary Diseases, İstanbul University-Cerrahpaşa, Cerrahpaşa Faculty of Medicine, İstanbul, Türkiye

Cite this article as: Gencer A, Atahan E, Gemicioğlu B, Börekçi Ş, Çalışkaner Öztürk B. Emphysema or bronchiectasis with pulmonary nodules: impact on the risk of malignancy. *Thorac Res Pract.* 2026;27(4):215-221

ABSTRACT

OBJECTIVE: The increasing use of computed tomography (CT) has led to a significant rise in the detection of pulmonary nodules. This has resulted in more studies aimed at identifying risk factors associated with increased detection of early-stage lung cancer. The aim of this study is to investigate the effect of coexisting emphysema or bronchiectasis on the incidence of malignancy in pulmonary nodules.

MATERIAL AND METHODS: The study included 212 patients with pulmonary nodules detected on CT images. They were divided into three groups based on the presence of emphysema or bronchiectasis. The effects of demographic factors and pulmonary nodule characteristics on malignancy were evaluated in these groups.

RESULTS: Comparison of the incidence of malignancy in pulmonary nodules across groups showed no significant difference in the emphysema or bronchiectasis groups compared with the control group ($P > 0.05$).

CONCLUSION: Contrary to what is frequently reported in the literature, the presence of emphysema was not found to be a risk factor for lung cancer in this study. The presence of bronchiectasis was also found not to be a risk factor; however, there are insufficient data in the literature on this point.

KEYWORDS: Pulmonary nodule with emphysema, pulmonary nodule with bronchiectasis, lung cancer risk factors, lung cancer screening

Received: 28.10.2025	***Revision Requested:*** 24.11.2025	***Last Revision Received:*** 22.12.2025
Accepted: 29.12.2025	***Epub:*** 02.03.2026	***Publication Date:*** 24.07.2026

INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths worldwide. As a result of studies conducted to enable early diagnosis, lung cancer screening programs have been developed.[1] Risk-based lung cancer screening programs evaluate patient characteristics such as age, smoking history, gender, ethnic origin, personal and family history of cancer, and presence of emphysema.[2] Risk models based on characteristics of pulmonary nodules detected on computed tomography (CT) scans have also been developed to predict malignancy.[3] The criteria in these models include large nodule size, upper lobe location, and irregular margins.[4]

More than 85% of patients diagnosed with lung cancer are smokers.[5] A history of smoking is also the most common cause of emphysema. Lung cancer and emphysema share this common primary risk factor.[6] Despite this common risk factor, studies investigating the relationship between these two diseases have reported conflicting results. While some studies have shown that the detection of emphysema on CT scans is a risk factor for lung cancer,[7,8] another study failed to find this correlation.[9]

Chronic inflammation and infection play important roles in the pathophysiology of lung cancer.[10] Bronchiectasis is one of the pulmonary diseases associated with chronic infection and systemic inflammation; few studies have evaluated whether it is a risk factor for lung cancer.[11]

Corresponding author: Ayşegül Gencer, MD, e-mail: aysegulbozkurtgencer@gmail.com

Conflicting data exist regarding the effect of pulmonary emphysema on the incidence of malignancy among patients with pulmonary nodules. The aim of this study is to investigate the effects of emphysema or bronchiectasis coexisting with pulmonary nodules, and of nodule characteristics, on the incidence of malignancy in the nodules.

MATERIAL AND METHODS

Study Design and Setting

The study was designed as a retrospective study. The study included patients who presented to the pulmonary diseases outpatient clinic between March 1, 2018, and March 1, 2022; had a pulmonary nodule detected on chest X-ray; underwent thoracic CT that showed a pulmonary nodule measuring 8–30 mm; and had persistent nodules despite non-specific antibiotic therapy. Patients with pulmonary nodules were grouped into three categories: those with emphysema (emphysema group), those with bronchiectasis (bronchiectasis group), and those in a control group. Patients with concurrent emphysema and bronchiectasis were excluded from the study because the number of such cases was small.

This study was approved by the İstanbul University-Cerrahpaşa Clinical Research Ethics Committee on May 16, 2022 (no: E83045809-604.01.01-381251).

Participants

Inclusion criteria:

- Between March 1, 2018, and March 1, 2022, a non-calcified solid pulmonary nodule was detected on a chest CT scan performed at the outpatient chest diseases clinic after a nodule was identified on a chest X-ray. The nodule persisted on a follow-up CT scan performed two months after non-specific antibiotic therapy
- Pulmonary nodule measuring 8–30 mm in size with further investigations performed to assess malignancy
- Aged between 30 and 90
- Detection of a pulmonary nodule measuring 8–30 mm on the initial CT scan
- Completion of a 2-year CT follow-up program from the date of nodule detection or diagnosis of malignancy during follow-up

Main Points

- In this study, contrary to reports in the literature, it was found that the presence of emphysema with pulmonary nodules had no significant effect on the incidence of malignancy and was not a risk factor for lung cancer.
- Bronchiectasis has not been identified as a risk factor with this study.
- The incidence of malignancy in pulmonary nodules of 8 mm or larger in the emphysema group was not different from that in the control group.

Exclusion criteria:

- Presence of emphysema and bronchiectasis
- Presence of regression after non-specific antibiotics
- Failure to complete 2 years of follow-up during nodule monitoring
- History of lung cancer or any other malignancy
- Presence of active lung infection
- Diagnosis or suspicion of interstitial lung disease
- Pulmonary vascular disease
- Pleural diseases, pleural effusion
- Alveolar hemorrhage
- Pregnancy
- Chronic diseases such as heart failure, kidney failure
- History of known rheumatological diseases

Variables and Data Collection

Demographic data(age, gender), family history, and smoking status (pack-years) were recorded from the patients' files. Pulmonary nodule characteristics and the presence of coexisting emphysema or bronchiectasis were recorded from chest CT scans.

Thorax CT Scans

Data on the single largest pulmonary nodule detected on the initial and follow-up lung parenchymal-window CT scans were included in the study. Each nodule that was regularly monitored according to the Fleischner guidelines and remained unchanged for at least 2 years was considered benign and included in the dataset. Data on patients who received a histopathological diagnosis of malignancy during follow-up were added. The presence of emphysema or bronchiectasis was recorded. Patients with emphysema were included in the emphysema group, those with bronchiectasis were included in the bronchiectasis group, and those with both conditions were excluded from the study. Patients without emphysema or bronchiectasis were included in the control group.

The size, location, and margin characteristics of pulmonary nodules were recorded. Nodules diagnosed as malignant during the 2-year follow-up period are included, along with their size at the time of diagnosis. Location data were recorded as peripheral or central. Nodules defined as central were located in the segmental bronchi or in more proximal airways, whereas peripheral nodules were located distally in the subsegmental bronchi.[12]

Diagnosis of Malignancy

Patients diagnosed with lung cancer by histopathological examination of biopsy samples were classified according to the World Health Organization classification of lung neoplasms.[13]

Study Size

All patients who met the eligibility criteria and provided informed consent were included in the study.

Statistical Analysis

All analyses were performed using the Statistical Package for the Social Sciences (SPSS) v22 (SPSS Inc., Chicago, IL, USA). Data were presented as mean ± standard deviation or median (minimum-maximum) for continuous variables and frequency (percentage) for categorical variables, depending on the normality of the distribution, and descriptive statistics were used in the analysis of the data. The chi-square test was used to compare categorical data. For normally distributed data, one-way analysis of variance was used; for non-normally distributed data, the Kruskal-Wallis test and the Mann-Whitney U test were used. Two-tailed *P* values less than 0.05 were considered statistically significant.

RESULTS

Of the 212 patients included in the study, 36.8% (n = 78) had emphysema, 24.5% (n = 52) had bronchiectasis, and 38.7% (n = 82) were in the control group. The mean age of the patients was 67.39±10.90, and the mean number of cigarettes smoked was 48.33±30.60 pack-years. 75.5% of the patients were male, 84.9% had a history of smoking, 80.7% had a family history of cancer, 0.5% had asbestos exposure, and 14.6% had a history of tuberculosis.

When comparing the variables of gender, smoking history, and tuberculosis history in patients with emphysema, bronchiectasis, and the control group, there were significant differences between the three groups (*P* = 0.000, *P* = 0.000, *P* = 0.005, respectively); 45.6% of male patients, 52.1% of active smokers, and 48.4% of patients with a history of tuberculosis were in the emphysema group. In addition, the average number of pack-years smoked by patients in the emphysema group was significantly higher (*P* = 0.000) (Table 1).

The characteristics of all pulmonary nodules, benign and malignant, were evaluated with respect to the presence of bronchiectasis, and no significant differences were found between the bronchiectasis and control groups (*P* > 0.05; Table 2).

The relationship between the size of all pulmonary nodules (benign and malignant) and the presence of emphysema was evaluated; no significant difference was found between nodules with and without emphysema (*P* > 0.05) (Table 3). When the relationship between the presence of emphysema and the edge characteristics and location of pulmonary nodules was evaluated, differences were observed between the two groups (*P* = 0.007 and *P* = 0.042, respectively). While nodules with smooth margins were frequently observed in the group without emphysema, nodules with irregular margins were more frequent in the emphysema group (odds ratio: 2.462; 95% confidence interval: 1.270–4.772). The likelihood of pulmonary nodules with irregular margins was 2.46 times higher in the emphysema group than in the non-emphysema group. 66.7% of centrally located pulmonary nodules were found in the emphysema group (odds ratio: 2.433; 95% confidence interval: 1.019–5.808). The likelihood of a pulmonary nodule being centrally located was 2.43 times higher in the presence of emphysema than in its absence (*P* = 0.041; Table 3).

Malignant pulmonary nodules are summarized in Table 4; no statistically significant differences were found between the emphysema, bronchiectasis, and control groups with respect to nodule size, margin characteristics, location, and histopathological type (*P* > 0.05). It was determined that 45.6% of malignant pulmonary nodules were >31 mm in size, 87.7% had irregular margins, and 63.2% were peripherally

Table 1. Demographic characteristics of patients (n = 212)

Demographic characteristics	Emphysema group (n = 78) Mean ± SD	Bronchiectasis group (n = 52) Mean ± SD	Control group (c) (n = 82) Mean ± SD	Total Mean ± SD	*P*
Age	69.44±8.44	67.00±11.58	65.69±12.25	67.39±10.90	*P* = 0.131
Amount of cigarettes smoked (pack-year) (pack-years) (n = 180)	57.42±28.21	47.23±31.75	37.80±29.58	48.33±30.60	***P* = 0.000*** a>c†
	n (%)	**n (%)**	**n (%)**	**n (%)**	
Gender					
Female	5 (9.6)	11 (21.2)	36 (69.2)	52 (24.5)	***P* = 0.000***
Male	73 (45.6)	41 (25.6)	46 (28.8)	160 (75.5)	
Smoking status					
Never smoked	2 (6.3)	9 (28.1)	21 (65.6)	32 (15.1)	
Ex-smoker	27 (31.4)	26 (30.2)	33 (38.4)	86 (40.6)	***P* = 0.000***
Smoker	49 (52.1)	17 (18.1)	28 (29.8)	94 (44.3)	
Family history of cancer	10 (24.4)	9 (22.0)	22 (53.7)	41 (19.3)	*P* = 0.074
Asbestos exposure	0 (0)	0 (0)	1 (100.0)	1 (0.5)	*P* = 0.451
Tuberculosis history	15 (48.4)	12 (38.7)	4 (12.9)	31 (14.6)	***P* = 0.005***

*$P < 0.05$, †The *P* value obtained after Bonferroni correction was $P < 0.0167$. The chi-square test was used for categorical variables. Since numerical variables did not show a normal distribution, the Kruskal-Wallis test was used. Note: Row percentage (row percent) and column percentage (column percent) are given for groups
SD: standard deviation

located. Comparison of the margin characteristics of malignant pulmonary nodules between the emphysema and control groups revealed a significant difference (P = 0.039). 60% of pulmonary nodules with irregular margins occurred in the emphysema group (odds ratio: 9.000; 95% confidence interval: 1.006–80.522). The likelihood of a malignant pulmonary nodule with irregular margins was 9 times higher in the presence of emphysema than in its absence.

DISCUSSION

This study aimed to determine the incidence of malignancy among pulmonary nodules in patients with emphysema or bronchiectasis. We found that the frequency of malignancy in pulmonary nodules in the presence of emphysema did not differ from that in the control group; emphysema, often identified in the literature as a risk factor for lung cancer, was not found to be a risk factor in this study. Similarly, bronchiectasis was not identified as a risk factor for lung cancer.

The literature reports conflicting results regarding whether the association between emphysema and pulmonary nodules contributes to nodule malignancy risk. Wachuła et al.[14] showed that pulmonary nodules were more frequently associated with active smoking among patients with emphysema, but they did not find that malignant pulmonary nodules were more common among patients with emphysema. While the results of this study support our data, other studies have reported that, among pulmonary nodules in patients with emphysema, malignant nodules are more common than benign nodules.[15,16] Hohberger

Table 2. Comparison of pulmonary nodule characteristics according to the presence of emphysema (n = 160)

Nodule characteristics	Emphysema		Total	*P*
	Presence (n = 78) Mean ± SD	Absence (n = 82) Mean ± SD	Mean ± SD	
Pulmonary nodule size	25.48±21.30	20.84±14.65	23.10±18.29	*P* = 0.470
	n (%)	n (%)	n (%)	
Margin characteristics				
Smooth	21 (35.0)	39 (65.0)	60 (37.5)	***P* = 0.007***
Irregular	57 (57.0)	43 (43.0)	100 (62.5)	
Location				
Peripheral	60 (45.1)	73 (54.9)	133 (83.1)	***P* = 0.041***
Near the hilum	18 (66.7)	9 (33.3)	27 (16.9)	
Malignancy status				
Benign	47 (45.6)	56 (54.4)	103 (64.4)	*P* = 0.289
Malignant	31 (54.4)	26 (45.6)	57 (35.6)	

P < 0.05, the chi-square test was used for categorical variables. Since numerical variables did not show a normal distribution, the Mann-Whitney U test was used.
Note: Row percentages were given for groups, and column percentages were given for the total
SD: standard deviation

Table 3. Comparison of pulmonary nodule characteristics according to the presence of bronchiectasis (n = 134)

Nodule characteristics	Bronchiectasis		Total	*P*
	Presence (n = 52) Mean ± SD	Absence (n = 82) Mean ± SD	Mean ± SD	
Pulmonary nodule size	19.27±14.75	20.84±14.65	20.23±14.65	*P* = 0.545
	n (%)	n (%)	n (%)	
Margin characteristics				
Smooth	22 (36.1)	39 (63.9)	61 (45.5)	*P* = 0.552
Irregular	30 (41.1)	43 (58.9)	73 (54.5)	
Location				
Peripheral	48 (39.7)	73 (60.3)	121 (90.3)	*P* = 0.531
Near the hilum	4 (30.8)	9 (69.2)	13 (9.7)	
Malignancy status				
Benign	42 (42.9)	56 (57.1)	98 (73.1)	*P* = 0.112
Malignant	10 (27.8)	26 (72.2)	36 (26.9)	

P < 0.05, the chi-square test was used for categorical variables. Since numerical variables did not show a normal distribution, the Mann-Whitney U test was used.
Note: Row percentages were given for groups, and column percentages were given for the total
SD: standard deviation

Table 4. Comparison of malignant pulmonary nodule characteristics between groups (n = 67)

Malignant nodule characteristics	Emphysema group (a) (n = 78) n (%)	Bronchiectasis group (b) (n = 52) n (%)	Control group (n=82) n (%)	Total n (%)	*P*
Margin characteristics					
Smooth	1 (11.1)	2 (22.2)	6 (66.7)	9 (13.4)	*P* = 0.073
Irregular	30 (51.7)	8 (13.8)	20 (34.5)	58 (86.6)	
Location					
Peripheral	17 (38.6)	8 (18.2)	19 (43.2)	44 (65.7)	*P* = 0.206
Near the hilum	14 (60.9)	2 (8.7)	7 (30.4)	23 (34.3)	
Histopathological type					
Adenocarcinoma	12 (40.0)	6 (20.0)	12 (40.0)	30 (44.8)	*P* = 0.843
Squamous cell carcinoma	11 (55.0)	2 (10.0)	7 (35.0)	20 (29.9)	
Small cell carcinoma	3 (50.0)	0 (0)	3 (50.0)	6 (9.0)	
Metastasis	5 (45.5)	2 (18.2)	4 (36.4)	11 (16.4)	

P < 0.05, the chi-square test was used for categorical variables. Since numerical variables did not show a normal distribution, the Mann-Whitney U test was used. Note: Row percentages were given for groups, and column percentages were given for the total

et al.[15] evaluated the relationship between emphysema severity and tumors in patients with lung cancer, and found that primary lung cancer developed more frequently in lung areas with severe emphysema.

The National Lung Screening Trial is one of the studies that have examined the relationship between pulmonary nodules and emphysema severity. Groups of patients with benign and malignant pulmonary nodules, mostly comprising individuals with emphysema who had similar smoking histories, were compared based on emphysema severity. Only severe emphysema was significantly associated with malignant pulmonary nodules; however, this association was not compared with a control group. The data also showed that the malignant pulmonary nodule group had a significantly higher incidence of diagnosed chronic obstructive pulmonary disease.[16] In our study, however, a comparison was made with a control group without emphysema, and the smoking amount in this control group was significantly lower than that in the emphysema group. In this respect, it is consistent with the data of Wachuła et al.[14]

Sanchez-Salcedo et al.[17] investigated the presence of emphysema in patients diagnosed with malignancy and demonstrated a association with lung cancer incidence; however, this association was not examined in relation to pulmonary nodules. A similar study by Henschke et al.[18] reported a significantly higher incidence of emphysema among patients diagnosed with lung cancer. However, the contribution of emphysema to the risk of malignancy in a pulmonary nodule was not evaluated in these studies. In our study, the relationship between emphysema and malignancy was evaluated based on pulmonary nodules.

Despite these studies, other studies did not find emphysema to be a significant risk factor for lung cancer. Maldonado et al.[19] found no significant association between severe emphysema and lung cancer. In our study, patients diagnosed with lung cancer were not evaluated for emphysema. Patients with pulmonary nodules were evaluated for emphysema, and the frequency of malignancy in pulmonary nodules associated with emphysema did not differ from that in the control group.

These conflicting results may be explained by interindividual variation in the biological damage caused by smoking; only a proportion of smokers develop emphysema, and the effect of pulmonary nodules on malignancy risk also varies.[20]

Worldwide, there are still more than one billion smokers, the vast majority of whom are men.[21,22] Chronic smoking, the primary cause of lung cancer and emphysema, is a common risk factor for these diseases and is more prevalent among men.[23-26] In line with these findings, the majority of our patients who smoked, as well as those with emphysema, were men. Tuberculosis has been identified as a risk factor for both bronchiectasis and emphysema.[27] We also found that patients with a history of tuberculosis frequently had emphysema.

Marginal characteristics of pulmonary nodules are among the parameters examined as risk factors for malignancy.[4,28] Upper-lobe localization is considered a predictor of lung cancer.[4,29] In this study, the incidence of irregular-edged pulmonary nodules in emphysematous lung parenchyma was 2.46 times higher, while the incidence of central localization of nodules was 2.43 times higher.

In the final model of their study evaluating the risk of malignancy associated with irregular margins in pulmonary nodules, McWilliams et al.[4] found no association between margin characteristics and risk of malignancy. The Dutch-Belgian randomized lung cancer screening trial (NELSON) reported similar findings: among pulmonary nodules with irregular margins, only the initial volume predicted malignancy, whereas morphology did not.[28] In our study, pulmonary nodules with irregular margins were more common in patients with emphysema; however, their effect on malignancy risk was assessed independently of emphysema.

Conflicting data exist on whether emphysema is a risk factor for lung cancer; however, our study did not identify emphysema as a risk factor.[14-19] However, irregularly shaped pulmonary

nodules were more common in the emphysema group. Data on the relationship between pulmonary nodule margin characteristics and malignancy risk are also variable.[4,28] We have not found any studies that evaluate the risk of malignancy in pulmonary nodules with irregular margins in patients with emphysema.

In a study by Sanchez-Carpintero Abad et al.[30] examining the role of bronchiectasis in lung cancer screening programs, the presence of bronchiectasis was associated with increased numbers of nodules and false-positive results. It has been emphasized that, although bronchiectasis is more common in smokers and therefore more common among patients included in lung cancer screening programs than in the general population, it has no effect on cancer incidence.[30] Similar findings were reported in the study by Cai et al.[31] In our study, we also found that the coexistence of bronchiectasis and pulmonary nodules did not increase the incidence of malignancy.

Study Limitations

The limitations of this study include that it was not prospective. Although a sample size appropriate for power analysis was obtained, a larger number of patients could have been included to allow for evaluation based on the severity of emphysema. Additionally, this suggests that an increased number of cases is associated with a higher risk of malignancy in emphysema cases and a lower risk of malignancy in nodules in bronchiectasis cases. The gender and smoking-status variables of the control and emphysema groups could have been made more similar. Since routine follow-up is not required for patients with pulmonary nodules smaller than 6 mm who do not have lung cancer risk factors, the absence of these patients in the control group may have introduced bias. The inclusion of metastatic nodules in the study was another limitation.

Strengths of the study include comparison of the emphysema group with the bronchiectasis group and with a control group. It investigated the effect of nodule-related characteristics on malignancy in the emphysema group. This study adds a contrary perspective to studies reporting emphysema as a risk factor for lung cancer and highlights the need for further research from a new perspective. It was a unique study in that it investigated the frequency of malignancy in pulmonary nodules in patients with bronchiectasis.

CONCLUSION

In conclusion, contrary to frequent reports in the literature, emphysema has not been identified as a risk factor for lung cancer among patients with pulmonary nodules. Bronchiectasis, which has not previously been investigated in this context, was also not identified as a risk factor in our study.

Ethics

Ethics Committee Approval: This study was approved by the İstanbul University-Cerrahpaşa Clinical Research Ethics Committee on May 16, 2022 (no: E83045809-604.01.01-381251).

Informed Consent: All patients who met the eligibility criteria and provided informed consent were included in the study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.G., B.Ç.Ö., Concept: A.G., B.Ç.Ö., Design: A.G., B.Ç.Ö., Data Collection or Processing: A.G., B.Ç.Ö., Analysis or Interpretation: A.G., E.A., B.G., Ş.B., B.Ç.Ö., Literature Search: A.G., Writing: A.G.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Wu Z, Wang F, Cao W, et al. Lung cancer risk prediction models based on pulmonary nodules: a systematic review. *Thorac Cancer*. 2022;13(5):664-677. **[Crossref]**
2. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365:395-409. **[Crossref]**
3. Toumazis I, Bastani M, Han SS, Plevritis SK. Risk-based lung cancer screening: a systematic review. *Lung Cancer*. 2020;147:154-186. **[Crossref]**
4. McWilliams A, Tammemagi MC, Mayo JR, et al. Probability of cancer in pulmonary nodules detected on first screening CT. *N Engl J Med*. 2013;369(10):910-919. **[Crossref]**
5. Smith BM, Pinto L, Ezer N, Sverzellati N, Muro S, Schwartzman K. Emphysema detected on computed tomography and risk of lung cancer: a systematic review and meta-analysis. *Lung Cancer*. 2012;77(1):58-63. **[Crossref]**
6. Mouronte-Roibás C, Leiro-Fernández V, Fernández-Villar A, Botana-Rial M, Ramos-Hernández C, Ruano-Ravina A. COPD, emphysema and the onset of lung cancer. A systematic review. *Cancer Lett*. 2016;382(2):240-244. **[Crossref]**
7. Loganathan RS, Stover DE, Shi W, Venkatraman E. Prevalence of COPD in women compared to men around the time of diagnosis of primary lung cancer. *Chest*. 2006;129(5):1305-1312. **[Crossref]**
8. Congleton J, Muers MF. The incidence of airflow obstruction in bronchial carcinoma, its relation to breathlessness, and response to bronchodilator therapy. *Respir Med*. 1995;89(4):291-296. **[Crossref]**
9. Strange C. COPD and lung cancer: are they both airways diseases? *Chest*. 2010;138(6):1289-1290. **[Crossref]**
10. Gomes M, Teixeira AL, Coelho A, Araújo A, Medeiros R. The role of inflammation in lung cancer. *Adv Exp Med Biol*. 2014;816:1-23. **[Crossref]**
11. Chung WS, Lin CL, Hsu WH, Kao CH. Increased risk of lung cancer among patients with bronchiectasis: a nationwide cohort study. QJM. 2016;109(1):17-25. **[Crossref]**
12. Bandoh S, Fujita J, Ueda Y, et al. Expression of carcinoembryonic antigen in peripheral- or central-located small cell lung cancer: its clinical significance. *Jpn J Clin Oncol*. 2001;31(7):305-310. **[Crossref]**
13. Nicholson AG, Tsao MS, Beasley MB, et al. The 2021 WHO Classification of Lung Tumors: Impact of Advances Since 2015. *J Thorac Oncol*. 2022;17(3):362-387. **[Crossref]**
14. Wachuła E, Szabłowska-Siwik S, Czyżewski D, Kozielski J, Adamek M. Emphysema affects the number and characteristics of solitary pulmonary nodules identified by chest low-dose computed tomography. A study on screenees with high-risk lung cancer

recruited in Upper Silesia. *Pol Arch Intern Med*. 2020;130(1):17-24. **[Crossref]**

15. Hohberger LA, Schroeder DR, Bartholmai BJ, et al. Correlation of regional emphysema and lung cancer: a lung tissue research consortium-based study. *J Thorac Oncol*. 2014;9(5):639-645. **[Crossref]**
16. Liu Y, Wang H, Li Q, et al. Radiologic features of small pulmonary nodules and lung cancer risk in the National Lung Screening Trial: a nested case-control study. *Radiology*. 2018;286(1):298-306. **[Crossref]**
17. Sanchez-Salcedo P, Berto J, de-Torres JP, et al. Lung cancer screening: fourteen year experience of the Pamplona early detection program (P-IELCAP). *Arch Bronconeumol*. 2015;51(4):169-176. English, Spanish. **[Crossref]**
18. Henschke CI, Yankelevitz DF, Mirtcheva R, McGuinness G, McCauley D, Miettinen OS; ELCAP Group. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *AJR Am J Roentgenol*. 2002;178(5):1053-1057. **[Crossref]**
19. Maldonado F, Bartholmai BJ, Swensen SJ, Midthun DE, Decker PA, Jett JR. Are airflow obstruction and radiographic evidence of emphysema risk factors for lung cancer? A nested case-control study using quantitative emphysema analysis. *Chest*. 2010;138(6):1295-1302. **[Crossref]**
20. Ueda K, Jinbo M, Li TS, Yagi T, Suga K, Hamano K. Computed tomography-diagnosed emphysema, not airway obstruction, is associated with the prognostic outcome of early-stage lung cancer. *Clin Cancer Res*. 2006;12(22):6730-6736. **[Crossref]**
21. Suraweera W, Warrell D, Whitaker R, et al. Trends in snakebite deaths in India from 2000 to 2019 in a nationally representative mortality study. *Elife*. 2020;9:e54076. **[Crossref]**
22. Le Foll B, Piper ME, Fowler CD, et al. Tobacco and nicotine use. *Nat Rev Dis Primers*. 2022;8(1):19.
23. Grydeland TB, Dirksen A, Coxson HO, et al. Quantitative computed tomography: emphysema and airway wall thickness by sex, age and smoking. *Eur Respir J*. 2009;34(4):858-865. **[Crossref]**
24. Martinez FJ, Curtis JL, Sciurba F, et al. Sex differences in severe pulmonary emphysema. *Am J Respir Crit Care Med*. 2007;176(3):243-252. **[Crossref]**
25. Dransfield MT, Washko GR, Foreman MG, Estepar RS, Reilly J, Bailey WC. Gender differences in the severity of CT emphysema in COPD. *Chest*. 2007;132(2):464-470. **[Crossref]**
26. Auerbach O, Hammond EC, Garfinkel L, Benante C. Relation of smoking and age to emphysema. Whole-lung section study. *N Engl J Med*. 1972;286(16):853-857. **[Crossref]**
27. Jin J, Li S, Yu W, Liu X, Sun Y. Emphysema and bronchiectasis in COPD patients with previous pulmonary tuberculosis: computed tomography features and clinical implications. *Int J Chron Obstruct Pulmon Dis*. 2018;13:375-384. **[Crossref]**
28. Xu DM, van der Zaag-Loonen HJ, Oudkerk M, et al. Smooth or attached solid indeterminate nodules detected at baseline CT screening in the NELSON study: cancer risk during 1 year of follow-up. *Radiology*. 2009;250(1):264-272. **[Crossref]**
29. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of ıncidental pulmonary nodules detected on CT ımages: from the Fleischner Society 2017. *Radiology*. 2017;284(1):228-243. **[Crossref]**
30. Sanchez-Carpintero Abad M, Sanchez-Salcedo P, de-Torres JP, et al. Prevalence and burden of bronchiectasis in a lung cancer screening program. *PLoS One*. 2020;15(4):e0231204. **[Crossref]**
31. Cai Q, Triphuridet N, Zhu Y, et al. Bronchiectasis in Low-Dose CT screening for lung cancer. *Radiology*. 2022;304(2):437-447. **[Crossref]**

Thorac Res Pract. 2026;27(4):222-228 **DOI:** 10.4274/ThoracResPract.2026.2025-11-1

Original Article

Endobronchial Ultrasonography (EBUS) with Lymph Node Aspiration: On-site Evaluation and Final Pathology Correlation

Ferhan Karataş[1], Makbule Çisel Aydın Meriçöz[2], Işıl Uzel[1], Tuğçe Yakut[3], Emine Karabul[1], Öner Dikensoy[1]

[1]Department of Pulmonology, Koç University Hospital, İstanbul, Türkiye
[2]Department of Pathology, Koç University Hospital, İstanbul, Türkiye
[3]Department of Allergy and Immunology, Koç University Hospital, İstanbul, Türkiye

Cite this article as: Karataş F, Aydın Meriçöz MÇ, Uzel I, Yakut T, Karabul E, Dikensoy Ö. Endobronchial ultrasonography (EBUS) with lymph node aspiration: on-site evaluation and final pathology correlation. *Thorac Res Pract.* 2026;27(4):222-228

ABSTRACT

OBJECTIVE: Rapid on-site evaluation (ROSE) during endobronchial ultrasound (EBUS)-guided transbronchial needle aspiration (TBNA) may improve sample adequacy and expedite clinical decision-making; however, its impact on diagnostic performance and concordance with final pathology remains unclear.

MATERIAL AND METHODS: We retrospectively analyzed consecutive adult patients who underwent EBUS-TBNA between February 2023 and January 2024. ROSE adequacy, final diagnostic adequacy, and ROSE-final pathology concordance were calculated. Factors associated with these outcomes were assessed using logistic regression analysis.

RESULTS: Ninety-five patients (mean age 65±13.8 years; 62.2% male) and 203 targets were sampled with 683 aspirations (3.4±1.8 passes/station). Based on lymph node (LN) level ROSE sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were 76.6%, 100.0%, 100.0%, 85.4%, and 90.1%, respectively. The adequacy rates were 89.7% for ROSE and 90.6% for final pathology, and the ROSE-final pathology concordance was 88.2%. Diagnostic discordance occurred in 11.8% of the nodes, most of which were related to limitations in tumor subtyping. A greater number of needle passes was positively associated with ROSE adequacy ($P = 0.038$) and final adequacy ($P = 0.021$), whereas a larger LN diameter favored final adequacy ($P = 0.045$). Larger LNs and greater experience of the final-diagnosis pathologist were associated with lower ROSE-final diagnosis concordance ($P = 0.026$ and $P = 0.015$, respectively).

CONCLUSION: ROSE during EBUS-TBNA yielded high sensitivity and overall accuracy, demonstrated strong sample adequacy, and showed substantial concordance with the final pathology. Optimizing the number of passes and considering LN size can enhance adequacy, and recognizing the limitations of cytological subtyping and potential observer variability is essential for multidisciplinary decision-making.

KEYWORDS: Endobronchial ultrasonography, lymph node aspiration, rapid on-site evaluation, final pathology correlation

Received: 03.11.2025 ***Revision Requested:*** 10.02.2026 ***Last Revision Received:*** 25.02.2026

Accepted: 12.03.2026 ***Epub:*** 07.04.2026 ***Publication Date:*** 24.07.2026

INTRODUCTION

Endobronchial ultrasound (EBUS) is a minimally invasive diagnostic modality that enables real-time evaluation of the bronchial mucosa, mediastinal and hilar lymph nodes (LN), and major vessels, as well as nodal tissue sampling in both malignant and benign diseases with mediastinal LN involvement.[1] EBUS-guided transbronchial needle aspiration (TBNA), which is used when a mediastinal and/or hilar LN biopsy is required for diagnostic purposes, staging, or molecular testing, has high sensitivity.[1,2] In recent years, EBUS has significantly reduced the need for mediastinoscopy for mediastinal LN sampling and has become the preferred initial procedure. Furthermore, tissue samples obtained using EBUS-TBNA are suitable for both pathological tumor subtyping and genotyping.[2]

Corresponding author: Ferhan Karataş, MD, e-mail: fkaratas19@hotmail.com

Rapid on-site evaluation (ROSE), an assessment technique used in interventional pulmonology, involves rapid specimen preparation, staining, and cytological examination to ensure safe access to the target lesion during the procedure.[3] Providing immediate feedback on sample quality facilitates decision-making regarding the need for additional sampling or termination of the procedure.[4] The impact of evaluating EBUS-TBNA specimens under ROSE guidance on the diagnostic yield remains controversial. While some studies[5,6] have reported that unnecessary biopsies are prevented, procedure-related complications are reduced, the number of sampled LNs decreases, the likelihood of encountering suspicious material declines, and diagnostic performance improves, other studies[7] have indicated that similar benefits could not be achieved.

Accurate and rapid diagnosis of primary tumors or metastases in lung cancer and other malignancies is critical for treatment planning and prognosis.[8] Therefore, concordance between ROSE and the final pathological diagnosis is of particular significance.

In this study, we aimed to evaluate ROSE, the final pathological diagnostic adequacy, diagnostic concordance, and factors influencing these outcomes in patients undergoing LN aspiration with EBUS.

MATERIAL AND METHODS

Study Design and Patient Population

This study was designed retrospectively in accordance with the Declaration of Helsinki and approved by the Koç University Ethics Committee (decision/protocol number: 2024.222.IRB2.103, date: 07.06.2024). All patients aged >18 years who underwent EBUS in the department of chest diseases between February 2023 and January 2024 for either benign or malignant disease were included in the study. Data on procedures and cases were obtained through retrospective review of the hospital's electronic medical records. There were no missing data for the primary outcome or key procedural variables used in the diagnostic performance and regression analyses. However, some morphological variables (e.g., LN margin, echogenicity, and shape) had missing values owing to incomplete procedural documentation. These variables were excluded from the regression analyses, and available-case analyses were used for descriptive statistics. No imputation method was used.

Main Points

- Rapid on-site evaluation (ROSE) during endobronchial ultrasound (EBUS)-guided transbronchial needle aspiration (TBNA) provided high diagnostic performance (accuracy 90.1%; sensitivity 76.6%). ROSE yielded robust adequacy rates (89.7%) and substantial concordance with final pathology (88.2%).
- Number of needle passes and lymph node (LN) diameter improved ROSE and final diagnostic adequacy.
- Diagnostic discordance occurred in 11.8% of nodes, related to subtyping limitations and observer variability.
- Larger LNs and greater pathologist experience were associated with lower ROSE-final diagnosis concordance.
- ROSE enhances EBUS-TBNA efficiency and yield while acknowledging cytologic subtyping constraints. Results can guide pass number targets and node selection strategies in routine practice.

Endobronchial Ultrasound Procedure

In our clinic, patients undergoing EBUS-TBNA are routinely informed of the procedure, including its indications and potential complications. Before the procedure, the patients are evaluated for bleeding diathesis, hypoxemia, cardiac, renal, or hepatic failure, and neurological disorders. Consultation and approval from relevant specialists are obtained when deemed necessary. Written informed consent was obtained from all patients prior to the procedure.

General anesthesia is administered by an anesthesiologist for all procedures. A 7.5 MHz convex probe bronchoscope (BF-UC180F, Olympus Optical Co., Tokyo, Japan) and an EU-C2000 processor (Olympus, Tokyo, Japan) are used for EBUS procedures. Standard bronchoscopic examination was performed first, followed by EBUS. During EBUS, LN stations are visualized according to the Mountain-Dresler LN map,[9] and systematic sampling is performed using an Olympus ViziShot NA-SX 22G EBUS needle. The decision to terminate the procedure is made jointly by the chest physician and a pathologist.

During EBUS-TBNA, morphological features of LNs, such as margins, echogenicity, short-axis diameter, number of needle passes per station, and ROSE-based diagnosis, were recorded.

Rapid on-Site Evaluation and Final Pathological Diagnosis

In our institution, materials obtained by EBUS-TBNA are routinely evaluated for bedside adequacy. After macroscopic assessment, selected portions of the aspirated material are smeared onto slides and stained using the Diff-Quik method, while the remaining material is fixed in formaldehyde to prepare a cell block. Diff-Quik-stained smears are examined microscopically in the procedure room by a pathologist. Samples containing >30% lymphocytes or anthracotic macrophages are considered representative and classified as "adequate." When malignant cells are identified, tumor subtyping is performed whenever possible; otherwise, the sample is recorded as malignant. In the absence of malignant cells, samples are categorized as benign if adequate, or as hypocellular if not fully representative. In the present study, the final pathological diagnoses were established by re-evaluating smear preparations and cell blocks in the pathology laboratory.

Statistical Analysis

All statistical analyses were conducted using the Statistical Package for the Social Sciences for Windows (SPSS), version 28.0 (IBM Corp., Armonk, NY, USA). Categorical variables were summarized as frequencies and percentages, while continuous variables were reported as mean ± standard deviation.

Diagnostic performance parameters, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall diagnostic accuracy, were calculated

with corresponding 95% confidence intervals (CI) using the exact binomial (Clopper–Pearson) method.

Associations between categorical variables were evaluated using the chi-square test or Fisher's exact test, as appropriate. The distribution of continuous variables was assessed using the Shapiro–Wilk test and confirmed by visual inspection of histograms and quantile–quantile plots. Normally distributed variables were compared using parametric tests, including the independent samples t-test or one-way analysis of variance (ANOVA), as appropriate. Non-normally distributed variables were analyzed using the Mann–Whitney U test or Kruskal–Wallis test. When ANOVA demonstrated statistically significant differences, Scheffé post-hoc analysis was performed to identify pairwise group differences.

Multivariable logistic regression analysis was performed to identify independent predictors of ROSE adequacy, final diagnostic adequacy, and ROSE-final diagnostic concordance. Model calibration was assessed using the Hosmer–Lemeshow goodness-of-fit test, and model discrimination was quantified using the area under the receiver operating characteristic curve (AUC) and Nagelkerke R^2. Multicollinearity was assessed using variance inflation factor, and no significant multicollinearity was detected.

Cases with missing data for outcome variables were excluded from the respective analyses, and complete-case analysis was performed without imputation. All statistical tests were two-sided, and a P value < 0.05 was considered statistically significant.

RESULTS

A total of 95 cases, with a mean age of 65±13.8 years, including 59 males (62.2%), were included in the study. EBUS-TBNA was performed at 203 stations (197 LN and 6 masses). All procedures were performed under general anesthesia, with 77 aspirations (37.9%) performed using an endotracheal tube and 126 (62.1%) performed using a laryngeal mask airway. A total of 683 aspirations were conducted, with a mean number of needle passes of 3.4±1.8 per station.

Of the LNs, 125 (61.5%) were in the mediastinum, 143 (88.3%) had a well-defined border, 139 (73.9%) exhibited heterogeneous echogenicity, and 65 (56.5%) were round in shape, with a mean diameter of 16.61±11.12 mm. No missing data were present for the primary outcome variables. Some morphological variables had missing values owing to incomplete documentation, and analyses were performed using the available case data (Table 1).

The most frequently identified diagnosis was "lymphoid tissue," which was observed in 107 cases of ROSE (52.7%) and in 95 cases on final pathological examination (46.8%). Detailed ROSE and final diagnostic results are shown in Figure 1.

Based on LN-level analysis, ROSE demonstrated a sensitivity of 76.6% (95% CI: 65.6–85.5), specificity of 100.0% (95% CI: 96.6–100.0), PPV of 100.0% (95% CI: 93.9–100.0), NPV of 85.4% (95% CI: 79.6–89.7), and an overall accuracy of 90.1% (95% CI: 84.8–94.0).

Table 1. Localization and morphological features of sampled targets

Localization n (%)		
Mediastinal		**125 (61.5)**
	7	56 (27.6)
	4R	48 (23.6)
	4L	17 (8.4)
	2R	4 (2)
Hilar		**78 (38.5)**
	11L	26 (12.8)
	11R	26 (12.8)
	10R	10 (4.9)
	10L	9 (4.4)
	Mass	6 (3)
	12L	1 (0.5)
Morphological features n (%)		
Edge features		
	Smooth	143 (88.3)
	Uncertain	17 (10.5)
	Lobulated	2 (1.2)
Echogenicity features		
	Heterogeneous	139 (73.9)
	Hypoechoic	38 (20.2)
	Isoechoic	11 (5.9)
Shape features		
	Round	65 (56.5)
	Conglomerate	40 (34.8)
	Oval	10 (8.7)
LN diameter, mm (mean ± SD)		16.6±11.1

LN: lymph node, SD: standard deviation

Diagnostic discordance was observed in 24 LNs (11.8%). The ROSE and final pathological results of discordant cases are presented in Table 2. Additionally, in 2 of the 9 LNs diagnosed as non-small cell lung cancer (NSCLC) by ROSE, and in 8 of the 18 LNs diagnosed as "carcinoma" (5 "NSCLC", 3 "carcinoma"), specific tumor subtyping could not be determined on final pathological analysis.

The mean number of needle passes in EBUS-TBNA procedures with ROSE adequacy (3.5±1.7) was significantly higher than that in those without adequacy (2.6±1.8) (P = 0.008). Similarly, the mean diameter of LN in ROSE-adequate EBUS-TBNA cases (15.7±9.3 mm) was larger than in those inadequate for ROSE (P = 0.028). In EBUS-TBNA procedures with final diagnostic adequacy, the mean number of needle passes was also higher (3.5±1.7; P = 0.002). Moreover, the mean LN diameter (15.8±9.5 mm) was significantly larger than that in cases without final diagnostic adequacy (10.9±6.8 mm) (P = 0.022). In EBUS-TBNA cases with ROSE-final diagnosis concordance, the mean LN diameter (14.9±9.1 mm) was lower than that in discordant cases (18.6±10.5 mm), although this difference was not statistically significant (P = 0.056). In contrast, the

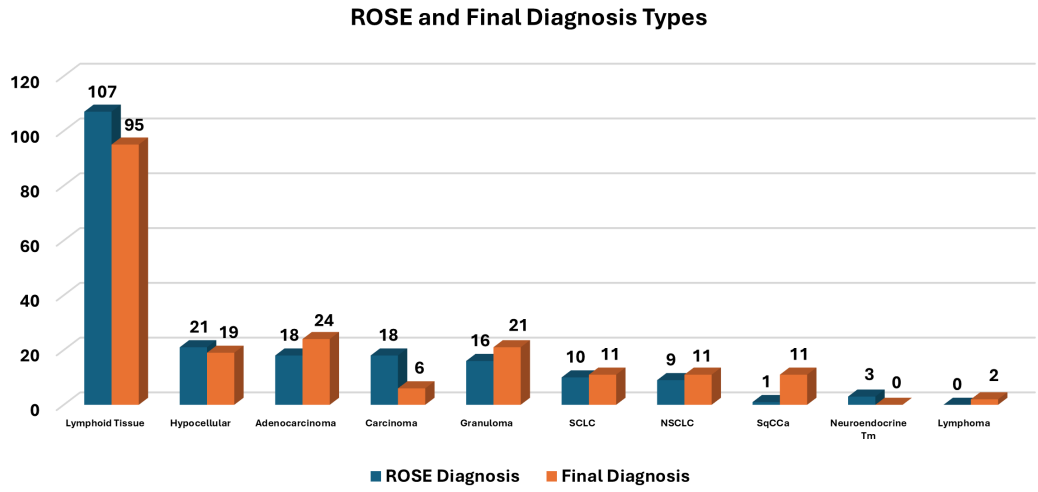

Figure 1. Comparison of ROSE and final pathological diagnosis categories in lymph nodes sampled by EBUS-TBNA. This figure illustrates the distribution of diagnostic categories obtained by ROSE and the corresponding final pathological diagnosis. Lymphoid tissue was the most frequently identified category in both ROSE and the final diagnosis. Differences between ROSE and the final diagnosis were observed in specific pathological subtypes, including carcinoma, adenocarcinoma, granuloma, squamous cell carcinoma, and lymphoma, reflecting cases of diagnostic discordance and refinement after definitive pathological evaluation.

ROSE: rapid on-site evaluation, EBUS: endobronchial ultrasound, TBNA: transbronchial needle aspiration, SqCCa: squamous cell carcinoma, SCLC: small cell lung cancer, NSCLC: non-small cell lung cancer

Table 2. ROSE and final pathology results in cases with diagnostic discordance

		ROSE diagnoses (n)			
		Lymphoid tissue	**Hypocellular**	**Granuloma**	**Total**
Final diagnoses (n)	Lymphoid tissue	-	4	1	5
	Hypocellular	4	-	-	4
	Granuloma	6	-	-	6
	RCC	3	-	-	3
	SqCCa	1	1	-	2
	SCLC	1	-	-	1
	NSCLC	1	-	-	1
	Lymphoma	1	1	-	2
	Total	17	6	1	24

ROSE: rapid on-site evaluation, RCC: renal cell carcinoma, SqCCa: squamous cell carcinoma, SCLC: small cell lung cancer, NSCLC: non-small cell lung cancer

professional experience of the pathologists evaluating the final diagnosis was significantly lower in concordant cases (18.2±11.3 years) compared with discordant cases (23.4±11.4 years) (P = 0.044) (Table 3).

Multivariate logistic regression analysis demonstrated that the number of needle passes was independently associated with both ROSE diagnostic adequacy [odds ratio (OR): 1.433, P = 0.038] and final diagnostic adequacy (OR: 1.593, P = 0.021). LN diameter was independently associated with final diagnostic adequacy (OR: 1.078, P = 0.045) and with diagnostic concordance (OR: 0.951, P = 0.026). Final pathologist experience was independently associated with diagnostic concordance (OR: 0.951, P = 0.015), but did not significantly affect diagnostic adequacy (Table 4).

The final diagnostic adequacy model demonstrated excellent calibration (Hosmer–Lemeshow P = 0.902), good discrimination (AUC: 0.758), and acceptable explanatory power (Nagelkerke R^2: 0.164). The ROSE adequacy model showed acceptable calibration (P = 0.073) and moderate discrimination, while the concordance model demonstrated near-acceptable discrimination with modest explanatory power (Nagelkerke R^2: 0.108); both models had AUC: 0.697. No evidence of multicollinearity was detected among the independent variables (Table 5).

DISCUSSION

In this study of LNs sampled using EBUS-TBNA, the diagnostic concordance rate between ROSE and final pathology was 88.2%. The sample adequacy rates were 89.7% for ROSE and 90.6% for the final pathology. Multivariate logistic regression analysis showed that the number of needle passes was independently associated with both ROSE and final diagnostic adequacy; LN diameter was associated with diagnostic adequacy; and greater final pathologist experience was independently associated with lower ROSE-final diagnostic concordance.

Table 3. Factors associated with ROSE adequacy, final diagnostic adequacy, and ROSE-final diagnostic concordance

LN features		ROSE adequacy			Final diagnostic adequacy			Diagnostic concordance		
		No	Yes	*P*	No	Yes	*P*	No	Yes	*P*
Gender n (%)	Male	15 (11.7)	113 (88.3)	0.401	12 (9.4)	116 (90.6)	0.992	18 (14.1)	110 (85.9)	0.197
	Female	6 (8.0)	69 (92.0)		7 (9.3)	68 (90.7)		6 (8.0)	69 (92.0)	
Type of intubation n (%)	ETT	8 (10.4)	69 (89.6)	0.986	8 (10.4)	69 (89.6)	0.693	7 (9.1)	70 (90.9)	0.346
	LMA	13 (10.3)	113 (89.7)		11 (8.7)	115 (91.3)		17 (13.5)	109 (86.5)	
Location n (%)	Hilar	11 (14.7)	64 (85.3)	0.121	9 (12.0)	66 (88.0)	0.322	12 (16.0)	63 (84.0)	0.158
	Mediastinal	10 (7.8)	118 (92.2)		10 (7.8)	118 (92.2)		12 (9.4)	116 (90.6)	
Age (mean ± SD)		65.2±14.3	66.5±12.9	0.843	65.9±15.5	66.4±12.8	0.736	65.5±12.4	66.4±13.1	0.823
LN diameter (mean ± SD)		11.9±9.1	15.7±9.3	**0.028**	10.9±6.8	15.8±9.5	**0.022**	18.6±10.5	14.9±9.1	0.056
Number of passes (mean ± SD)		2.6±1.8	3.5±1.7	**0.008**	2.5±1.7	3.5±1.7	**0.002**	3.6±2.0	3.3±1.7	0.438
ROSE pathologist experience (years) (mean ± SD)		13.5±7.9	12.7±7.1	0.327	13.7±8.3	12.6±7.1	0.43	11.3±4.8	12.9±7.5	0.814
Final diagnostic pathologist experience (years) (mean ± SD)		24.3±11.1	18.1±11.2	**0.023**	23.5±11.4	18.3±11.3	0.067	23.4±11.4	18.2±11.3	**0.044**

ROSE: rapid on-site evaluation, SD: standard deviation, LN: lymph node, ETT: endotracheal tube, LMA: laryngeal mask airway

Table 4. Multivariate binary logistic regression analysis identifying independent predictors of ROSE diagnostic adequacy, final diagnostic adequacy, and diagnostic concordance

	ROSE diagnostic adequacy		Final diagnostic adequacy		Diagnostic concordance	
Independent variable	**OR (95% CI)**	***P* value**	**OR (95% CI)**	***P* value**	**OR (95% CI)**	***P* value**
LN diameter	1.058 (0.995–1.125)	0.073	1.078 (1.002–1.161)	**0.045**	0.951 (0.909–0.994)	**0.026**
Number of passes	1.433 (1.020–2.015)	**0.038**	1.593 (1.074–2.361)	**0.021**	0.911 (0.717–1.157)	0.444
ROSE pathologist experience	0.982 (0.924–1.043)	0.550	-	-	1.058 (0.975–1.147)	0.175
Final diagnosis pathologist experience	-	-	0.960 (0.918–1.003)	0.069	0.951 (0.914–0.990)	**0.015**

Multivariate binary logistic regression analyses were performed using the enter method. Statistically significant results are shown in bold
ROSE: rapid on-site evaluation, OR: odds ratio, CI: confidence interval, LN: lymph node,

Table 5. Calibration, discrimination, and explanatory power of multivariate logistic regression models

Model	Hosmer–Lemeshow *P* value	Nagelkerke R^2	AUC
ROSE diagnostic adequacy	0.073	0.091	0.697
Final diagnostic adequacy	0.902	0.164	0.758
Diagnostic concordance	0.021	0.108	0.697

Hosmer–Lemeshow test was used to assess model calibration. Nagelkerke R^2 was used to evaluate explanatory power. Model discrimination was assessed using the area under the receiver operating characteristic curve (AUC). AUC values ≥0.7 indicate acceptable discrimination
ROSE: rapid on-site evaluation

Previous studies have demonstrated that ROSE significantly enhances the diagnostic performance of EBUS-TBNA, reporting sensitivity of 88–97%, specificity of 93–99%, PPV of 94–98%, and NPV of 87–98%.[5,6,10] These findings indicate that ROSE strengthens sample adequacy and concordance with final pathology. In a recent study, Magdy et al.[11] emphasized the superiority of this technique by reporting that the diagnostic accuracy reached 100% in the ROSE-applied group. In our study, ROSE demonstrated a sensitivity of 76.6%, a specificity of 100.0%, a PPV of 100.0%, an NPV of 85.4%, and an overall diagnostic accuracy of 90.1%. Consistent with the high sensitivity and diagnostic accuracy reported in the literature, these results confirm the value of ROSE in providing a rapid and reliable diagnosis. These results support the potential of ROSE to significantly increase the effectiveness of EBUS-TBNA and reduce the need for invasive surgical intervention.

In another study conducted during the severe acute respiratory syndrome coronavirus 2 pandemic, EBUS-TBNA was performed without ROSE, using only the cell block technique to reduce the risk of coronavirus disease-2019 transmission.[12] A total of 1,097 and 806 procedures were included in the ROSE and non-ROSE groups, respectively. In this study, sample adequacy from stations 4R and 7 was higher, likely because of their more accessible anatomical locations. The number of biopsy sites per patient was 2.4±2.3 in the ROSE group and 2.9±2.7 in the non-ROSE group, with a statistically significant difference between the groups ($P < 0.001$). While no difference in overall sample adequacy was observed between the two groups ($P = 0.785$), the malignancy detection rate was significantly higher in the ROSE group ($P = 0.036$). These findings may have been influenced by the triage strategy during the pandemic, in which cases with

lower clinical suspicion were followed up, whereas cases with higher clinical suspicion were referred for EBUS-TBNA. In this retrospective analysis, sample adequacy was similar between the ROSE and non-ROSE groups; nevertheless, ROSE increased the malignancy detection rate, and EBUS-TBNA should not be delayed in highly suspicious cases.

The diagnostic yield of EBUS-TBNA is influenced by factors such as lesion size, number of samples, needle gauge, and number of passes; reported accuracy rates range from 71% to 99%.[13] Current guidelines recommend at least three passes for diagnosis and four to six passes for molecular analysis.[1] Kim et al.[14] reported that increasing the number of passes improved the success rates from 52.6% to 100%, with a significant improvement observed after the third pass. In contrast, Hu et al.[15] indicated that the use of general anesthesia and ROSE reduced the number of punctures but did not improve diagnostic efficacy. In our study, the number of needle passes and the LN diameter were significantly higher in cases with ROSE adequacy; similar findings were observed in cases with final diagnostic adequacy. These results align with the literature that supports optimal pass numbers and LN diameters as determinants of diagnostic performance. Fielding et al.[16] also demonstrated that three agitations provided a cell and DNA yield equivalent to ten agitations, reduced blood contamination, and shortened the procedure.

Although ROSE provides high diagnostic accuracy for malignancy, limitations in subtype characterization have been reported.[17] Cytological smears are typically adequate; however, histological biopsies may be required for molecular analysis.[13] In our study, tumor subtyping could not be determined on final pathology in two of nine nodes diagnosed as "NSCLC" by ROSE and in 8 of 18 nodes diagnosed as "carcinoma" by ROSE, which illustrates this limitation. The literature attributes this issue to the lack of preserved tissue architecture and the presence of reactive cells that mimic malignant features.[4,11] Despite its lower NPV (85.4%), the relatively high specificity (100.0%) may be explained by cytological limitations and sample heterogeneity. Technical parameters are also important, as Fielding et al.[16] highlighted the advantages of reduced agitation, and Sakaguchi et al.[18] reported lower blood contamination with 25G needles. In the study by Caupena et al.,[10] the ROSE-final pathology concordance was 96.1% with a false-negative rate of 1.2% and a false-positive rate of 0.2%. In our cohort, the discordance rate was 11.8% with most discrepancies observed in NSCLC and carcinoma cases owing to difficulties in subtyping.

We found that LN diameter was negatively associated with concordance between ROSE and final pathology, with larger LNs being more frequently observed in discordant cases. This finding may be due to necrosis or cystic degeneration, which are more common in larger LNs and may impair cytological evaluation. Additionally, greater professional experience among pathologists who evaluated the final diagnosis (23.4 years vs. 18.2 years; $P = 0.044$) was associated with lower concordance. Previous studies have documented that inter-observer variability and more meticulous assessments by highly experienced pathologists in complex cases may increase the risk of discordance.[11,12] These findings emphasize the crucial roles of morphological features and observer-dependent factors in diagnostic concordance and support the need for multidisciplinary evaluation in challenging cases.

In the study by So et al.,[19] one of the causes of EBUS-TBNA failure in mediastinal LN staging was failure to sample other nodes, owing to false-positive ROSE interpretation. Although the final pathology is rarely positive, this risk exists during ROSE interpretation. The similarity between reactive epithelial cells and malignant cells, particularly for cytopathologists, may complicate their differentiation. As similar issues may occur in granulomatous disease, the ROSE results must be carefully interpreted and confirmed during staging. A systematic review reported a pooled sensitivity of 89.6%, a specificity of 95.9%, and a false-positive rate of 4.1% for ROSE performed by pulmonologists.[2] Pulmonologists have achieved high diagnostic accuracy in malignancy detection, with most errors arising from misinterpretation of bronchial epithelial cells. The concordance rate between pulmonologists and pathologists was 90%, suggesting that pulmonologists can serve as reliable alternatives when a pathologist is not available. A recent meta-analysis by Chen et al.[20] further confirmed the high diagnostic accuracy of ROSE for lung cancer during EBUS-TBNA. When used for mediastinal or hilar LN staging, diagnostic performance improved, with subgroup analyses showing a sensitivity of 90.1% and a specificity of 96.9%. ROSE increases sample adequacy, reduces the need for repeated procedures, shortens diagnostic time, ensures sufficient material for molecular testing, and enhances patient safety by enabling high diagnostic accuracy with fewer needle passes. These findings confirm that ROSE is a feasible, cost-effective technique that enhances the diagnostic yield and specificity of EBUS-TBNA and can be safely performed by trained pulmonologists in the absence of pathology support. Furthermore, the integration of artificial intelligence-assisted digital ROSE systems is expected to reduce processing time and improve standardization.

Study Limitations

This study has several limitations. First, its retrospective, single-center design may have introduced selection bias and limited the generalizability of the findings. The relatively small sample size may have reduced the statistical power, particularly in the subgroup analyses. ROSE and the final diagnosis were evaluated by different pathologists, introducing inter-observer variability that could affect diagnostic concordance. Although the final pathology is considered to be the gold standard, false-negative or false-positive results remain possible. Although no a priori power calculation was performed because of the retrospective design, a post hoc power analysis indicated that the study had approximately 82% power to detect moderate effect sizes in the concordance analysis. However, smaller subgroup sizes may limit the ability to detect weaker associations. Cytological limitations, particularly the lack of architectural preservation and sample heterogeneity, may hinder the subtype classification. Technical parameters (needle gauge, number of agitations, and vacuum use) were not standardized, which could have influenced diagnostic performance. Additionally, the lack of multidisciplinary evaluation and the absence of methods (e.g., cryobiopsy or molecular testing) in complex cases are additional limitations.

CONCLUSION

This study demonstrates that ROSE during EBUS-TBNA provides high diagnostic accuracy and sensitivity in evaluating mediastinal and hilar LNs. Our findings show that ROSE enhances sample adequacy, ensures strong concordance with final pathology, and expedites clinical decision-making. Nevertheless, limitations in sensitivity and negative predictive value, along with challenges in tumor subtypingassociated with cytological interpretation and sample heterogeneity underscore the importance of multidisciplinary assessment and supplementary pathological methods. Implementation of ROSE offer significant clinical advantages by reducing the need for invasive surgical interventions.

Ethics

Ethics Committee Approval: This study was designed retrospectively in accordance with the Declaration of Helsinki and approved by the Koç University Ethics Committee (decision/protocol number: 2024.222.IRB2.103, date: 07.06.2024).

Informed Consent: Written informed consent was obtained from all patients prior to the procedure.

Footnotes

Authorship Contributions

Surgical and Medical Practices: F.K., M.Ç.A.M., I.U., Ö.D., Concept: F.K., Design: F.K., Data Collection or Processing: F.K., M.Ç.A.M., I.U.,T.Y., E.K., Analysis or Interpretation: F.K., Literature Search: F.K., Writing: F.K., M.Ç.A.M.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Gilbert CR, Dust C, Argento AC, et al. Acquisition and handling of endobronchial ultrasound transbronchial needle samples: an American College of Chest Physicians Clinical Practice Guideline. *Chest.* 2025;167(3):899-909. [Crossref]
2. Ehab A, Elgenidy A, Abobakr OA, et al. The accuracy and adequacy of rapid on-site evaluation conducted by trained pulmonologist on transbronchial needle aspiration samples: a systemic review and meta-analysis. *Respir Med.* 2025;236:107903. [Crossref]
3. Li C, Xie W, Cao J, Feng J. Detailed procedure and clinical application overview of rapid on-site evaluation in diagnostic interventional pulmonology. *J Res Med Sci.* 2020;25:35. [Crossref]
4. Vasugi GA, Mathivanan KM, Rajendiran S, Sundaram S, Ayub II. Traveling with ROSE in EBUS-TBNA – experience from a Tertiary Care Hospital in South India. *J Cytol.* 2021;38(4):175-179. [Crossref]
5. Al-Najjar H, Almadhi M, Rajai A, et al. Comparing diagnostic performance across two high-volume endobronchial ultrasound services in greater manchester with and without rapid on-site evaluation. *Respiration.* 2025;104(8):586-596. [Crossref]
6. Silav ZK, Çıkrıkçıoğlu M, Aydemir B, Yıldırım M. EBUS ROSE-guided cytological evaluations of intrathoracic lesions. Assessments of the pathological material adequacy and the diagnostic efficacy. *Diagn Cytopathol.* 2023;51(7):423-433. [Crossref]
7. Mercer C, Brackenborough K, Davidson C, et al. Rapid onsite evaluation: a prospective observational study of endobronchial ultrasound-guided transbronchial needle aspirates to expedite diagnosis. *Br J Biomed Sci.* 2024;81:12938. [Crossref]
8. Jaiyesimi IA, Leighl NB, Ismaila N, et al. Therapy for stage IV non-small cell lung cancer with driver alterations: ASCO living guideline, version 2023.3. *J Clin Oncol.* 2024;42(11):1-22. [Crossref]
9. Mountain CF, Dresler CM. Regional lymph node classification for lung cancer staging. *Chest.* 1997;111:1718-1723. [Crossref]
10. Caupena C, Esteban L, Jaen A, et al. Concordance between rapid on-site evaluation and final cytologic diagnosis in patients undergoing endobronchial ultrasound-guided transbronchial needle aspiration for non-small cell lung cancer staging. *Am J Clin Pathol.* 2020;153:190-197. [Crossref]
11. Magdy O, AbdelDayem A, ELMaraghi A, Ali M, Hafez F. The value of rapid on-site evaluation during conventional and endobronchial ultrasound needle aspiration in the diagnosis of mediastinal lymphadenopathy and lung cancer. *The Egyptian Journal of Bronchology.* 2024;18:70. [Crossref]
12. Zhao X, Boothe P, Hussnain Naqvi SM, et al. Assessing ROSE for adequacy of EBUS-TBNA compared with a direct-to-cell block approach as a response to the COVID-19 pandemic. *J Am Soc Cytopathol.* 2022;11(6):368-374. [Crossref]
13. Bandiera A, Arrigoni G. The impact of pathological analysis on endobronchial ultrasound diagnostic accuracy. *Mediastinum.* 2020;4:19. [Crossref]
14. Kim BK, Choi H, Kim CY. Factors associated with increased diagnostic yield of endobronchial ultrasound-guided transbronchial needle aspiration: an observational single center study. *J Thorac Dis.* 2024;16(1):439-449. [Crossref]
15. Hu YH, Xu WH, Li YY. The impact of anesthesia methods and rapid on-site evaluation on the diagnosis of lung cancer using endobronchial ultrasound-guided transbronchial needle aspiration: a retrospective study. *J Thorac Dis.* 2024;16(12):8493-8502. [Crossref]
16. Fielding D, Dalley AJ, Singh M, et al. Prospective optimization of endobronchial ultrasound-guided transbronchial needle aspiration lymph node assessment for lung cancer: three needle agitations are noninferior to 10 agitations for adequate tumor cell and DNA yield. *JTO Clin Res Rep.* 2022;3(10):100403. [Crossref]
17. Murthi M, Donna E, Arias S, et al. Diagnostic accuracy of endobronchial ultrasound-guided transbronchial needle aspiration (EBUS-TBNA) in real life. *Front Med (Lausanne).* 2020;7:118. [Crossref]
18. Sakaguchi T, Inoue T, Miyazawa T, Mineshita M. Comparison of the 22-gauge and 25-gauge needles for endobronchial ultrasound-guided transbronchial needle aspiration. *Respir Investig.* 2021;59(2):235-239. [Crossref]
19. So C, Matsumoto Y, Imabayashi T, et al. Identifying factors causing failure of nodal staging by endobronchial ultrasound-guided transbronchial needle aspiration in non-small cell lung cancer. *Transl Lung Cancer Res.* 2023;12:2169-2180. [Crossref]
20. Chen CC, Lu SC, Chang YK, et al. Diagnostic performance of rapid on-site evaluation during bronchoscopy for lung cancer: a comprehensive meta-analysis. *Cancer Cytopathol.* 2025;133(1):22908. [Crossref]

Thorac Res Pract. 2026;27(4):229-234 **DOI:** 10.4274/ThoracResPract.2026.2025-11-4

Original Article

CT Histogram Analysis for Predicting Malignancy in Anterior Mediastinal Masses

Nilay Çavuşoğlu Yalçın[1], Mustafa Yalçın[2], Okan Karataş[1]

[1]Clinic of Thoracic Surgery, University of Health Sciences Türkiye, Antalya Training and Research Hospital, Antalya, Türkiye
[2]Clinic of Radiology, University of Health Sciences, Antalya Training and Research Hospital, Antalya, Türkiye

Cite this article as: Çavuşoğlu Yalçın N, Yalçın M, Karataş O. CT histogram analysis for predicting malignancy in anterior mediastinal masses. *Thorac Res Pract.* 2026;27(4):229-234

ABSTRACT

OBJECTIVE: Anterior mediastinal masses encompass a broad spectrum from benign thymic hyperplasia to malignant thymomas and thymic complementary computed tomography (CT) histogram and texture analysis have emerged as complementary, non-invasive quantitative biomarkers for malignancy prediction. This study investigates whether CT histogram-derived parameters, particularly mean attenuation and first-order metrics, can differentiate benign from malignant anterior mediastinal masses to support preoperative risk stratification.

MATERIAL AND METHODS: We retrospectively analyzed data from 102 patients who underwent surgical resection of anterior mediastinal masses. Non-contrast CT scans were evaluated, and mean Hounsfield unit (HU), standard deviation, skewness, and kurtosis values were extracted using region-of-interest-based histogram analysis. Histopathological diagnosis is considered the gold standard for classification. Group comparisons were performed using appropriate parametric or non-parametric tests, and the diagnostic performance was evaluated using receiver operating characteristic (ROC) analysis.

RESULTS: Of the 102 patients, 60 (58.8%) had malignant and 42 (41.2%) had benign pathologies. The mean CT attenuation was higher in malignant lesions than in benign lesions (63.3±16.9 HU vs. 48.3±39.0 HU; $P = 0.023$), whereas the standard deviation was similar between the groups ($P = 0.810$). Skewness and kurtosis were higher in malignant lesions ($P < 0.05$), indicating denser and more peaked attenuation distributions. ROC analysis of mean HU demonstrated moderate discriminative performance (area under the curve: 0.64), with high sensitivity (96.7%) and limited specificity (50.0%).

CONCLUSION: CT histogram analysis, particularly mean attenuation, may serve as a complementary biomarker for differentiating benign from malignant anterior mediastinal masses. Given its high sensitivity but limited specificity, it should be used as an adjunct to clinical and morphologic assessment rather than as a determinant of surgical management. Prospective validation incorporating volumetric and advanced radiomic parameters is required.

KEYWORDS: Anterior mediastinal mass, computed tomography, histogram analysis, thymoma, malignancy prediction

Received: 22.12.2025 ***Revision Requested:*** 12.01.2026 ***Last Revision Received:*** 18.01.2026

Accepted: 26.01.2026 ***Epub:*** 11.03.2026 ***Publication Date:*** 24.07.2026

INTRODUCTION

Anterior mediastinal masses comprise a heterogeneous group of lesions with varying clinical behaviors, ranging from benign entities, such as thymic hyperplasia and cysts, to malignant neoplasms, including thymomas and thymic carcinomas. These lesions are increasingly identified incidentally due to the widespread use of computed tomography (CT) in thoracic imaging, particularly in asymptomatic individuals undergoing lung cancer screening or trauma evaluation.[1]

Surgical resection remains the cornerstone of thymic epithelial tumor treatment. However, a significant proportion of autoimmune myasthenia gravis are ultimately found to be non-therapeutic, with up to 27%–30% of thymectomy revealing benign histology.[2] This has raised concerns regarding overtreatment and highlighted the need for improved risk stratification tools that can assist in distinguishing benign from malignant lesions using non-invasive imaging modalities.

Corresponding author: Nilay Çavuşoğlu Yalçın, MD, e-mail: drnili@hotmail.com

Conventional CT imaging is routinely used to evaluate anterior mediastinal masses, focusing on size, shape, location, and presence of features such as fat, calcification, and invasion. However, these qualitative assessments are often insufficient to reliably distinguish benign from malignant lesions, particularly in borderline or equivocal cases.[3] As such, there is growing interest in the application of advanced quantitative imaging techniques, including histogram and texture analysis, to enhance diagnostic accuracy.

Histogram analysis, a component of radiomics, involves extracting quantitative metrics from voxel intensity distributions within a region-of-interest (ROI). This technique enables the objective evaluation of lesion density, uniformity, and internal heterogeneity, which may correlate with histological behavior.[4] Histogram-based parameters, such as mean attenuation, standard deviation, and percentiles, can effectively differentiate thymic hyperplasia from thymoma and carcinoma, with high diagnostic performance.[5-7]

Despite its promise, the integration of CT histogram analysis into routine clinical practice remains limited, and further validation is required. In this context, the present study aimed to evaluate the ability of simple histogram parameters—mean and standard deviation of attenuation on non-contrast CT—to differentiate benign from malignant anterior mediastinal masses in patients undergoing surgical resection. We also included skewness and kurtosis values in the analysis to provide a more complete characterization of the voxel intensity distribution, as recommended in the recent literature. We hypothesize that malignant lesions will exhibit higher and more homogeneous attenuation patterns, thereby supporting the utility of CT histogram analysis as a non-invasive preoperative risk stratification tool.

MATERIAL AND METHODS

A retrospective analysis was conducted on 102 consecutive patients who underwent surgical resection for anterior mediastinal masses between January 2015 and December 2024. This retrospective study was approved by the Ethics Committee of Antalya Training and Research Hospital (approval number: 2025-125, approval date: 08 May 2025). The requirement for informed consent was waived due to the retrospective nature of the study. Only patients with available preoperative non-contrast thoracic CT imaging, complete clinical information, and definitive histopathological diagnosis were included in the study. Patients were excluded if they had received prior chemotherapy or radiotherapy, had incomplete imaging data, or if the lesion was evaluated only by biopsy without subsequent surgical resection. A STROBE-compliant flow diagram summarizing patient selection and exclusion criteria is provided in Figure 1.

Main Points

- Malignant anterior mediastinal lesions show significantly higher mean Hounsfield unit (HU) values than benign ones on non-contrast computed tomography (CT).
- A histogram mean cutoff of 39.3 HU yields high sensitivity (96.7%) for predicting malignancy.
- Simple first-order CT histogram parameters can aid preoperative risk stratification and may reduce unnecessary surgical resections.

Non-contrast CT was examined using multi-detector CT scanners with standard clinical protocols (120 kVp tube voltage, 1–2 mm slice thickness). Axial images were reconstructed using soft-tissue kernels and exported in the DICOM format for analysis. Histogram-derived attenuation values were extracted using a standardized ROI approach. A fixed-diameter circular ROI was manually placed on the axial slice demonstrating the largest cross-sectional area of each lesion while carefully avoiding calcified, cystic, necrotic, fatty, or vascular components to minimize measurement variability. A thoracic radiologist with 10 years of experience placed all ROIs.

Demographic variables (age and sex), lesion size, and histogram parameters were compared between the benign and malignant groups based on the final surgical pathology. The normality of continuous variables was assessed using the Shapiro–Wilk test. Independent-samples t-tests were used for normally distributed variables; for non-normally distributed variables, the Mann–Whitney U test was applied. Categorical variables were analyzed using the χ^2 test or Fisher's exact test, where appropriate. Receiver operating characteristic (ROC) curve analysis was performed to evaluate the diagnostic performance of the histogram mean Hounsfield unit (HU). The area under the curve (AUC), sensitivity, specificity, and optimal threshold values were calculated using Youden's index. Skewness and kurtosis values were compared between the malignant and benign groups to evaluate differences in distribution shape characteristics.

Statistical Analysis

All statistical analyses were performed using the Statistical Package for the Social Sciences version 26.0 (IBM Corp., Armonk, NY). A *P* value < 0.05 was considered statistically significant. Because this was a retrospective study, an a priori

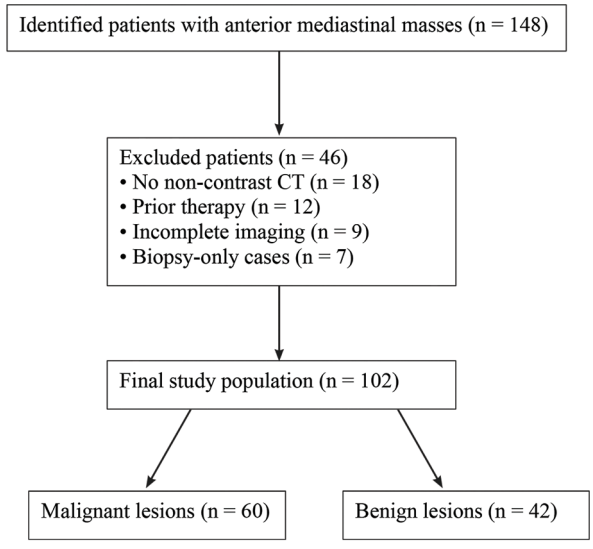

Figure 1. STROBE-compliant patient selection flowchart

power analysis could not be performed; however, a post-hoc effect size analysis for the mean HU difference was conducted and demonstrated sufficient discriminatory power for the primary outcome measure.

RESULTS

This retrospective study included 102 patients with anterior mediastinal masses. Of these, 60 patients (58.8%) were diagnosed with malignant lesions and 42 patients (41.2%) with benign lesions based on the final histopathological evaluation.

The overall mean age was 51.4 years, with male patients exhibiting a higher mean age (53.6 years) than females (48.9 years) (Figure 2). In the benign group, 50% were male and 50% were female, whereas the malignant group consisted of 55.0% males and 45.0% females (Figure 3).

Among the malignant lesions (n = 60), the World Health Organization (WHO) thymoma subtypes were as follows: type A (n = 2), type B1 (n = 1), type B2 (n = 14), and type B3 (n = 5). Thymic carcinoma occurred in 8 cases, while lymphomas (n = 18) and other malignant tumors (n = 12) comprised the remaining cases (Table 1).

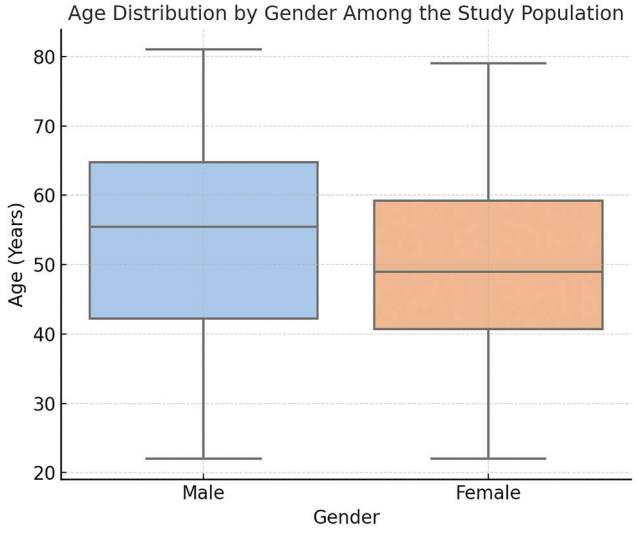

Figure 2. Age distribution by sex among the study population

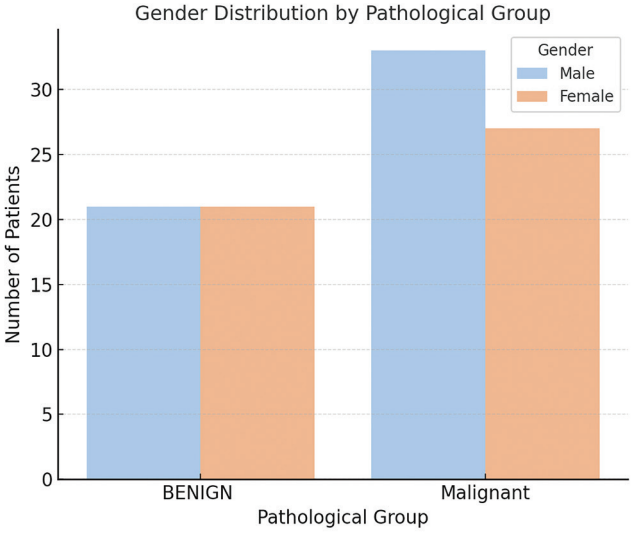

Figure 3. Gender distribution across the benign and malignant pathology groups

CT histogram analysis demonstrated that malignant anterior mediastinal lesions exhibited significantly higher mean HU values than benign lesions (63.3 ± 16.9 HU vs. 48.3 ± 39.0 HU, $P = 0.023$), supporting the diagnostic utility of mean attenuation as a radiologic biomarker of malignancy (Figure 4). In contrast, the standard deviation of HU values was higher in benign lesions (39.0 HU vs. 16.9 HU), reflecting greater internal heterogeneity, although this difference was not statistically significant ($P = 0.810$).

The shape descriptors of the voxel intensity distribution revealed further distinctions between groups. Malignant lesions demonstrated greater right-sided skewness (0.81 vs. 0.53) and higher kurtosis (0.47 vs. 0.76), both of which reached statistical significance ($P = 0.045$ and $P = 0.018$, respectively), suggesting more homogeneously dense tissue profiles in malignancy, whereas benign lesions exhibited broader and flatter attenuation distributions (Table 2).

Table 2 summarizes the histogram-derived parameters of the anterior mediastinal lesions. Malignant masses demonstrated significantly higher mean HU and greater peakedness (kurtosis), whereas benign lesions exhibited greater attenuation variability and flatter distribution profiles.

These findings highlight that while histogram-derived mean attenuation offers high sensitivity but limited specificity, its diagnostic performance remains moderate and should therefore be interpreted as a complementary rather than standalone biomarker.

ROC curve analysis demonstrated moderate discriminative performance of the histogram mean HU (AUC: 0.64), with high sensitivity (96.7%) and limited specificity (50%) (Figure 5). In contrast, the standard deviation of HU values was higher in benign lesions (39.0 HU vs. 16.9 HU), reflecting greater internal heterogeneity, although this difference was not statistically significant ($P = 0.810$).

Malignant lesions demonstrated greater right-sided skewness (0.81 vs. 0.53) and higher kurtosis (0.47 vs. −0.76), both of which reached statistical significance ($P = 0.045$ and $P = 0.018$, respectively), suggesting more homogeneously dense

Table 1. WHO classification of malignant thymic lesions in the study cohort (n = 60)

WHO subtype/malignant category	n	%
Type A	2	3.3%
Type B1	1	1.7%
Type B2	14	23.3%
Type B3	5	8.3%
Thymic carcinoma	8	13.3%
Lymphoma	18	30.0%
Other malignant tumors	12	20.0%

WHO: World Health Organization

Table 2. Descriptive summary of statistics

Group	Mean (HU)	Standard deviation (HU)	Skewness	Kurtosis
Benign	48.26	39.03	0.53	-0.76
Malign	63.25	16.91	0.81	0.47

HU: Hounsfield unit

Table 3. ROC curve performance metrics

Metric	Value
AUC	0.64
Optimal threshold (HU)	39.3
Sensitivity	96.7%
Specificity	50%

ROC: receiver operating characteristic, AUC: area under the curve, HU: Hounsfield unit

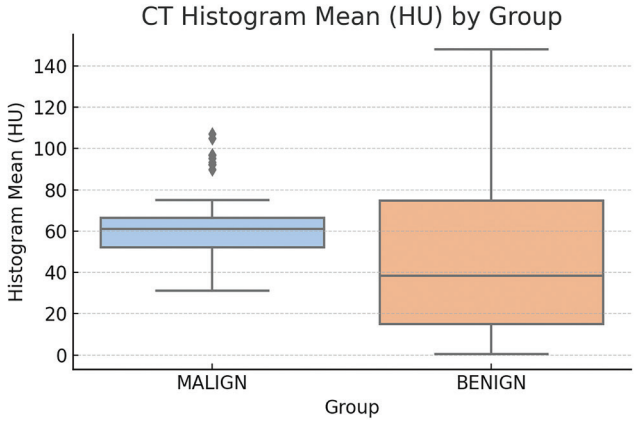

Figure 4. CT histogram mean (HU) by groups

CT: computed tomography, HU: Hounsfield unit

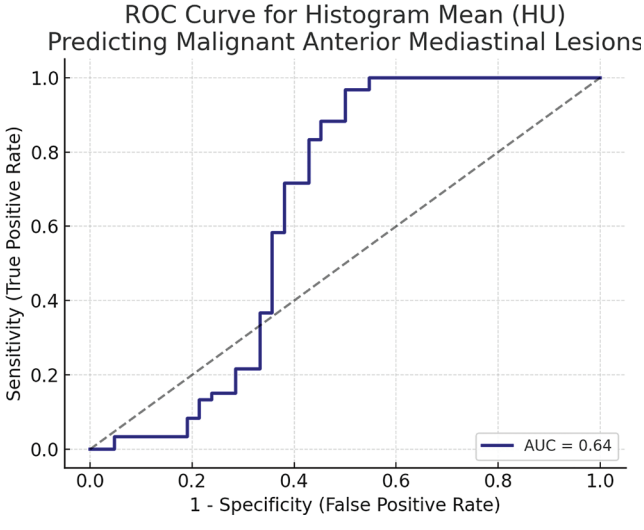

Figure 5. ROC curve for the histogram mean (HU) in predicting malignant anterior mediastinal lesions

ROC: receiver operating characteristic, HU: Hounsfield unit

tissue profiles in malignancy, whereas benign lesions exhibited broader and flatter attenuation distributions (Table 3).

These findings highlight that while histogram-derived mean attenuation offers high sensitivity but limited specificity, its diagnostic performance remains moderate. Therefore, it should be interpreted as a complementary rather than standalone biomarker.

Although benign lesions exhibited a notably higher standard deviation (39.03 HU) than malignant lesions (16.91 HU), this variability did not reach statistical significance, indicating that heterogeneity alone may not reliably differentiate pathology.

Skewness was greater in the malignant group (0.81) than in the non-malignant group, reflecting a stronger right-tailed distribution. This suggests that higher HU values occur more frequently in malignancy, reinforcing the pattern of denser tissue characteristics. Kurtosis was also higher in malignant lesions (0.47), consistent with a more compact and homogeneous tissue density pattern.

These descriptive parameters indicate that malignant anterior mediastinal masses tend to exhibit higher and more homogeneous attenuation, whereas benign masses are more heterogeneous, with wider variation in internal density. Therefore, these metrics may serve as informative adjuncts in preoperative radiologic assessment and risk stratification. Although benign lesions exhibited a notably higher standard deviation (39.03 HU) than malignant lesions (16.91 HU), this variability did not reach statistical significance, indicating that heterogeneity alone may not reliably differentiate pathology.

DISCUSSION

Our findings demonstrate that malignant anterior mediastinal lesions are characterized by significantly higher mean CT attenuation values compared to benign counterparts, supporting the hypothesis that such lesions exhibit greater internal density and structural homogeneity—a notion previously corroborated by Liu et al.[6], who identified histogram-derived median attenuation as an independent discriminator between thymic tumors and non-neoplastic lesions. This alignment with existing literature reinforces the biological plausibility of our results and supports the potential utility of simple, first-order histogram features as practical imaging biomarkers.

Although prior research has explored various radiomics-based texture and entropy metrics,[7,8] our data emphasize the clinical applicability of quantitative values, such as mean attenuation, derived from routine non-contrast CT without specialized software. Importantly, the present study confirms that the mean HU is significantly higher in malignant lesions (P = 0.023),

whereas the standard deviation ($P = 0.810$) fails to demonstrate discriminatory power, indicating that global attenuation values may serve as a more stable surrogate marker of malignancy than heterogeneity measures alone. This concept is consistent with that of Roden and Szolkowska,[8] who reported a substantial overlap in qualitative CT characteristics between benign and malignant thymic lesions.

Our cohort demographics, including a mean patient age of 51.4 years and a modest male predominance among malignant cases (55%), closely mirror previous studies evaluating anterior mediastinal lesions.[9-11] This demographic concordance strengthens external validity and suggests that our findings can be generalized to similar clinical populations.

Furthermore, among malignant cases, the distribution of WHO thymoma subtypes—dominated by B2 and B3 patterns—parallels epidemiologic data from prior thymic tumor registries.[10] This detailed classification (Table 3) addresses reviewer concerns regarding the malignant subgroup composition and clarifies the proportional burden represented by lymphoma (30%) and thymic carcinoma (13.3%), both of which may influence attenuation profiles.

Despite the significant mean HU difference, the ROC analysis demonstrated only moderate discriminative performance, with an AUC of 0.64. Although the sensitivity was very high (96.7%), the specificity remained limited (50%), indicating that the histogram mean HU is more suitable as a rule-out test than a confirmatory diagnostic tool. This finding directly supports the reviewer's comment regarding the need for cautious interpretation of statements suggesting the reduction of unnecessary Thymectomy. Accordingly, we have revised the manuscript to clarify that histogram parameters should be considered adjunctive decision-support tools rather than determinants of surgical candidacy.

Prior work by Rajamohan et al.[12] similarly suggested that single-slice histogram parameters can complement, but not replace, comprehensive morphologic evaluation.[13] Our data are consistent with these conclusions and reinforce the concept that when integrated into structured reporting systems, first-order metrics can enhance radiologic assessment.

The significant elevation of skewness and kurtosis in malignant lesions further supports the notion that malignancies demonstrate more peaked and right-skewed attenuation distributions. These statistically significant findings ($P = 0.045$ and $P = 0.018$, respectively) were added and highlighted per reviewer request, ensuring that all shape descriptors used in the analysis were reported with corresponding p-values. Although benign lesions exhibited greater heterogeneity (higher standard deviation), the difference was not statistically significant, implying that heterogeneity alone may not be a reliable hallmark of benignity.

Several methodological considerations warrant discussion. First, histogram analysis was performed on a single axial slice with a fixed-diameter ROI, chosen to standardize measurement and reflect real-world clinical workflow. However, this approach may not fully capture whole-tumor heterogeneity, particularly in larger or cystic-necrotic lesions. Therefore, we have expanded the limitation section to explicitly acknowledge that volumetric or multi-ROI radiomics techniques could yield more comprehensive density assessments.

Second, areas of calcification, cystic change, and necrosis were intentionally excluded from the ROI placement, a methodological decision intended to minimize measurement noise. Nonetheless, the reviewer's concerns are valid: these elements may represent biologically relevant components of certain thymic malignancies. Accordingly, the revised discussion acknowledges that exclusion of such areas may artificially homogenize the attenuation profile and represents a limitation of our approach.

Third, no a priori power analysis was performed due to the retrospective nature of the study. We have clarified this in the Methods section and noted that adequate statistical power was supported for the primary outcome measure by a post-hoc effect size assessment.

Finally, the statistical analysis was streamlined to remove repetition, as recommended by the reviewers. Redundant descriptions of applied tests (lines 89–95 vs. 97–108 in the original submission) were consolidated to ensure clarity and methodological coherence.

Taken together, our findings support the emerging role of CT histogram-derived metrics—particularly mean attenuation and distribution shape parameters—as non-invasive, accessible, and reproducible biomarkers that can assist in differentiating benign from malignant anterior mediastinal masses. However, given the moderate AUC and limited specificity, these parameters should be interpreted cautiously and used in conjunction with clinical evaluation, morphologic CT findings, and multidisciplinary judgment rather than as standalone diagnostic criteria.

Prospective multicenter studies incorporating volumetric texture analysis, radiomics-based machine learning models, and external validation cohorts are needed to establish standardized diagnostic thresholds and further delineate the role of CT quantitative imaging in thoracic oncology.

CONCLUSION

This study underscores the emerging diagnostic potential of CT histogram analysis as a non-invasive, quantitative imaging modality for preoperative risk stratification in anterior mediastinal masses. By evaluating first-order features, such as mean attenuation and distribution shape metrics (i.e., skewness and kurtosis), we demonstrated that malignant lesions are characterized by significantly higher, more peaked, and right-skewed HU profiles, whereas benign lesions exhibit broader heterogeneity.

The high sensitivity (96.7%) and moderate specificity (50%) associated with a data-driven HU threshold suggest that histogram mean values may serve as a complementary imaging biomarker for malignancy detection rather than a reliable standalone biomarker. In accordance with the reviewer's recommendations, all statements implying direct reduction of unnecessary surgery were revised to emphasize that CT histogram metrics do not independently determine surgical candidacy but instead function as supportive adjuncts within

a broader diagnostic framework. Although not sufficient in isolation, histogram metrics can substantially complement morphological assessment, especially in cases with ambiguity.

Importantly, this approach leverages routine non-contrast CT imaging, offering a low-cost, widely accessible tool that requires no additional scanning protocols or contrast agents, making it particularly appealing in resource-limited settings or when contrast is contraindicated. Relying on standard imaging techniques also enhances reproducibility and facilitates integration into existing radiology workflows without additional software requirements.

When integrated into the preoperative workflow alongside clinical and morphologic data, CT histogram analysis can support surgical triage and assist—rather than determine—in identifying cases that may warrant closer surveillance versus earlier surgical intervention, thereby promoting more targeted, patient-specific decision-making.

Prospective multicenter validation studies incorporating advanced texture and radiomic parameters are needed to further delineate the role of CT-based quantitative imaging in thoracic oncology and to develop standardized thresholds that can be implemented in everyday practice. To address the current methodological limitations acknowledged in this study, future work should specifically evaluate volumetric histogram analysis, machine-learning classifiers, and inter-observer reproducibility.

Ethics

Ethics Committee Approval: This retrospective study was approved by the Ethics Committee of Antalya Training and Research Hospital (approval number: 2025-125, approval date: 08 May 2025).

Informed Consent: The requirement for informed consent was waived due to the retrospective nature of the study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: N.Ç.Y., Concept: N.Ç.Y., M.Y., Design: N.Ç.Y., O.K., Data Collection or Processing: N.Ç.Y., M.Y., Analysis or Interpretation: N.Ç.Y., M.Y., Literature Search: N.Ç.Y., O.K., Writing: N.Ç.Y.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Ke L, Liu J, Shuai Y, et al. Intercostal approach VATS is feasible for large-sized anterior mediastinal tumors. *Sci Rep.* 2024;14(1):17227. **[Crossref]**
2. Ackman JB, Verzosa S, Kovach AE, et al. High rate of unnecessary thymectomy and its cause. Can computed tomography distinguish thymoma, lymphoma, thymic hyperplasia, and thymic cysts? *Eur J Radiol.* 2015;84:524-533. **[Crossref]**
3. Heider J, Koch RA, Kross R, editors. Diseases of the chest, heart and vascular system 2025–2028. *Diagnostic Imaging.* 2025. **[Crossref]**
4. Gao C, Yang L, Xu Y, et al. Differentiating low-risk thymomas from high-risk thymomas: preoperative radiomics nomogram based on contrast enhanced CT to minimize unnecessary invasive thoracotomy. *BMC Med Imaging.* 2024;24:197. **[Crossref]**
5. Zhang Z, Mi K, Wang Z, et al. Using optimized CT type to predict histological classifications of thymic epithelial tumors: a radiomics integrated analysis. *Insights Imaging.* 2025;16(1):67. **[Crossref]**
6. Liu M, Zhang Y, Liu LH. Histogram analysis based on unenhanced CT for identifying thymoma and lymphoma among prevascular mediastinal incidentalomas. *Cancer Imaging.* 2024;24(1):5. **[Crossref]**
7. Xia H, Yu J, Nie K, Yang J, Zhu L, Zhang S. CT radiomics and human-machine hybrid system for differentiating mediastinal lymphomas from thymic epithelial tumors. *Cancer Imaging.* 2024;24(1):163. **[Crossref]**
8. Roden AC, Szolkowska M. Common and rare carcinomas of the thymus. *Virchows Arch.* 2021;478(1):111-128. **[Crossref]**
9. Nam JG, Goo JM, Park CM, Lee HJ, Lee CH, Yoon SH. Age- and gender-specific disease distribution and the diagnostic accuracy of CT for resected anterior mediastinal lesions. *Thorac Cancer.* 2019;10:1378-1387. **[Crossref]**
10. Al-Khalifa M, Alsaad S, Al-Kulaib A, Al-Muawda M, Al-Tareef H, Darwish A. Histopathological characteristics of mediastinal tumors through 11 years. *Bahrain Med Bull.* 2020;42:13-19. **[Crossref]**
11. Juanpere S, Cañete N, Ortuño P, Martínez S, Sanchez G, Bernado L. A diagnostic approach to the mediastinal masses. *Insights Imaging.* 2013;4(1):29-52. **[Crossref]**
12. Rajamohan N, Goyal A, Kandasamy D, et al. CT texture analysis in evaluation of thymic tumors and thymic hyperplasia: correlation with the international thymic malignancy interest group (ITMIG) stage and WHO grade. *Br J Radiol.* 2021;94(1128):20210583. **[Crossref]**
13. Yasaka K, Akai H, Abe O, Ohtomo K, Kiryu S. Quantitative computed tomography texture analyses for anterior mediastinal masses: differentiation between solid masses and cysts. *Eur J Radiol.* 2018;100:85-91. **[Crossref]**

Thorac Res Pract. 2026;27(4):235-240 **DOI:** 10.4274/ThoracResPract.2026.2025-12-2

Original Article

Not All Corridors are Equal: Corridor Length Affects 6-minute Walk Test Performance–A Randomized Crossover Trial

Melissa Köprülüoğlu, Elvan Felekoğlu, Büşra Türk, İlknur Naz

Department of Physiotherapy and Rehabilitation, İzmir Katip Çelebi University Faculty of Health Sciences, İzmir, Türkiye

Cite this article as: Köprülüoğlu M, Felekoğlu E, Türk B, Naz İ. Not all corridors are equal: corridor length affects 6-minute walk test performance–a randomized crossover trial. *Thorac Res Pract.* 2026;27(4):235-240

ABSTRACT

OBJECTIVE: The recommended corridor length for the 6-minute walk test (6MWT) is 30 meters. However, it is not always feasible in clinical practice. The objective of this study was to compare walking distance and hemodynamic responses during 6MWTs performed in 10-, 20-, and 30-meter corridors in healthy young adults.

MATERIAL AND METHODS: Thirty-four healthy young adults (mean age: 21.2±1.5 years; female/male: 25/9) completed three randomized 6MWTs in 10-, 20-, and 30-meter corridors on separate days and with a crossover design. Walking distance, mean walking speed, and step counts were measured. Heart rate, blood pressure, respiratory rate, peripheral oxygen saturation, dyspnea, and fatigue were recorded at baseline (Pre), immediately post-test (Post), and at 3 (R3) and 5 (R5) minutes of recovery. Data were analyzed using the Friedman test with post-hoc Wilcoxon signed-rank tests and Bonferroni correction.

RESULTS: Walking distance and speed were significantly lower in the 10-meter corridor than in the 20- and 30-meter corridors ($P < 0.05$), whereas step counts did not differ ($P = 0.065$). Hemodynamic responses were similar in all tests, except for increased post-test fatigue and dyspnea at R5 during the 10-meter corridor test ($P < 0.05$).

CONCLUSION: Although hemodynamic responses were comparable across different corridor lengths, the 10-meter corridor led to an underestimation of walking distance by ~60 meters and to increased subjective fatigue and dyspnea. A 20-meter corridor may serve as a feasible alternative when the standard 30-meter corridor is unavailable in clinical practice, whereas a 10-meter corridor should be avoided.

KEYWORDS: 6-minute walk test, walking distance, corridor length, exercise testing

Received: 03.12.2025 ***Revision Requested:*** 03.03.2026 ***Last Revision Received:*** 22.04.2026

Accepted: 05.06.2026 ***Epub:*** 02.07.2026 ***Publication Date:*** 24.07.2026

INTRODUCTION

The 6-minute walk test (6MWT) is one of the most widely used field tests to evaluate functional exercise capacity in various populations. Its simplicity, reproducibility, and ability to reflect the integrated responses of the musculoskeletal and cardiopulmonary systems have made it a popular tool in both clinical and research settings.[1] The walking distance is the primary outcome in 6MWT, which asks the patient to walk as far as they can along a 30-meter corridor in six minutes.[2] Reference equations incorporating variables such as age, gender, height, and weight have been established to interpret test results.[3-5] Functional capacity is then assessed based on the percentage of predicted values achieved.

The American Thoracic Society (ATS) and the European Respiratory Society (ERS) guidelines state that the 6MWT should be performed in a 30-meter corridor.[2] However, in clinical practice, environmental and spatial constraints often necessitate using shorter corridors. This limitation has prompted several investigations into whether corridor length affects test performance. Previous studies, mostly conducted in patient populations such as individuals with stroke or chronic respiratory diseases, have suggested that shorter corridor lengths may alter test outcomes.[6-10]

Corresponding author: Melissa Köprülüoğlu, PhD, PT, e-mail: kopruluoglumelissa@gmail.com

Despite this evidence, the majority of research has focused on clinical populations, and data on the influence of corridor length on 6MWT performance in healthy individuals remain limited. In a current study, Chuatrakoon et al.[11] revealed that <20-meter corridors varied results compared to the standard length, and a walkway of >15 meter had a very strong agreement with a standard walkway. Understanding these effects in healthy young adults is important to establish baseline evidence and to inform clinical interpretation. The present study aimed to compare 6MWT performance and hemodynamic responses in healthy young adults across 10-, 20-, and 30-meter corridors. We hypothesized that a 10-meter corridor would lead to earlier onset of fatigue, more rapid increases in vital signs, and shorter walking distances.

MATERIAL AND METHODS

Participants and Pre-test Assessments

The study was conducted at the Research, Education, and Innovation Laboratories of the Department of Physiotherapy and Rehabilitation, İzmir Katip Çelebi University. An a priori sample size calculation was performed using G*Power V3.1.9.7. With an effect size of 0.25, a 95% confidence interval, and a 5% margin of error, a minimum of 33 participants was required.[12] Anticipating a 20% dropout, 40 participants were initially recruited. Ultimately, 34 participants completed all stages of the study (Figure 1).

Eligible participants were between 18 and 25 years of age, and we excluded participants who had chronic diseases diagnosed by a physician, were receiving regular medical treatment, and had physical limitations that could impair walking ability. Demographic characteristics, including age, sex, height, weight, and smoking status, were recorded. Before testing, all participants rested for 10 minutes, and their baseline hemodynamic parameters were assessed.

Test Procedures

This study was conducted using a randomized crossover design and was completed in accordance with the CONSORT Extension for Crossover Trials checklist. Each participant completed the 6MWT on three occasions following a familiarization session, once at each corridor length (10, 20, and 30 meters), with the order randomized, using computer-generated sequences, by a blinded physiotherapist who assessed the participants. The ATS guidelines were followed for the familiarization session. In this session, we explained the task, had the participant walk alone at their own pace for 6 minutes, and used the standardized encouragement script on a separate day.[2] Tests were scheduled on separate days between 08:00 a.m. and 12:00 p.m. to minimize circadian variability, and the order of tests was randomized to reduce potential learning effects.

Main Points

- The 6-minute walk test (6MWT) is an important tool fto assess exercise capacity.
- In conditions where a 30-meter corridor is not available, a 20-meter corridor may be sufficient for a 6MWT.
- The results of a 6MWT conducted in a 10-meter corridor will differ from those of a 20- and 30-meter corridor.

A one-day break was given between the three tests. Participants were instructed to avoid activities that would cause excessive fatigue and to abstain from heavy alcohol consumption or supplement use during this period. Participants were excluded if they were fasting, had eaten immediately before testing, had engaged in strenuous exercise within the 24 hours prior to testing, had smoked, or had used medications/supplements that might influence performance.

Every test was conducted in a calm, indoor corridor with a level, straight floor and markings spaced three meters apart. The turning points were indicated with two cones. Participants were instructed to walk as far as possible for six minutes at their own pace. Standardized encouragement was provided at one-minute intervals, consistent with ATS/ERS guidelines.[2,13]

Outcome Measures

The primary outcome was walking distance. Secondary outcomes included mean walking speed, step count, heart rate, systolic and diastolic blood pressure, respiratory rate, peripheral oxygen saturation (SpO_2), dyspnea, and fatigue. Walking distance, speed, and step counts were recorded using a smartwatch (Fitbit Charge 5, United States). Predicted walking distance was calculated according to Cazzoletti et al.'s[14] formula, considering age, sex, height, and values of the participants. In addition, the percentage of the predicted maximum heart rate according to Tanaka et al.[15] was recorded for each participant. Dyspnea and fatigue were evaluated using a visual analogue scale (0–100) because it is a sensitive and reproducible tool.[16,17] Hemodynamic variables were assessed at four time points: pre-test (Pre), immediately post-test (Post), 3 minutes into recovery (R3), and 5 minutes into recovery (R5).

Ethical Considerations

The entire study process was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained from the İzmir Katip Çelebi University Non-Interventional Research Ethics Committee (approval no: 382, date: 22.09.2022, 2024/08/05). All study details were explained to all participants, and written informed consent was obtained.

Statistical Analysis

The data obtained from the research were analyzed using the SPSS 22.0 statistical software package. Descriptive data are presented as mean ± standard deviation for normally distributed continuous variables, median (25/75 interquartile range) for non-normally distributed continuous variables, and percentage (%) for categorical variables. The walking distance was defined as the primary outcome measure. The Friedman test was used to compare the effects of the test corridors on outcomes. If significant differences were found, the Wilcoxon signed-rank test with Bonferroni correction was applied for pairwise comparisons.

Intraclass correlation coefficient (ICC) analysis with a two-way mixed-effects model was performed to examine the reliability of the test across three conditions. ICC values were interpreted as follows: <0.50, poor; 0.50–0.75, moderate; 0.75–0.90, good; and >0.90, excellent reliability. Effect sizes for the Friedman

test were expressed as Kendall's coefficient of concordance (Kendall's W). Kendall's W values of 0.10, 0.30, and 0.50 were interpreted as small, moderate, and large effects, respectively.

RESULTS

Participant Characteristics

A total of 34 participants who met the study's inclusion criteria were enrolled (Figure 1). Six participants did not complete all tests. Descriptive characteristics are presented in Table 1. Most participants were female, non-smokers, and had body mass index within the normal range. Mean age was 21.2±1.5 years.

Hemodynamic and Symptom Responses

Pre-test values for heart rate, blood pressure, respiratory rate, SpO_2, dyspnea, and fatigue did not differ significantly across the three test conditions ($P > 0.05$). Post-test and R3 values were also comparable across corridor lengths, except for fatigue, which was significantly higher following the 10-meter 6MWT than after the 20- and 30-meter corridors ($P = 0.048$). At R5, dyspnea was significantly greater after the 10-meter corridor than after the 20-meter corridor ($P = 0.029$), while all other parameters remained similar (Table 2).

Walking Performance

Walking distance differed significantly across the three corridor lengths ($P < 0.001$). Pairwise comparisons indicated that distances achieved in the 10-meter corridor were significantly shorter than those in the 20- and 30-meter corridors (both $P < 0.001$), whereas no difference was observed between the 20- and 30-meter corridors (Figure 2). On average, the walking distance in the 10-meter corridor was approximately 60 meters shorter than in the longer corridors.

Step counts did not differ significantly among conditions ($P = 0.065$). Walking speed, however, was considerably lower in the 10-meter corridor than in the 20-meter corridor ($P = 0.033$), whereas no difference was found between the 20- and 30-meter corridors (Table 3). Additionally, the percentage of predicted walking distance was lower and the percentage of maximum heart rate was higher during the 6MWT performed in a 10-meter corridor (Table 3).

Table 1. Descriptive variables of participants

Variables	Statistics (n = 34)
Age, (year)	
$\bar{x}$ ± SD	21.2±1.5
M (min-max)	21 (18-25)
Gender, n (%)	
Female	25 (73.5)
Male	9 (26.5)
Height, cm	
$\bar{x}$ ± SD	168.0±8.5
M (min-max)	165.0 (161.8-172.3)
Weight, kg	
$\bar{x}$ ± SD	62.3±12.7
M (min-max)	60.0 (51.5-67.6)
Body mass index, kg/m²	
$\bar{x}$ ± SD	22.0±2.7
M (min-max)	22.2 (18.6-24.1)
Cigarette consumption, n (%)	
Active smoker	5 (14.7)
Ex-smoker	2 (5.9)
Never smoked	27 (79.4)
Cigarette consumption amount, pack-year	
$\bar{x}$ ± SD	4.0±1.2
M (min-max)	4 (3-5)

$\bar{x}$: mean, SD: standard deviation, M: median, min: minimum, max: maximum

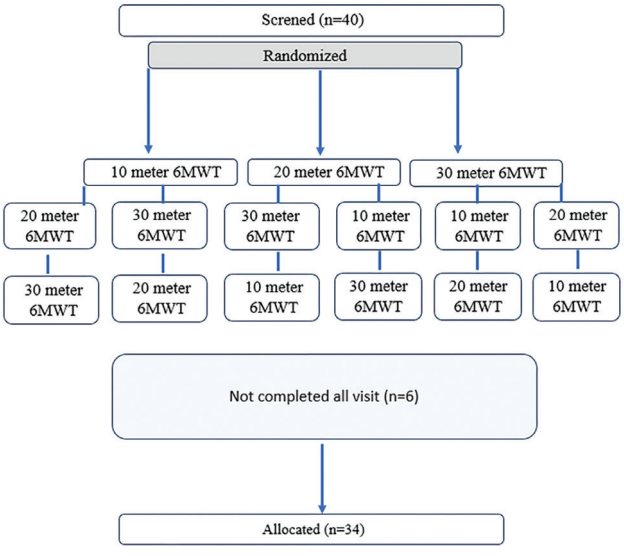

Figure 1. Flowchart and test procedures of the research

6MWT: 6-minute walk test

DISCUSSION

This study investigated whether corridor length influences performance and hemodynamic responses during the 6MWT in healthy young adults. The main findings were: (1) walking distance and walking speed were significantly lower in the 10-meter corridor compared with the 20- and 30-meter corridors; (2) step counts remained similar across all conditions; and (3) hemodynamic responses were largely comparable, except for higher post-test fatigue and greater dyspnea during recovery in the 10-meter corridor. This study aimed to investigate the impact of corridor length on various performance metrics (distance, step count, walking speed, and percentage of predicted values) and on detailed hemodynamic responses measured before, after, and during recovery, in contrast to earlier research that primarily focused on walking distance. In this respect, it differs from previous studies and offers a particular perspective.

The ATS and ERS recommended a 30-meter corridor for 6MWT.[2,13] However, it may not always be possible in a clinical setting due to the clinic's physical space and environmental factors. The circumstances prompted researchers to investigate whether different corridor lengths affected the 6MWT distance

Table 2. Comparison of the variables in pre-test, post-test, recovery 3rd min, and recovery 5th min

	10 m-6MWT	20 m-6MWT	30 m-6MWT	ICC (95%)	*P* values
Pre-test variables					
Heart rate, bpm	88.78±12.71	86.41±13.41	87.06±13.71	0.66 (0.50/0.80)	0.780
SBP, mmHg	113.74±13.69	116.79±14.78	114.32±17.19	0.27 (0.12/0.43)	0.639
DBP, mmHg	72.94±11.31	74.71±9.86	73.09±9.60	0.41 (0.23/0.54)	0.579
Respiratory rate, breaths per minute	25±3	25±3	25±3	0.36 (0.12/0.39)	0.631
Oxygen saturation, %	98.29±1.09	98.09±1.20	98.18±1.32	0.37 (0.24/0.51)	0.826
Dyspnea, 0-100	0.15±0.56	0.12±0.40	0.12±0.41	0.41 (0.29/0.68)	0.969
Fatigue, 0-100	0.41±0.16	0.50±0.17	0.50±0.05	0.71 (0.56/0.83)	0.947
Post-test (recovery 1st minute) variables					
Heart rate, bpm	98.21±18.02	96.74±19.76	98.47±22.18	0.58 (0.39/0.74)	0.269
SBP, mmHg	121.47±13.31	124.94±23.84	126.41±20.86	0.36 (0.19/0.53)	0.321
DBP, mmHg	73.65±13.83	75.55±8.22	77.29±15.03	0.29 (0.21/0.42)	0.234
Respiratory rate, breaths per minute	29±4	30±4	29±5	0.37 (0.30/0.53)	0.640
Oxygen saturation, %	97.80±1.03	97.09±3.81	97.88±1.87	0.51 (0.26/0.61)	0.580
Dyspnea, 0-100	2.85±0.83	2.76±0.52	2.88±0.53	0.70 (0.53/0.82)	0.827
Fatigue, 0-100	3.61±0.81[a]	3.15±0.33[b]	3.27±0.54[b]	0.53 (0.33/0.70)	0.048*
Recovery 3rd minute variables					
Heart rate, bpm	92.88±14.64	92.21±16.95	92.06±15.23	0.77 (0.63/0.87)	0.895
SBP, mmHg	117.50±13.17	113.91±24.67	118.53±17.48	0.30 (0.18/0.52)	0.620
DBP, mmHg	73.41±10.51	76.56±13.60	74.97±12.90	0.35 (0.28/0.48)	0.352
Respiratory rate, breaths per minute	25±4	25±3	24±4	0.57 (0.49/0.68)	0.154
Oxygen saturation, %	97.71±1.85	97.82±2.37	98.09±1.66	0.39 (0.24/0.51)	0.697
Dyspnea, 0-100	2.00±0.23	1.74±0.13	1.71±0.30	0.62 (0.44/0.77)	0.241
Fatigue, 0-100	2.35±0.39	2.29±0.22	2.35±0.50	0.52 (0.32/0.70)	0.697
Recovery 5th minute variables					
Heart rate, bpm	89.47±15.18	89.88±13.63	91.02±14.29	0.76 (0.62/0.86)	0.673
SBP, mmHg	114.97±20.35	112.64±10.88	113.52±12.86	0.42 (0.21/0.63)	0.647
DBP, mmHg	73.53±11.30	72.97±8.87	72.82±9.96	0.46 (0.25/0.66)	0.421
Respiratory rate, breaths per minute	24±4	24±4	24±3	0.40 (0.32/0.59)	0.941
Oxygen saturation, %	97.68±1.85	98.18±1.31	97.82±2.78	0.29 (0.17/0.51)	0.252
Dyspnea, 0-100	0.50±0.82[c]	0.29±0.58[d]	0.24±0.70[d]	0.58 (0.39/0.74)	0.029*
Fatigue, 0-100	1.25±0.76	1.02±0.74	1.46±0.91	0.51 (0.31/0.70)	0.708

All three groups were compared with the Friedman test. For pairwise comparison, the Wilcoxon signed-rank test with Bonferroni correction was applied. *$P < 0.05$ indicates a significant difference. The letters a, b, c, and d are used to indicate statistical differences. Test values with different letters indicate a statistically significant difference

6MWT: 6-minute walk test, ICC: intraclass correlation coefficient, SBP: systolic blood pressure, DBP: diastolic blood pressure, min: minute

Table 3. Comparison of test results

	10 m-6MWT	20 m-6MWT	30 m-6MWT	ICC (95%)	*P* values	Kendall's W
Walking distance, meter	684.19±55.48[a]	739.24±59.08[b]	742.73±61.84[b]	0.82 (0.71/0.90)	**<0.001***	0.72
Percentage of predicted walking distance, %	89.95±7.73	93.67±8.05	94.03±7.93	0.67 (0.59/0.81)	**<0.001***	0.66
Percentage of predicted maximum heart rate, %	64.82±9.24	58.51±10.17	59.96±10.49	0.57 (0.51/0.63)	**<0.001***	0.69
Step counts, step	757±52	750±56	745±55	0.87 (0.78/0.93)	0.065	0.16
Walking speed, meter/sec	1.34±0.15[c]	1.50±0.16[d]	1.48±0.17[d]	0.57 (0.49/0.71)	**0.30***	0.21

All three groups were compared with the Friedman test. For pairwise comparison, the Wilcoxon signed-rank test with Bonferroni correction was applied. *$P < 0.05$ indicates a significant difference. The letters a, b, c, and d are used to indicate statistical differences. Test values with different letters indicate a statistically significant difference

6MWT: 6-minute walk test, ICC: intraclass correlation coefficient, sec: second

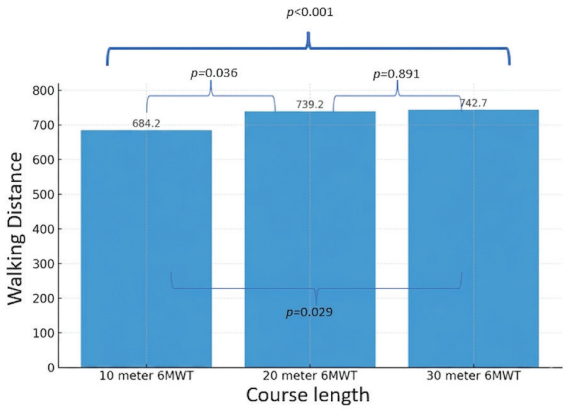

Figure 2. Comparing to walking distance in 10-, 20-, and 30-meter 6MWTs

6MWT: 6-minute walk test

and hemodynamic response to the test. Most of the research conducted on chronic obstructive pulmonary disease[7,9,10] examined individuals using one corridor length; however, studies of individuals with different chronic diseases and healthy individuals used different corridor lengths.[6,8,18,19] According to the findings of the systematic review by Ngueleu et al.[20], there is evidence that the conditions under which the 6MWT is administered have a substantial impact on the score. However, there is a study that suggests this impact was not clinically relevant.[21] Studies used corridors of 10, 15, 20, and 30 meters showed that shorter corridor lengths negatively affected walking distance, with the longest distances being achieved in corridors of 30 meters,[6,8,18] and the reference equations created for 30-meter corridors cannot be used for different corridor lengths.[7,9] Only one study by Gochicoa-Rangel et al.[8], which compared the results of the 6MWT performed in a corridor of 15 and 30 meters in patients with chronic respiratory disease, found similar test results. In healthy adults, Chuatrakoon et al.[11] found that <20-meter corridors varied results compared to the standard length. Our study supported previous studies' findings regarding 6MWT results in different corridors. The longest walking distance was recorded in the 30-meter corridor and was statistically similar to that recorded in the 6MWT performed in the 20-meter corridor. The shortest walking distance, the lowest predicted walking distance, and the highest predicted maximum heart rate were recorded in the 10-meter corridor. The shorter 6MWT distance achieved on a 10-meter corridor may be due to the increased number of turns required on such a short corridor.[7,18] More turns require more deceleration. This is supported by the observation that participants had a lower average speed in the 10-meter corridor despite similar step counts in our study. In addition, our participants exhibited higher fatigue values during the 10-meter 6MWT. This may be another reason for the reduced walking distance observed in the 10-meter corridor. Chuatrakoon et al.'s[11] study, conducted on healthy adults, showed similar results, although the average walking performance of the participants was found to be lower than our results. Since that study did not include predicted values, step count, and average speed, as our study did, no conclusions can be drawn. This difference may be attributable to the age range, gender distribution, and height of the included population. In this study, walking performance was interpreted using predictive equations adjusted for age, sex, weight, and height thereby reducing potential bias related to demographic differences. Despite these adjustments, age- and sex-related physiological variability in cardiovascular responses may still influence hemodynamic outcomes. Thus, caution is warranted when extrapolating these findings beyond the studied population.

The 6MWT differed significantly across corridor lengths ($P < 0.001$), with a large effect size Kendals W = 0.72. Although the ICC indicated good agreement between the tests (ICC: 0.82), participants walked a substantially shorter distance in the 10-m corridor than in the 20-m and 30-m corridors. The distances obtained from the 20-m and 30-m corridors were very similar, indicating that the 20-m corridor may provide comparable results to the standard 30-m corridor. Despite relatively high ICC values, the systematic reduction in walking distance observed in the 10-m corridor indicates that this configuration should not be considered a comparable alternative to longer corridors.

Responses to the 6MWT, including walking distance, are also important. For this reason, post-test heart rate, SpO_2, dyspnea, and fatigue were frequently examined in studies.[7,8,10,18] According to the study results, at the end of the test there was a minimal increase in heart rate, dyspnea, and fatigue, and a minimal decrease in SpO_2, particularly among patient groups. Additionally, different corridor lengths in 6MWTs did not affect test responses. However, recovery responses were not examined in these studies. In our study, we measured heart rate, SpO_2, blood pressure, respiratory rate, dyspnea, and fatigue at pre-test, post-test, R3, and R5. Responses were similar across all time periods. However, post-test fatigue and R5 dyspnea were higher when the 6MWT was performed in the 10-meter corridor. This difference may be due to a greater number of turns in the shorter corridor. The deceleration-acceleration cycle during turns may have contributed to increased fatigue and dyspnea.

Strengths of this study include the randomized crossover design, standardized testing procedures, and inclusion of recovery measurements, which have been underreported in prior research. However, several limitations should be acknowledged. Because only healthy young adults aged 18–25 years were included, the findings may not be generalizable to older adults or individuals with comorbidities. Age-related physiological and clinical changes could alter the observed relationships. Further research involving a wider age spectrum is warranted. Second, the sample was predominantly female, which may have influenced outcomes. Third, physical activity levels were not assessed, which could have affected walking performance. Another limitation of this study is that period, sequence, and potential carryover effects inherent to the crossover design were not formally evaluated. Although the order of the corridor conditions was randomized and rest periods were provided between tests to minimize these effects, their potential influence cannot be completely ruled out. Finally, although a smartwatch was used to objectively measure walking speed and step counts, validation of this device for use in the 6MWT is warranted.

CONCLUSION

This study demonstrated that corridor length significantly affects 6MWT outcomes in healthy young adults. While hemodynamic responses were largely comparable across 10-, 20-, and 30-meter corridors, the 10-meter corridor consistently produced shorter walking distances—approximately 60 meters shorter—and higher subjective fatigue and dyspnea than in the longer corridors. These findings suggest that using a 10-meter corridor may compromise the validity of the 6MWT. In contrast, a 20-meter corridor produced results comparable to those obtained under the standard 30-meter condition, suggesting that it may be a feasible alternative for clinical settings with limited space.

Ethics

Ethics Committee Approval: The entire study process was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained from the İzmir Katip Çelebi University Non-Interventional Research Ethics Committee (approval no: 382, date: 22.09.2022, 2024/08/05).

Informed Consent: All study details were explained to all participants, and written informed consent was obtained.

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.K., Concept: M.K., B.T., Design: M.K., E.F., İ.N., Data Collection or Processing: M.K., B.T., Analysis or Interpretation: M.K., E.F., İ.N., Literature Search: M.K., B.T., E.F., İ.N., Writing: M.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: This study was supported by TÜBİTAK 2209-a projects.

REFERENCES

1. Matos Casano HA, Anjum F. Six-minute walk test. In: StatPearls. StatPearls Publishing; 2023. [Crossref]
2. Holland AE, Spruit MA, Troosters T, et al. An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. *Eur Respir J.* 2014;44(6):1428-1446. [Crossref]
3. Gibbons WJ, Fruchter N, Sloan S, Levy RD. Reference values for a multiple repetition 6-minute walk test in healthy adults older than 20 years. *J Cardiopulm Rehabil.* 2001;21(2):87-93. [Crossref]
4. Chetta A, Zanini A, Pisi G, et al. Reference values for the 6-min walk test in healthy subjects 20–50 years old. *Respir Med.* 2006;100(9):1573-1578. [Crossref]
5. Halliday SJ, Wang L, Yu C, et al. Six-minute walk distance in healthy young adults. *Respir Med.* 2020;165:105933. [Crossref]
6. Ng SS, Tsang WW, Cheung TH, Chung JS, To FP, Yu PC. Walkway length, but not turning direction, determines the six-minute walk test distance in individuals with stroke. *Arch Phys Med Rehabil.* 2011;92(5):806-811. [Crossref]
7. Beekman E, Mesters I, Hendriks EJ, et al. Course length of 30 metres versus 10 metres has a significant influence on six-minute walk distance in patients with COPD: an experimental crossover study. *J Physiother.* 2013;59(3):169-176. [Crossref]
8. Gochicoa-Rangel L, Ramírez-José MC, Troncoso-Huitrón P, et al. Shorter corridors can be used for the six-minute walk test in subjects with chronic lung diseases. *Respir Investig.* 2020;58(4):255-261. [Crossref]
9. Klein SR, Gulart AA, Venâncio RS, et al. Performance difference on the six-minute walk test on tracks of 20 and 30 meters for patients with chronic obstructive pulmonary disease: Validity and reliability. *Braz J Phys Ther.* 2021;25(1):40-47. [Crossref]
10. Saiphoklang N, Pugongchai A, Leelasittikul K. Comparison between 20 and 30 meters in walkway length affecting the 6-minute walk test in patients with chronic obstructive pulmonary disease: a randomized crossover study. *PLoS One.* 2022;17(1):e0262238. [Crossref]
11. Chuatrakoon B, Seepang N, Chaiwong D, Nanavichit R, Rerkasem K, Nantakool S. The agreement of the various distance walkway in the 6-minute walk test in healthy adults. *PLoS One.* 2025;20(4):e0321503. [Crossref]
12. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Erlbaum; 1988. [Crossref]
13. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med.* 2002;166(1):111-117. [Crossref]
14. Cazzoletti L, Zanolin ME, Dorelli G, et al. Six-minute walk distance in healthy subjects: reference standards from a general population sample. *Respir Res.* 2022;23(1):83. [Crossref]
15. Tanaka H, Monahan KD, Seals DR. Age-predicted maximal heart rate revisited. *J Am Coll Cardiol.* 2001;37(1):153-156. [Crossref]
16. Grant S, Aitchison T, Henderson E, Christie J, Zare S, McMurray J, Dargie H. A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. Chest. 1999;116(5):1208-1217. [Crossref]
17. Capodaglio EM. Comparison between the CR10 Borg's scale and the VAS (visual analogue scale) during an arm-cranking exercise. *J Occup Rehabil.* 2001;11(2):69-74. [Crossref]
18. Ng SS, Yu PC, To FP, Chung JS, Cheung TH. Effect of walkway length and turning direction on the distance covered in the 6-minute walk test among adults over 50 years of age: a cross-sectional study. *Physiotherapy.* 2013;99(1):63-70. [Crossref]
19. Fell B, Hanekom S, Heine M. A modified six-minute walk test (6MWT) for low-resource settings – a cross-sectional study. *Heart Lung.* 2022;52:117-122. [Crossref]
20. Ngueleu AM, Barrette S, Buteau C, et al. Impact of pathway shape and length on the validity of the 6-minute walking test: a systematic review and meta-analysis. *Sensors (Basel).* 2024;25(1):17. [Crossref]
21. Heinz P, Gulart AA, Klein SR, et al. A performance comparison of the 20 and 30 meter six-minute walk tests among middle aged and older adults. *Physiother Theory Pract.* 2021;37(7):817-825. [Crossref]

Thorac Res Pract. 2026;27(4):241-249 **DOI:** 10.4274/ThoracResPract.2026.2025-12-3

Original Article

Use of Impulse Oscillometry in Patients with Chronic Obstructive Pulmonary Disease

Burcu Öztürk Şahin[1], Nurdan Köktürk[2]

[1]Department of Pulmonary Diseases, University of Health Sciences Türkiye, Gülhane Faculty of Medicine, Ankara, Türkiye
[2]Department of Pulmonary Diseases, Gazi University Faculty of Medicine Ankara, Türkiye

Cite this article as: Öztürk Şahin B, Köktürk N. Use of impulse oscillometry in patients with chronic obstructive pulmonary disease. *Thorac Res Pract.* 2026;27(4):241-249

ABSTRACT

OBJECTIVE: Although spirometry is a standard diagnostic tool for obstructive lung diseases, it is rarely performed, and the quality of basic spirometry is low in many countries. The impulse oscillometry system (IOS) is a non-invasive method that requires minimal patient cooperation and, therefore, can be performed even in the pediatric population. It has been also suggested that IOS may help to clarify the relationship between small airway disease and the underlying mechanisms of chronic obstructive pulmonary disease (COPD). The primary objective of the study was to compare IOS parameters in COPD patients with those in healthy individuals. Our secondary objective was to determine the relationship between IOS parameters and standard pulmonary function tests (PFTs). Our hypothesis was that airway resistance detected by IOS would be higher in COPD patients than in controls. Hence, IOS would provide findings comparable to and correlated with those of standardized PFTs for small airway obstruction.

MATERIAL AND METHODS: A total of 104 subjects (62 patients with COPD and 42 healthy non-smoking individuals) were included in the study. All subjects underwent spirometry, diffusing capacity for carbon monoxide (DLCO), lung-volume measurements, and IOS.

RESULTS: COPD patients showed significant decreases in forced expiratory volume in one second (FEV_1), forced vital capacity (FVC), FEV_1/FVC, FEF 25–75, and DLCOadj, and significant increases in residual volume (RV) and total lung capacity (TLC) compared with the control group. In terms of IOS parameters, R5%, Z5%, Fres, and R5–R20/R5% values were significantly higher in the COPD group ($P = 0.029$, $P = 0.022$, $P = 0.009$, $P = 0.004$, respectively). The reactance area (AX) value, defined as the AX, was also significantly increased in the COPD group ($P = 0.004$). The correlation between FEF 25–75% (L/s) and R5–R20 was moderate and negative ($r = -0.491$, $P < 0.001$). A weak correlation ($r = 0.240$, $P = 0.017$) was also found between the RV/TLC ratio and R5–R20.

CONCLUSION: This study showed that airway resistance was increased in the COPD group and that IOS parameters were associated with measures of small-airway function in standard PFTs. IOS can be used as a non-invasive, patient-friendly method that complements PFTs by providing a comprehensive assessment of COPD pathology and pathophysiological changes and detecting changes in symptomatic patients.

KEYWORDS: Impulse oscillometry, chronic obstructive pulmonary disease, impedance

Received: 05.12.2025 ***Revision Requested:*** 31.12.2025 ***Last Revision Received:*** 18.01.2026

Accepted: 22.01.2026 ***Epub:*** 09.03.2026 ***Publication Date:*** 24.07.2026

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a common, preventable, and treatable chronic inflammatory disease characterized by airflow limitation and respiratory symptoms resulting from airway and/or alveolar destruction.[1,2] It results from an inflammatory process initiated by harmful gases and particles and by developmental host factors. COPD affects millions of people worldwide and is a major cause of morbidity and mortality.[3]

Pulmonary function tests (PFT) should be performed in patients with dyspnea, chronic cough and/or mucus production, a history of exposure to various particles, or risk factors for COPD. Spirometry is the most objective and best-standardized method for demonstrating airflow limitation; it is simple and reproducible.[2]

Corresponding author: Burcu Öztürk Şahin, MD, e-mail: brcztrk7@gmail.com

Impulse oscillometry system (IOS) is a variant of the forced oscillation technique and is an easy, non-invasive method for assessing lung function. It can be easily used with young children and older adults, since the procedure does not require patient cooperation. Moreover, it provides more detailed information than spirometry, allowing it to detect differences between central and peripheral airways, as well as early changes in small airways. The oscillations are applied at a fixed square-wave frequency of 5 Hz, from which all other frequencies of interest are derived; these are typically multiples of 5 Hz (5 to 35 Hz) and last between 30 and 40 ms each.[4] Low-frequency waves (5 Hz) reach the periphery of the lungs, whereas high-frequency waves (20 Hz) reach the upper airways.[4] In IOS, impedance (Z) represents the opposition to oscillatory airflow in the respiratory system and reflects the combined resistive and reactive components. The two main components of impedance are resistance (R) and reactance (X). These are measured in the frequency range 5–20 Hz and are named according to these frequencies. For example, R and X measured at 5 Hz are denoted R5 and X5. These values are recorded simultaneously with each breath during the measurement. By measuring resistance and reactance, IOS may better assess distal airway pathology, a key component of COPD.[5,6] Since IOS is a non-invasive method that requires minimal active participation by the patient, it can also be used routinely for the diagnosis and follow-up of COPD.[7] In theory, the elastic properties of the lung are reflected in low-frequency reactance, which corresponds to the distal airways.[8] Studies have shown that IOS can be used to evaluate respiratory system impedance, assess lung heterogeneity at low frequencies, detect airway resistance earlier than spirometry in symptomatic patients, and define different subgroups of COPD patients.[9] In addition to these data, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2022 states that IOS may help to better understand the relationships between small airway disease and the underlying disease mechanisms of COPD, such as clinically significant symptoms.[9] IOS can be used to complement spirometry for this purpose. In addition, IOS is thought to be valuable for detecting individuals with pathological airway changes (small airway disease) typical of COPD.[10] With these considerations in mind, we hypothesized that parameters derived from IOS that reflect small airway dysfunction –particularly the reactance area (AX) and indices representing the relative contribution of peripheral airway resistance (R5–R20/R5)– would be more sensitive indicators of COPD-related pathophysiology than absolute resistance values alone. Therefore, IOS can provide complementary information to standard PFTs in the assessment of small airway involvement in COPD. The aim of this study was to compare IOS parameters in patients with stable COPD and healthy individuals, to examine the relationship between spirometry parameters, and to evaluate whether IOS parameters reflecting small airway dysfunction, particularly AX and R5–R20/R5 values, differ between the two groups. A secondary aim was to investigate the relationship between these IOS parameters and traditional pulmonary function indices [including FEF 25–75 and residual volume (RV)/total lung capacity (TLC)] associated with small airway obstruction and air trapping. Our hypothesis is that relative and reactance-based IOS indices will show stronger associations with small airway dysfunction than absolute resistance measurements.

Main Points

- One of the main points of this article is to demonstrate changes in chronic obstructive pulmonary disease (COPD) patients by performing impulse oscillometry system (IOS) testing on healthy adults and patients diagnosed with COPD.
- Another key point is the correlation between standard spirometry and IOS.

MATERIAL AND METHODS

Our study is a cross-sectional case-control study. Patients aged 40 and older who presented to the Department of Pulmonology Outpatient Clinic of Gazi University Hospital, Health Research and Application Center between October 2020 and September 2021 were included in the study. An informed consent form was obtained from all participants. Ethics committee approval was obtained from the Gazi University Faculty of Medicine Clinical Research Ethics Committee with the decision numbered 533 and dated 07.09.2020.

Inclusion Criteria

Two groups of patients were enrolled in the study.

COPD Group

- Patients aged 40 and over,
- Patients with stable COPD diagnosed with GOLD spirometric criteria,[1]
- Patients with no other pulmonary pathology,
- COPD patients who applied for routine outpatient controls and were requested a PFT.

Control Group

- Those aged 40 and over,
- Healthy, non-active smoking volunteers.

Smokers were defined as follows:[11]

- Smoking at least one cigarette per week for at least 6 months,
- Smoking at least 100 cigarettes in a lifetime.

Individuals who smoked and quit within the last 15 years were classified as ex-smokers.

Those who have smoked previously but have quit, and those who have not smoked in the last 15 years, were included in the non-smokers group.

Exclusion Criteria

- Patients with any pulmonary disease,
- The subjects who have coronavirus disease-2019 (COVID-19) positive test results,
- Those who did not agree to participate in the research.

Study Design

Demographic characteristics of the patients (age, sex, smoking history, occupation), current complaints (cough, phlegm, shortness of breath, hemoptysis, chest pain), history of pulmonary disease, current pulmonary diseases, and comorbidities were assessed and recorded. The cases were evaluated with symptoms, physical examination and spirometry measurements [forced expiratory volume in 1 second (FEV_1) (L and %), forced vital capacity (L and %), FEV_1/FVC (%), peak expiratory flow-(PEF) (%), forced expiratory flow (FEF) 25-75 (L/s)], GOLD stage and GOLD spirometric grade for patients with a diagnosis of COPD, lung volumes [TLC (L and %) if measured, functional residual capacity (FRC) (L and %), expiratory residual volume (ERV) (L and %), RV (L and %) and the ratio of RV to TLC (%)] and diffusing capacity for carbon monoxide (DLCO) hemoglobin adjusted (%) test results.

IOS measurement results were recorded as R5 [kPa/(L/s) and %], X5 [kPa/(L/s) and %], AX (kPa/L), R20 [kPa/(L/s) and %], X20 [kPa/(L/s) and %], Z5 [kPa/(L/s) and %] and R5–R20.

Spirometric Maneuvers

Spirometry was performed using the MasterScreen (CareFusion, Germany) instrument. All tests were performed consecutively on the same day, with the patient resting; the patient had rested for at least 1 hour and had not smoked. The patient rested for one hour before undergoing IOS. The FVC maneuver was performed by an experienced technician with the patient seated upright and the nose clipped. The best of the 3 measurements in which patients took a deep inspiration followed by a forced expiration was considered. The largest observed FVC and largest FEV_1 of all acceptable values, were reported, even if they did not come from the same maneuver.[12] In the report, prebronchodilator FEV_1 (L and %), FVC (L and %), FEV_1/FVC (%), PEF (%), and FEF 25–75 (L/s) values were recorded.

The single-breath-hold technique was used to measure DLCO. Hemoglobin values (g/dL) were recorded prior to measurements. The DLCO single-breath maneuver begins with several cycles of tidal breathing followed by a full exhalation to near RV. The patient then rapidly inhales the test gas mixture to full lung capacity and holds their breath for approximately ten seconds. Following the breath-hold, the patient exhales completely and rapidly; the first portion of exhaled gas is expelled, and an alveolar sample is then collected. DLCO is then calculated from the concentrations of carbon monoxide and tracer gas in the sample. Following DLCO measurement by the single-breath-hold method, lung volumes were measured by nitrogen washout using the same device. Measured TLC (L and %), FRC (L and %), ERV (L and %), RV (L and %), and the ratio of RV to TLC (%) were recorded. When clinically feasible, lung volume measurements were obtained using the nitrogen washout method; they were not available for all participants, reflecting real-life clinical practice during the COVID-19 period.

Patients who underwent PFT were subsequently assessed using the Masterscreen IOS (CareFusion, Germany, 2016). For spirometric measurements using IOS, the device was calibrated each morning before measurements. Three acceptable measurements were made for each patient, and the average of the three acceptable R measurements across all frequencies was recorded. During the test, the patients were kept in a sitting position with their heads upright. While seated in this position, patients were asked to hold the disposable mouthpiece between their teeth and to close their lips tightly around it. The patient was asked to support both cheeks to prevent swelling of the upper airway shunt. While the patient maintains relaxed breathing, the device delivers small pressure impulses into the airway. No forced or deep breaths are required. The system then analyzes the resulting pressure–flow oscillations during tidal breathing to measure respiratory resistance and reactance across a range of frequencies.

The study was conducted during the COVID-19 pandemic, a period in which PFT was subject to significant restrictions. Therefore, spirometry and IOS were performed only in patients with documented negative severe acute respiratory syndrome-coronavirus-2 test results. To minimize procedure duration and potential aerosol exposure, measurements were limited to pre-bronchodilator assessments; post-bronchodilator testing could not be routinely performed.

Statistical Analysis

We reported descriptive statistics as mean and standard deviation for continuous variables with a normal distribution, and as median, minimum, and maximum values for continuous variables with a non-normal distribution. Frequencies and percentages were used to describe categorical variables. To determine the normality of continuous variables, the Shapiro-Wilk test results and the histograms, box plots, and Q-Q plots were evaluated. Based on the assumption of normal distributions, two independent groups (study groups) were compared using either the independent-samples t-test or the Mann-Whitney U test. For categorical variables, differences between groups were examined using the chi-square test or Fisher's exact test. The relationship between continuous variables was evaluated using Spearman's rank correlation coefficient because the assumption of normality was not met. To evaluate diagnostic accuracy, receiver operating characteristic curve analysis was performed. *P* values of 0.05 or lower were considered statistically significant. Statistical analyses were performed using IBM SPSS version 23.

Given the real-world and pandemic-related constraints of the study, no a priori power analysis was performed. The study was therefore designed as an exploratory and hypothesis-generating investigation.

RESULTS

A total of 104 patients aged 40 years and older who presented to the Pulmonology department outpatient clinic between October 2020 and September 2021 were included in the study. The COPD group included 62 patients aged 40 and over who were diagnosed with stable COPD and were attending routine outpatient clinic follow-up visits. In the control group, 42 healthy, non-smoking patients aged 40 years and over were included. Table 1 summarizes the demographic characteristics and pulmonary complaints in the two groups.

When the patients' spirometry and DLCO measurements were examined, all spirometric parameters and DLCO values were lower in the COPD group. RV (L) and RV/TLC% values were increased in the COPD group. These differences were found to be statistically significant. Comparisons of spirometry and DLCO values between the groups are presented in Tables 2 and 3.

While the patients' IOS values were examined, resistance was calculated using pressure waves at different frequencies (5-20 Hz) [R5 kPa/(L/s), R5%, R20 kPa/(L/s), R20%, R5-20 kPa/(L/s)]. A significant difference was observed in R5% values between the COPD and control groups ($P = 0.029$). The median R5 value, which provides information about the resistance of the entire respiratory system, was 0.59 kPa/(L/s) (184%) in the COPD group and 0.52 kPa/(L/s) (172%) in the control group. When the R20 values, which provide information about the large airways were examined, they were 0.36 kPa/(L/s) (133.5%) in the COPD group and 0.4 kPa/(L/s) (133%) in the control group. No significant difference was observed between R20 values ($P = 0.782$). When the R5–R20 results were examined, no significant difference was found between the groups. No statistically significant difference was found ($P = 0.079$).

Reactance was calculated using pressure waves at different frequencies (5–20 Hz): X5 kPa/(L/s), X5%, X20 kPa/(L/s), and X20%. The X5 value, which is expected to decrease in distal airway obstruction, was significantly lower in the COPD group than in the control group ($P < 0.009$). Similarly, the X20 value, which is expected to decline more slowly than X5 in distal airway obstruction, decreased significantly in the COPD group compared with the control group ($P < 0.014$).

Impedance (Z) was calculated using pressure waves at a frequency of 5 Hz in two ways: Z5 [kPa/(L/s)] and Z5%. Although no significant difference was observed between the groups for Z5 kPa/(L/s) —the sum of the forces required to propagate the pressure wave applied to the respiratory system— a significant difference was observed for Z5% values, which were higher in the COPD group ($P = 0.200$ and 0.022, respectively). The reactance area (AX) value, was significantly higher in the

Table 1. Comparative demographic characteristics of patient groups

	COPD group (n = 62)	Control group (n = 42)	Test statistic	*P* value
Age (years)	63.5 (40–80)	56.5 (40–72)	2.699	**0.007**
Sex				
Female	9 (14.52)	19 (45.24)	12.012	**0.001**
Male	53 (85.48)	23 (54.76)		
Height (cm)	166 (146–182)	162.5 (143–185)	1.114	0.265
Weight (kilograms)	76.5 (42–133)	80.5 (36–120)	1.286	0.198
BMI	28.13±5.93	30.10±5.34	1.721	0.088
Smoking history				
Non-smoker	5 (8.06)	31 (73.81)	54.010	**<0.001**
Ex-smoker	26 (41.94)	11 (26.19)		
Active smoker	31 (50.00)	0		
Pack year	38 (2–135)	12 (2–35)	4.017	**<0.001**
Symptoms				
Active				
Pulmonary	56 (90.32)	0 (0)	82.194	**<0.001**
Complaints	23 (37.10)	1 (2.38)	16.999	**<0.001**
Cough	23 (37.10)	0 (0)	20.005	**<0.001**
Phlegm	0 (0)	0 (0)	NA	NA
Hemoptysis	3 (4.84)	0 (0)	NA	0.271
Chest pain	55 (88.71)	1 (2.38)	75.085	**<0.001**
Shortness of breath	56 (90.32)	0 (0)	82.194	**<0.001**
mMRC				
Scale				
0	1 (1.82)		NA	NA
1	12 (21.82)			
2	27 (49.09)			
3	13 (23.64)			
4	2 (3.64)			

Descriptive statistics are given as median (minimum–maximum) or mean ± standard deviation
COPD: chronic obstructive pulmonary disease, BMI: body mass index, mMRC: modified Medical Research Council, NA: not applicable

Table 2. Comparison of spirometric values by patient groups

	COPD group (n = 62)	Control group (n = 42)	Test statistic	*P* value
FEV_1 (L)	1.52 (0.38–3.69)	2.86 (1.62–5.23)	6.238	**<0.001**
FEV_1 (%)	62.02±23.84	108.14±16.97	11.524	**<0.001**
FVC (L)	2.79 (0.83–5.20)	3.53 (1.94–5.03)	3.124	**0.002**
FVC (%)	81.03±23.80	108.83±17.38	6.486	**<0.001**
FEV_1/FVC (%)	62 (27–70)	81 (72.5–96)	8.632	**<0.001**
PEF (L)	4.31 (1.26–11.7)	6.64 (3.49–10.63)	4.790	**<0.001**
PEF (%)	59 (17–143)	101 (68–137)	6.411	**<0.001**
FEF 25 (L)	2.3 (0.44–8.05)	6.14 (3.34–10.38)	6.593	**<0.001**
FEF 25 (%)	36.5 (8–116)	105 (67–140)	7.633	**<0.001**
FEF 50 (L)	1.1 (0.15–3.73)	3.8 (1.85–6.54)	8.023	**<0.001**
FEF 50 (%)	28 (3–90)	92 (58–154)	8.302	**<0.001**
FEF 75 (L)	0.35 (0.14–1.52)	1.12 (0.52–2.16)	7.450	**<0.001**
FEF 75 (%)	28 (4–102)	69.5 (40–141)	7.467	**<0.001**
FEF 25–75(L)	0.80 (0.17–3.19)	2.89 (1.31–4.88)	8.083	**<0.001**
FEF 25–75 (%)	27 (5–94)	89 (51–138)	8.241	**<0.001**

Descriptive statistics are given as median (minimum–maximum) or mean ± standard deviation
COPD: chronic obstructive pulmonary disease, FEV_1: forced expiratory volume in one second, FVC: forced vital capacity, FEF: forced expiratory flow

Table 3. Comparison of DLCO measurement results by groups

	COPD group (n = 62)	Control group (n = 42)	Test statistic	*P* value
DLCO (%)	71.316±24.575	100.927±18.227	6.528	**<0.001**
TLC (L)	4.935±1.317	4.936±1.018	0.015	0.988
TLC (%)	82.632±16.802	90.293±13.950	2.386	0.019
FRC (L)	3.14 (1.42–6.1)	2.99 (1.06–26.2)	1.412	0.158
FRC (%)	98 (45–181)	97 (60–140)	0.187	0.851
ERV (L)	0.80 (0.01–2.35)	0.85 (0.03–2.67)	0.709	0.478
ERV (%)	83 (1–284)	97 (4–236)	0.699	0.485
RV (L)	2.18 (1.12–4.51)	1.91 (1.2–3.02)	3.104	**0.002**
RV (%)	100 (54–195)	94 (73–154)	1.192	0.233
RV/TLC (%)	48 (27–71)	37 (27–55)	4.091	**<0.001**

Descriptive statistics are given as median (minimum–maximum)or mean ± standard deviation
DLCO: carbon monoxide, COPD: chronic obstructive pulmonary disease, TLC: total lung capacity, FRC: functional residual capacity, ERV: expiratory residual volume, RV: residual volume,

COPD patient group ($P < 0.004$). When the Fres values of other parameters measured with the IOS were compared, a significant difference was observed, with higher Fres values in the COPD group ($P < 0.009$). All results of the IOS measurements, along with comparisons between the groups, are presented in Table 4. Under physiological conditions, peripheral airway resistance is known to constitute approximately 10–30% of total respiratory resistance.[13] In obstructive diseases with peripheral airway obstruction, peripheral resistance contributes a significantly greater proportion of total resistance. Based on this, the (R5–R20)/R5 ratio was examined and found to be 36.3% in the COPD group and 28.4% in the control group ($P = 0.004$). A comparison of the R5–R20/R5 ratios is shown in Table 4.

To examine the relationship between the severity of obstruction and oscillometric parameters in patients with COPD, GOLD grades 2 and 3 were compared. Because each of the GOLD grade 1 and GOLD grade 4 groups contained only one patient, these groups were excluded from the analysis. No significant differences were observed between the groups in the values of R5 [kPa/(L/s)], indicating total airway resistance, and R5–R20 [kPa/(L/s)], indicating ventilation heterogeneity. Similarly, no significant difference was observed between the groups in the values of X5 kPa/(L/s), X20 kPa/(L/s), and Z5 kPa/(L/s), the latter of which represents the sum of the forces required to move the pressure wave in the respiratory system. Only the AX kPa/L value increased as obstruction severity increased, and this increase was significant ($P = 0.032$). Table 5 shows a comparison of IOS results across spirometric GOLD grades.

Correlation

We examined the correlation between the spirometry measurement FEF 25–75, which indicates small airway obstruction, and the IOS measurement R5–R20. A moderate negative correlation was observed between FEF 25–75% and R5–R20 (Spearman correlation coefficient r = -0.49, $P < 0.001$; Figure 1). When the correlation between the RV/TLC ratio, an

Table 4. Comparison of IOS measurement results by groups

	COPD group (n = 62)	Control group (n = 42)	Test statistic	*P* value
R5 [kPa/(L/s)]	0.59 (0.24–1.29)	0.52 (0.28–0.99)	1.165	0.244
R5 (%)	184 (79–406)	172 (92–252)	2.185	**0.029**
R20 [kPa/(L/s)]	0.36 (0.2–0.68)	0.4 (0.25–0.63)	0.551	0.582
R20 (%)	133.5 (65–257)	132.5 (80–198)	0.391	0.696
R5–R20	0.18 (0.03–0.82)	0.13 (0.02–0.51)	1.754	0.079
X5 [kPa/(L/s)]	-0.17 (-0.7–0.21)	-0.12 (-0.33–0.42)	2.613	**0.009**
X5 (%)	656 (-2996–6443)	277 (-7432–5129)	3.707	**<0.001**
X20 [kPa/(L/s)]	-0.02 (-0.24–0.11)	0 (-0.11–0.07)	2.459	**0.014**
X20 (%)	-58.5 (-4492–729)	-5.5 (-2122–5675)	1.729	0.084
AX (kPa/L)	1.77 (0.14–8.52)	0.87 (0.19–4.22)	2.879	**0.004**
Z5 [kPa/(L/s)]	0.62 (0.24–1.39)	0.54 (0.29–1.03)	1.282	0.2
Z5 (%)	197 (82–474)	175 (93–268)	2.299	**0.022**
Fres	24.03 (9.97–39.54)	19.94 (12.96–34.6)	2.607	**0.009**
R5–R20/R5 (%)	36.311±16.015	28.462±11.255	2.921	**0.004**

Descriptive statistics are given as median (minimum–maximum) or mean ± standard deviation
COPD: chronic obstructive pulmonary disease, AX: reactance area

Table 5. Comparison of IOS results by spirometric GOLD grade

	GOLD grade 2 (n = 50)	GOLD grade 3 (n = 10)	Test statistic	*P* value
R5 [kPa/(L/s)]	0.57 (0.29–1.29)	0.77 (0.39–1.24)	1.413	0.158
R5 (%)	173.5 (95–406)	263 (127–406)	1.666	0.096
R20 [kPa/(L/s)]	0.37 (0.21–0.68)	0.37 (0.22–0.55)	0.010	0.992
R20 (%)	132.5 (65–257)	144 (85–207)	0.436	0.663
R5–R20 [kPa/(L/s)]	0.17 (0.03–0.82)	0.35 (0.04–0.76)	1.330	0.183
R5–R20/R5 (%)	32.19 (6.25–67.44)	49.07 (25.64–65.52)	1.908	0.056
X5 [kPa/(L/s)]	-0.15 (-0.62–0.21)	-0.19 (-0.7– -0.12)	1.953	0.051
X5 (%)	656 (–2996–6443)	1035.5 (527–5255)	1.533	0.125
X20 [kPa/(L/s)]	–0.02 (–0.24–0.01)	-0.09 (-0.24–0.04)	1.520	0.129
X20 (%)	–39.5 (–4492–729)	-166.5 (-792–95)	1.210	0.226
AX (kPa/L)	1.59 (0.14–8.12)	4.39 (0.77–8.52)	2.151	**0.032**
Z5 [kPa/(L/s)]	0.61 (0.30–1.39)	0.81 (0.32–1.35)	0.903	0.367
Z5 (%)	186 (98–421)	275 (104–474)	1.240	0.215
Fres	23.65 (9.97–39.54)	29.09 (15.22–36.28)	1.230	0.219

Descriptive statistics are given as median (minimum–maximum) or mean ± standard deviation
IOS: impulse oscillometry system, GOLD: Global Initiative for Chronic Obstructive Lung Disease, AX: reactance area

indicator of air trapping, and R5–R20 was examined, a weak association was observed (Spearman correlation coefficient r = 0.24, *P* = 0.017) (Figure 2).

DISCUSSION

Current COPD guidelines recommend using spirometry to assess airflow limitation along with symptoms and exacerbation history.[1] Spirometry is the most widely used PFT. However, spirometry may fail to detect early changes in respiratory function. Some studies have shown that IOS parameters can reflect pathological changes such as airway obstruction, air trapping, and decreased compliance in COPD patients to a certain extent, as with traditional PFT parameters.[14,15] For this reason, IOS can be used as a non-invasive, patient-friendly method that complements PFTs, provides a comprehensive assessment of COPD pathology and associated pathophysiological changes, and detects early changes in symptomatic patients. Therefore, IOS has emerged as a complementary approach to traditional PFTs.[4] IOS is performed during normal breathing for approximately 30–40 seconds and is more physiological than spirometry because it does not require extra effort. It provides more detailed information than spirometry and should be considered a complement to it for comprehensive assessment of pulmonary function. We conducted this study to compare IOS parameters between patients with COPD and healthy individuals.

In our study, the non-smoking COPD group differed significantly from the control group in R5% oscillometric values. However, no significant difference was observed between R20 and R5–R20 values for the two groups. Other significant IOS parameters

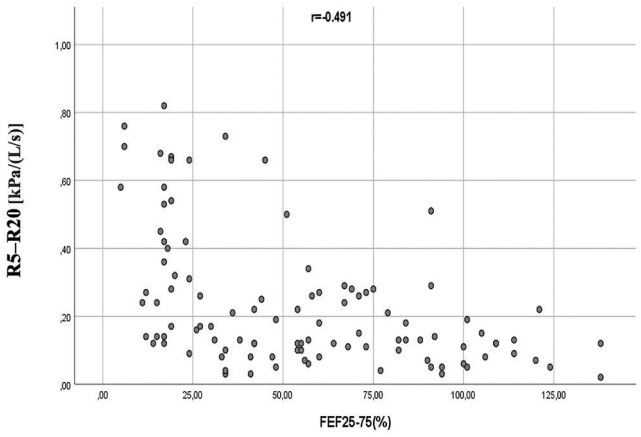

Figure 1. Spearman's correlation analysis between FEF 25–75% and R5–R20 [kPa/(L/s)]

FEF: forced expiratory flow

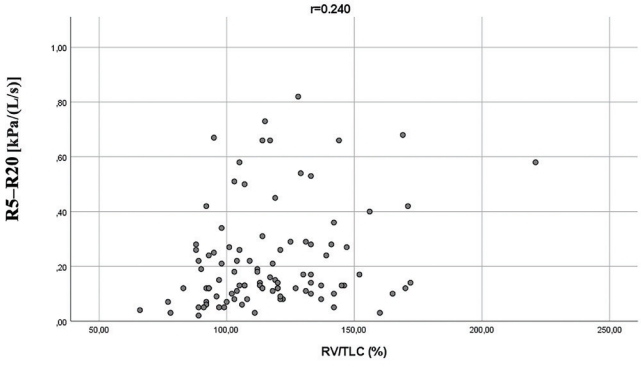

Figure 2. Spearman's correlation analysis between RV/TLC and R5–R20 [kPa/(L/s)]

RV/TLC: residual volume/total lung capacity

included X5 [kPa/(L/s)], X5 (%), X20 [kPa/(L/s)], AX (kPa/L), Z5 (%), Fres, and R5–R20/R5 (%). The results for the control and COPD groups in our study are similar to those reported in the ECLIPSE study by Crim et al.[14] The ECLIPSE study reported that the oscillometric parameter values in the COPD group were significantly different from those in both healthy smokers and healthy non-smokers.

Potentially, the most important application area of IOS is the assessment and control of peripheral airways. In 2007, Oppenheimer et al.[16] reported that IOS allows assessment of peripheral airways in a manner similar to dynamic compliance. In our study, the median value of R5–R20, a sensitive index of small airway obstruction, was found to be 0.13 kPa/(L/s) in the control group and 0.18 kPa/(L/s) in the COPD group. This difference was not statistically significant. In a study by Tomalak et al.[17], the median R5–R20 in the non-obstructed group was 0.13 kPa/(L/s) [range, 0.01 to 0.72 kPa/(L/s)], whereas in the obstructed group it was 0.32 kPa/(L/s) [range, 0.01 to 1.07 kPa/(L/s)]. Although the absolute values in our study were similar those in literature, no significant difference was observed between the control group and the COPD group. Although R5–R20 is frequently used as an indicator of small airway resistance, it is influenced by multiple physiological factors, including lung volume, age, and ventilation heterogeneity. In cross-sectional studies, absolute R5–R20 values may therefore lack the sensitivity to discriminate between groups with varying degrees of small airway involvement. In our study, the age difference between the two groups may explain this finding. The median age was 56.5 years in the control group and 63.5 years in the COPD group. One of the limitations of this study is our inability to control for age-related factors. Age and sex are known to influence respiratory mechanics and IOS measurements. In this study, the COPD group was older and predominantly male compared with the control group, which is consistent with the epidemiological characteristics of COPD. Although this imbalance may have contributed to differences in IOS parameters, performing formal multivariable adjustment would have led to both a substantial reduction in sample size and the introduction of potential selection bias. Therefore, the results should be interpreted as associations rather than causal effects.

Under physiological conditions, peripheral airway resistance accounts for approximately 10-30% of total respiratory resistance.[13] In obstructive diseases with obstruction in the peripheral airways, peripheral and total airway resistance are markedly increased. In our analysis, the R5–R20/R5 ratio was 28.4% in the control group, 36.3% in the COPD group, 32% in the GOLD 2 group, and 49% in the GOLD 3 group. In an analysis by Tomalak et al.[17], the R5–R20/R5 ratio was 33% in the group without airway obstruction and 50.7% in the group with airway obstruction.

In the present study, the absence of a statistically significant difference in absolute R5–R20 between COPD patients and controls does not contradict the presence of peripheral airway dysfunction. Instead, the significantly higher R5–R20/R5 ratio suggests a relatively greater contribution of peripheral airway resistance to total respiratory resistance in COPD. This proportional index may better capture the pathophysiological heterogeneity of COPD than absolute resistance values alone.

In the ECLIPSE study, the AX value was found to be 0.38 kPa/L in the non-smoking control group, 0.34 kPa/L in the smoker group without obstruction, and 1.99 kPa/L in the COPD group.[14] As the severity of obstruction increased from GOLD 2 to GOLD 4, the AX values were 1.37 kPa/L, 2.25 kPa/L, and 3.23 kPa/L, respectively. In our study, the AX values were 0.87 kPa/L in the non-smoking control group, 1.77 kPa/L in the COPD group, and 1.59 kPa/L and 4.39 kPa/L in the GOLD 2 and GOLD 3 groups, respectively. In our study, a significant difference in the AX value was observed between the two groups. Although the patient population in our study was smaller than in the ECLIPSE study, the AX values were similar. In a study by Wei et al.[15] of 215 patients, values for GOLD grades 1–4 were 0.66 kPa/L, 1.43 kPa/L, 2.07 kPa/L, and 2.5 kPa/L, respectively. Reactance-based parameters, particularly AX, are thought to reflect abnormalities in distal airway compliance and lung elastic properties. The significant increase in AX observed in the COPD group and its association with disease severity support the notion that reactance-based IOS indices may be more sensitive markers of small airway dysfunction in COPD than resistance-based parameters alone.

In the study conducted by Wei et al.[15], IOS resistance parameters Z5%, R5 kPa/L/s, R5%, R5–R20 kPa/L/s, and R5–R20/R5% showed a gradual increase with increasing GOLD grade, while no such increase was observed for R20 kPa/L/s and R20%. This suggests that the increase in total resistance was solely attributable to increased resistance in the peripheral airways. In our study, although a gradual increase in the absolute values of R5% and R5 kPa/L/s, Z5 kPa/L/s and Z5%, R5–R20 kPa/L/s, and R5–R20/R5% with increasing GOLD grade was observed, these increases were not statistically significant. In the study by Wei et al.[15], the reactance indices X5 kPa/L/s, Fres, and AX, reflecting peripheral airways, lung tissue, and thoracic elastic resistance, showed a gradual change with increasing GOLD grade. Kanda et al.[18] reported significantly more negative X5 values, reflecting small airway collapse, in COPD patients with severe pulmonary dysfunction.

Our results indicate that although total airway resistance was higher in the COPD group, no significant difference was found in R5–R20 values, which reflect small-airway resistance. A significant difference was found between the R5–R20/R5 (%) and AX (kPa/L) values. This may be related to differences in age and demographic characteristics between the two groups, suggesting that these parameters may be more sensitive. In our study, consistent with the literature, X5 (kPa/L/s), Fres, and AX values were higher in GOLD grade 3 than in grade 2. A significant difference in AX values was found between GOLD grades 2 and 3. However, because the number of patients in GOLD grades 1 and 4 was 1, only GOLD grades 2 and 3 were compared. Although the lack of uniform GOLD grades might be a limitation, the aim of the study was not to re-evaluate COPD diagnoses determined according to GOLD criteria but to investigate the relationship between IOS parameters and markers of small airway dysfunction in a real-world clinical setting.

Studies have shown a good correlation between IOS parameters and conventional lung function parameters in elderly patients.[19-21] As an alternative, IOS offers the advantage of ease of use. Wei et al.[15] observed moderate negative correlations between FEV_1% and Z5%, R5–R20, R5–R20/R5%, Fres, and AX (Spearman correlation coefficients r = 0.435, 0.425, 0.474, 0.5, and 0.521, respectively, $P < 0.001$). They also found moderate negative correlations between FEF 75–25% and %Z5, R5–R20, %R5–R20/R5, Fres and AX, reflecting small airway dysfunction (Spearman correlation coefficients r = 0.439, 0.452, 0.489, 0.510, and 0.540, respectively, $P < 0.001$). In our study, we observed a moderate negative correlation between R5–R20 and FEF 75–25% (Spearman correlation coefficient r = -0.49, $P < 0.001$), consistent with the findings of Wei et al.[15] found a weak correlation between RV/TLC and R5–R20 with r = 0.346, and we also found a weak correlation with r = 0.240 in our study ($P = 0.017$).

Study Limitations

IOS may have technical limitations. IOS-derived indices, as in the ECLIPSE study, have been compared with spirometry, and a lack of reference values obtained from healthy non-smokers has been reported.[14] According to this study, IOS can assess lung heterogeneity at lower oscillation frequencies and be used not only to monitor spirometric changes but also to identify subgroups of COPD patients.

The main limitation of this study is that it is a single-center study. A multicenter study could be designed to improve population representativeness. The age and gender distributions between groups were not equal. One major limitation of this study is that it was conducted during the COVID-19 pandemic. The absence of post-bronchodilator spirometry and lung volume measurements represents a significant limitation of this study. This situation was primarily related to the COVID-19 pandemic; during this period, lung function testing was restricted, and procedures were modified to reduce aerosol generation and patient exposure. As a result, all measurements were obtained under baseline conditions. Recruitment of a larger number of subjects because of the COVID-19 pandemic. The lack of prior power calculations and the relatively small sample size reflect the exploratory nature of the study. The distribution of GOLD grades was uneven, limiting the interpretation of subgroup analyses based on disease severity. These limitations should be considered when interpreting the results and underscore the need for larger prospective studies.

CONCLUSION

Although spirometry is the GOLD standard method for diagnosing COPD, its application can be challenging. IOS can be used in addition to spirometry, or in patients who are unable to perform spirometry. IOS provides complementary information regarding respiratory mechanics, particularly in relation to small airway dysfunction. In this study, reactance-based and relative IOS parameters, especially AX (kPa/L) and R5–R20/R5 (%), were more consistently associated with established markers of small airway involvement than absolute resistance measures alone. Total airway resistance was higher in the control group. Among the standard PFT parameters, FEF 25–75 and RV/TLC ratio were correlated with R5–R20; among the IOS parameters, R5–R20 was correlated with FEF 25–75 and RV/TLC ratio.

Given the cross-sectional design and methodological limitations, IOS should be regarded as an adjunct rather than a substitute for conventional PFT. Prospective studies with standardized post-bronchodilator measurements and larger sample sizes are warranted to further define the clinical utility of IOS in COPD.

Although IOS has been used since the 1970s, this test is still not widely adopted in clinical practice. Consistent with results in the literature, we highlight the use of IOS. IOS is an effortless, non-invasive method and is as effective as spirometry.

Ethics

Ethics Committee Approval: Ethics committee approval was obtained from the Gazi University Faculty of Medicine Clinical Research Ethics Committee with the decision numbered 533 and dated 07.09.2020.

Informed Consent: An informed consent form was obtained from all participants.

Footnotes

Authorship Contributions

Surgical and Medical Practices: B.Ö.Ş., Concept: N.K., Design: N.K., Data Collection or Processing: B.Ö.Ş., Analysis or Interpretation: B.Ö.Ş., N.K., Literature Search: B.Ö.Ş., N.K., Writing: B.Ö.Ş., N.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease (2021 Report) [Internet]. Fontana (WI): GOLD; 2020. Accessed February 27, 2026. **[Crossref]**
2. Şen E, Özgen Alpaydın A, eds. Türk Toraks Derneğiının GOLD 2021 Kronik Obstrüktif Akciğer Hastalığı (KOAH) Raporuna Bakışı [Turkish Thoracic Societyıs Perspective on the GOLD 2021 Chronic Obstructive Pulmonary Disease (COPD) Report] [Internet]. Ankara: Optimus Yayıncılık; 2021. Accessed February 27, 2026. ISBN: 978-605-74460-0-8. **[Crossref]**
3. Mirici A. KOAH (Kronik Obstrüktif Akciğer Hastalığı) [Chronic Obstructive Pulmonary Disease (COPD)]. In: Mirici A, editör. Obstruktif Akciğer Hastalıkları (Havayolu Hastalıkları) [Obstructive Lung Diseases (Airway Diseases)]. Göğüs Hastalıkları, Bölüm 3. Ankara: TÜSAD; 2020. s. C1-C9. **[Crossref]**
4. Oostveen E, MacLeod D, Lorino H, et al. The forced oscillation technique in clinical practice: methodology, recommendations and future developments. *Eur Respir J.* 2003;22(6):1026-1041. **[Crossref]**
5. Contoli M, Bousquet J, Fabbri L, et al. The small airways and distal lung compartment in asthma and COPD: a time for reappraisal. *Allergy.* 2010;65(2):141-151. **[Crossref]**
6. Hogg JC, Chu F, Utokaparch S, et al. The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med.* 2004;350(26):2645-2653. **[Crossref]**
7. Frantz S, Nihlén U, Dencker M, Engström G, Löfdahl C-G, Wollmer P. Impulse oscillometry may be of value in detecting early manifestations of COPD. *Respir Med.* 2012;106(8):1116-1123. **[Crossref]**
8. Goldman MD. Clinical application of forced oscillation. *Pulm Pharmacol Ther.* 2001;14(5):341-350. **[Crossref]**
9. Jarenbäck L, Ankerst J, Bjermer L, Tufvesson E. Flow-volume parameters in COPD related to extended measurements of lung volume, diffusion, and resistance. *Pulm Med.* 2013;2013:782052. **[Crossref]**
10. Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Prevention, Diagnosis and Management of COPD (2022 Report) [Internet]. Fontana (WI): GOLD; 2021. Accessed February 27, 2026. **[Crossref]**
11. World Health Organization. WHO Report on the Global Tobacco Epidemic, 2008: The MPOWER Package [İnternet]. Geneva: WHO; 2008. Accessed February 27, 2026. ISBN: 9789241596282. **[Crossref]**
12. Miller M, Hankinson J, Brusasco V, et al. ATS/ERS task force: standardisation of lung function testing. *Eur Respir J.* 2005;26(2):319-338. **[Crossref]**
13. Gong SG, Yang WL, Liu JM, Liu WZ, Zheng W. Change in pulmonary function in chronic obstructive pulmonary disease stage 0 patients. *Int J Clin Exp Med.* 2015;8(11):21400-21406. **[Crossref]**
14. Crim C, Celli B, Edwards LD, et al. Respiratory system impedance with impulse oscillometry in healthy and COPD subjects: ECLIPSE baseline results. *Respir Med.* 2011;105(7):1069-1078. **[Crossref]**
15. Wei X, Shi Z, Cui Y, et al. Impulse oscillometry system as an alternative diagnostic method for chronic obstructive pulmonary disease. *Medicine.* 2017;96(46):8543. **[Crossref]**
16. Oppenheimer BW, Goldring RM, Herberg ME, et al. Distal airway function in symptomatic subjects with normal spirometry following World Trade Center dust exposure. *Chest.* 2007;132(4):1275-1282. **[Crossref]**
17. Tomalak W, Czajkowska-Malinowska M, Radliński J. Application of impulse oscillometry in respiratory system evaluation in elderly patients. *Pneumonol Alergol Pol.* 2014;82(4):330-335. **[Crossref]**
18. Kanda S, Fujimoto K, Komatsu Y, Yasuo M, Hanaoka M, Kubo K. Evaluation of respiratory impedance in asthma and COPD by an impulse oscillation system. *Intern Med.* 2010;49(1):23-30. **[Crossref]**
19. Williamson PA, Clearie K, Menzies D, Vaidyanathan S, Lipworth BJ. Assessment of small-airways disease using alveolar nitric oxide and impulse oscillometry in asthma and COPD. *Lung.* 2011;189:121-129. **[Crossref]**
20. Haruna A, Oga T, Muro S, et al. Relationship between peripheral airway function and patient-reported outcomes in COPD: a cross-sectional study. *BMC Pulm Med.* 2010;10(1):10. **[Crossref]**
21. Kolsum U, Borrill Z, Roy K, et al. Impulse oscillometry in COPD: identification of measurements related to airway obstruction, airway conductance and lung volumes. *Respir Med.* 2009;103(1):136-143. **[Crossref]**

Thorac Res Pract. 2026;27(4):250-260 **DOI:** 10.4274/ThoracResPract.2026.2026-2-5

Review

Working in Extreme Heat, an Invisible Occupational Risk: Legislation, Risk Management and Policy Recommendations for Türkiye: Turkish Thoracic Society Workshop Report on the Regulation of Working Conditions in Extreme Heat

Sebahat Genç[1], Selin Çakmakcı Karakaya[2], Sabri Serhan Olcay[3], Mikdat Kadıoğlu[4], İstemi Ceysu Dinç Ütükler[5]

[1]Clinic of Pulmonology, Private Medlife Medical Center, Muğla, Türkiye
[2]Subdivision of Work and Occupational Diseases, Clinic of Pulmonology, Ankara Atatürk Sanatorium Training and Research Hospital, Ankara, Türkiye
[3]Department of Pulmonology, Muğla Sıtkı Koçman University Faculty of Medicine, Muğla, Türkiye
[4]Department of Climate Science and Meteorological Engineering, İstanbul Technical University, İstanbul, Türkiye
[5]Department of Labour and Social Security Law, Yozgat Bozok University Faculty of Law, Yozgat, Türkiye

Cite this article as: Genç S, Çakmakcı Karakaya S, Olcay SS, Kadıoğlu M, Dinç Ütükler İC. Working in extreme heat, an invisible occupational risk: legislation, risk management and policy recommendations for Türkiye: Turkish Thoracic Society workshop report on the regulation of working conditions in extreme heat. *Thorac Res Pract.* 2026;27(4):250-260

ABSTRACT

Extreme heat, amplified by climate change, has become a predictable and escalating occupational hazard in Türkiye, increasing the burden of heat-related illnesses, traumatic workplace injuries, and productivity losses among both outdoor and indoor workers, particularly in labor-intensive sectors and among vulnerable populations. This Turkish Thoracic Society workshop report synthesizes interdisciplinary expert input and available evidence to propose a practical, prevention-focused framework for regulating working conditions in extreme temperatures. The recommendations were systematically developed through four thematic interdisciplinary working groups and finalized by plenary consensus. Key recommendations include explicitly defining employers' duties to assess, prevent, and manage extreme weather–related risks within Occupational Health and Safety Law No. 6331; establishing secondary regulations that operationalize temperature/Wet-Bulb Globe Temperature (WBGT)-triggered work–rest schedules, hydration and cooling requirements, and acclimatization protocols; and standardizing objective heat-stress monitoring using the WBGT index (aligned with ISO 7243). The report prioritizes engineering and administrative controls (shading, ventilation/cooling, shift adjustments, rest areas, and "buddy systems") over reliance on personal protective equipment, and calls for strengthened labor inspection capacity "through the lens of climate change." Additional policy actions include integrating heat-related outcomes into occupational disease and injury surveillance systems, enhancing inter-institutional coordination for rapid reporting and work-stoppage mechanisms during extreme heat events, and implementing multilingual, sector-specific training and emergency preparedness pathways. Protecting informal and migrant workers, seasonal agricultural laborers, and individuals with chronic cardiopulmonary disease is highlighted as a core priority for equity and climate justice.

KEYWORDS: Health policies, environmental disorders, occupational and environmental lung diseases, occupational disorders, workshop report

Received: 18.02.2026 ***Revision Requested:*** 18.02.2026 ***Last Revision Received:*** 23.04.2026

Accepted: 19.05.2026 ***Epub:*** 23.07.2026 ***Publication Date:*** 24.07.2026

Corresponding author: Sabri Serhan Olcay, MD, e-mail: serhanolcay@mu.edu.tr

INTRODUCTION

In the era we are currently in, which we call the Anthropocene-Capitalocene, the earth has begun to warm up due to the increasing production of carbon dioxide, a commonly known greenhouse gas. Each year is hotter than the previous one, and temperature records are being broken.

The World Health Organization (WHO) reports that there are approximately 175,000 deaths annually in Europe due to extreme heat.[1] Globally, this number is around 500,000, and deaths are increasing due to climate change and urbanization.[2] If this continues, deaths could reach 2.3 million by the end of the century.[3]

Türkiye is located in the Mediterranean climate zone, a region highly vulnerable to both the current and projected impacts of global climate change. Research indicates that the greatest increases in temperature in our country are predicted to occur in the Southeast, Aegean, and Mediterranean regions. If no measures are taken, it is predicted that summer temperatures in these regions will be 4–7 °C higher by 2100.[4] Rising summer temperatures and increasingly frequent heat waves adversely affect worker health, exposing both indoor and outdoor workers to increased risk of severe morbidity and potentially fatal outcomes. The human body's tolerance to heat has a critical limit between 40 °C and 50 °C. In conditions of extreme heat and humidity, and in direct sunlight, the risk of heat stress, heat syncope, and heatstroke increases after 20–30 minutes if adequate precautions are not taken.[2] If left untreated, it can rapidly worsen and lead to death. Children, older adults, individuals with chronic illnesses, and individuals with low socioeconomic status are at higher risk. To protect against heatstroke, exposure time should not exceed two hours at temperatures above 32 °C and one hour during high humidity and physical activity. It should also be noted that individual differences may exist.

Main Points

- Predictable occupational risk: Extreme heat is no longer an exceptional event but a permanent occupational hazard in Türkiye, significantly increasing the risk of heat-related morbidities, workplace accidents, and traumatic injuries.
- Legislative necessity: Current Turkish legislation lacks specific provisions for extreme temperatures; the report recommends explicitly defining the employer's obligation to manage climate-related risks within Occupational Health and Safety Law No. 6331 and establishing secondary regulations for work–rest arrangements.
- Standardized risk assessment: The Wet-Bulb Globe Temperature index should be the standard for objective heat stress monitoring, facilitating the implementation of dynamic, threshold-based work–rest schedules and acclimatization protocols.
- Protection of vulnerable groups: Strengthened protection mechanisms are required for high-risk populations, including informal migrant workers, seasonal agricultural workers, and individuals with chronic respiratory or cardiovascular diseases.
- Integrated action framework: A comprehensive policy approach must prioritize engineering and administrative controls—such as hydration plans, shading systems, and inter-institutional coordination—to ensure effective emergency response and medical surveillance.

In a study evaluating Occupational Health and Safety (OHS) in the construction sector in Türkiye,[5] it was reported that most occupational accidents occurred during the summer in 2007 and 2008. Long working hours and the adverse effects of hot summer weather on workers can be cited as reasons for this situation. Especially for those working in heavy, labor-intensive jobs in hot and humid environments, heat is a significant source of occupational hazard.[6]

When the body fluids lost through sweating in very hot environments are not adequately replenished, the risk of health problems increases. In work environments with high temperatures, acute health risks increase, and short-term fainting occurs when 10–12% of body fluid is lost. It is stated that overweight, underweight, and elderly workers are at greater risk.[7]

In our country, heat waves are not defined as disasters. However, in the European Bank's classification, it ranks first and is considered a silent, insidious disaster. In addition, deaths caused by heat waves are not recorded in our country.

In light of these findings and the increasing temperatures in our country, there is a clear need for specific, policy-oriented frameworks tailored to the Turkish occupational health context. Therefore, the Turkish Thoracic Society (TTS, *Türk Toraks Derneği*) Environmental Issues and Lung Health Working Group planned to hold a workshop titled "Regulation of Working Conditions in Extreme Heat". At the workshop held in Ankara on April 5, 2025, interdisciplinary knowledge sharing took place. The primary objective of this workshop was to address the existing knowledge gap by synthesizing interdisciplinary experts' opinions to develop a practical, prevention-focused framework and to formulate concrete policy recommendations for regulating working conditions in extreme temperatures.

WORKSHOP DEFINITIONS AND METHODOLOGY

Definitions

Heat Wave: Defined as an abnormally and disturbingly high level of heat, usually humid, lasting or expected to last for three days or more over a wide area.

Heat Stress: The net (overall) heat load an employee may be exposed to due to the combined contributions of metabolic temperature, environmental factors, and clothing requirements. It begins when the body cannot compensate for this excessive load. As heat stress approaches the limits of human tolerance, further increases become unacceptable and can lead to heat strain and the development of heat-related illnesses. Occupational heat stress and the risk of heat-related illness increase among outdoor and indoor workers who are exposed to excessive heat or whose work involves hot environments.

Heat Exhaustion: A more advanced stage of heat stress. Fluid and electrolyte loss occurs.

Heat Stroke: A stage requiring immediate medical intervention, where thermoregulation has completely collapsed**.**

Outdoor Workplace: Defined as workplaces where workers are exposed to weather conditions (such as solar radiation, humidity and temperature).

The Wet-Bulb Globe Temperature (WBGT) index is one of the most common environmental measures used to assess the risk of heat stress in work environments.[8] WBGT estimates the heat load of the human body by considering the ambient temperature, humidity, radiant heat (heat emitted by the sun or hot surfaces), and airflow. Thus, it helps to determine the level of risk to thermal comfort and health (e.g., heatstroke, dehydration, fatigue) for workers in hot environments.

Methods

This study was designed as a qualitative, consensus-based workshop aimed at formulating policy recommendations. The workshop aimed to consult the views and knowledge of all disciplines and institutions related to the subject, using purposive sampling. Invitation letters were sent to the Ministry of Health and the Ministry of Labour and Social Security, requesting the participation of their representatives in occupational health matters. The Association of Public Health Specialists (*Halk Sağlığı Uzmanları Derneği*), the Turkish Medical Association (*Türk Tabipleri Birliği*), the Occupational Lung Diseases Working Group of our society, the Ankara Municipality, the Confederation of Turkish Trade Unions (TÜRK-İŞ), the Revolutionary Workers' Unions Confederation (DİSK), and the International Labour Organization's (ILO) Türkiye Representation were also invited to participate, each with at least one representative. In addition, a lawyer working on "labor and social security law" and an academic meteorologist were also invited to the workshop. At least one representative from each component participated, resulting in a total of 24 experts.

In the morning sessions, experts presented theoretical information on the subjects to establish a shared scientific baseline. In the afternoon, participants were divided into four groups, each representing a different area of expertise, and each group developed proposals focusing on a specific theme. Each group discussion was guided by a moderator and recorded by a rapporteur. The topics were determined as follows: 1) legal regulations; 2) risk assessment; 3) training; and 4) emergency interventions. Suggestions were made after discussions with all groups. Following group discussions, the proposals were shared with all participants in a plenary assembly, revised based on feedback from open discussion, and finalized by qualitative consensus among all attendees rather than through a formal voting mechanism. This report contains the systematic recommendations developed during the workshop and the findings that can be evaluated at the policy level. As the primary objective was to synthesize multidisciplinary expert opinions into a practical framework, no statistical analysis was performed.

KEY COMPONENTS OF OCCUPATIONAL HEALTH AND SAFETY MANAGEMENT

The following sections present the direct outputs and consensus-based recommendations generated by the four thematic working groups. To distinguish the workshop's findings from general explanatory text, the specific proposals developed by the participants are outlined below as targeted recommendations. These outputs are structured around four main pillars identified as critical during the group discussions:

Legal Regulations

The legal system of the Republic of Türkiye is based on the principle of a hierarchy of norms. At the top of this hierarchy is the Constitution of the Republic of Türkiye, followed by international conventions, laws, Presidential decrees, and regulations. Each lower-level regulatory norm must be consistent with the norm at the next higher level. Strengthening protection against heat stress in occupational settings requires legal regulations at different levels of this hierarchy. Accordingly, the proposals developed in the workshop have been grouped under two main headings: primary (at the legal level) and secondary (at the regulatory level).

Primary legislation: proposals for regulation at the law level

The basic legal framework regarding OHS in Türkiye is determined by the OHS Law No. 6331.

Article 4 of the law requires the employer to conduct a risk assessment. Article 5 regulates the "principles of protection against risks". Clause (c) of this article stipulates that the employer must "combat risks at their source".

Scientific data clearly show that extreme temperatures resulting from climate change pose an OHS risk to employees. The increased frequency of deaths and illnesses due to exposure to extreme temperatures necessitates addressing this risk at its source in workplaces. Therefore, the employer's obligation to combat climate change is already implicit in current legal provisions.

Recommendation:

Add the following statement to Article 5 of Law No. 6331: "The employer is obliged to develop a policy for combating extreme weather conditions in the workplace and to analyze and prevent workplace risks related to extreme weather conditions".

This regulation is of great importance in raising awareness of occupational health risks associated with extreme weather events and climate change, institutionalizing a preventive approach, and protecting employee health.

Secondary legislation: proposals for regulations at the regulation level

In the hierarchy of norms, the second-level instruments subordinate to laws are Presidential Decrees and regulations. The changes and additions that may be made within the scope of secondary legislation are listed below:

a. Regulation on Labour Inspection

Currently, labor inspections are carried out by the Guidance and Inspection Directorate of the Ministry of Labour and Social Security. Inspections are carried out in two ways: "programmed inspection" and "unprogrammed inspection upon complaint".

Recommendations:

1. A change should be made in the Regulation on Labour Inspection to increase the competence of labor inspectors in recognizing and evaluating OHS risks related to climate change by providing them with specialized training. In addition,

2. A clear provision should be introduced that the inspections should be carried out "through the lens of climate change".

Thus, the inspection processes will cover not only the existing physical conditions but also the effects of climatic factors on occupational safety.

b. New Regulation Proposals

Current legislation does not contain specific regulations for workers in extreme heat or other extreme weather conditions. Therefore, it is proposed that new regulations addressing climate change and extreme weather events be developed. The proposed regulation titles are:

1. "Regulation on the Effects of Extreme Weather Events on Worker Health",

2. "Worker Health Regulation" (expanded to include climatic and environmental factors),

3. "Regulation on Workers in Hot, Enclosed Environments".

These regulations should specify implementation principles for topics such as risk assessment, workplace organization, protective measures, rest areas, health surveillance, and training.

c. Working Hours and Rest Arrangements

Recommendations:

1. Through an amendment to the Working Hours Regulation under the Labour Law, working hours and break arrangements during high temperatures should be defined by a temperature-threshold-based system.

For this purpose: a table should be prepared that determines working hours, break times, and fluid supplementation requirements according to air temperature,

2. Special provisions should be established for vulnerable groups and for workers who cannot take breaks because of the nature of their work. Accordingly, it is proposed that a regulation be introduced to prevent these groups from working between 10:30 and 15:30 when air temperatures are 32 °C or higher.

This provision will contribute both to reducing health risks caused by heat stress and to preventing workplace accidents.

Risk assessment and key topics

The most critical step in managing working conditions in extreme heat is establishing a comprehensive and dynamic risk assessment process.

Recommendations:

1. Defining the scope and context: clarifying the boundaries of the workplace, the activities performed, job descriptions, and monitoring mechanisms;

2. Identifying the hazard: defining the physiological and environmental components of heat stress and determining the influencing factors;

3. Determining the environmental heat: identifying indicators such as temperature, humidity, airflow, and radiant heat through field measurements; The use of the WBGT index (according to ISO 7243 standard),

4. Identification of risky areas: determining critical processes, work points, and time periods in terms of extreme heat,

5. Identification of vulnerable groups: identifying susceptible or unprotected groups (pregnant women, the elderly, those with chronic diseases, children and young workers, women, migrants, disabled individuals, etc.),

6. Classification of heat-related illnesses: identifying clinical presentations, early warning signs, and possible complications,

7. Assessment of probability and severity: analyzing the risk considering exposure duration, intensity, workload, and recovery possibilities,

8. Rating and determining the acceptability of the risk,

9. Establishment of control steps: planning engineering, administrative arrangements, and personal protective equipment (PPE) according to the prevention hierarchy,

10. Implementation, monitoring, and assignment of responsibilities: environmental and health monitoring and surveillance, training programs, emergency procedures,

11. Recording and communication of findings, including documentation, reporting, notification, and periodic review processes, should be implemented.

Legislation and policy recommendations

1. It is recommended that workplaces with a high risk of heat stress be categorized as the "very hazardous class", taking into account the WBGT index, and that appropriate OHS processes be implemented.

2. It is necessary to identify groups that require special policies, to consider thermal comfort a risk factor during the workplace establishment phase, and to determine measurement periods (e.g., the hottest months of the year).

3. The list of occupational diseases included in the current "Republic of Türkiye Regulation on the Determination of Loss of Earning Capacity in the Workforce and Occupation" does

not cover health effects related to thermal comfort. Addressing this deficiency requires publishing a new version.

Appendix 1 details how the WBGT index is used in the fields of OHS, sports, and military training, as well as the measurement methods and application examples. Thanks to these applications, heat stress is objectively monitored across different environments, and preventive and corrective measures are implemented according to risk level, thereby protecting the health and performance of workers, athletes, or soldiers.

Workplace implementation and prevention strategies

Recommendations:

1. Establishing early warning systems or seeking advice from scientific experts;

2. Defining an emergency communication chain and creating a higher authority notification mechanism;

3. Providing appropriate ventilation and shading systems;

4. Regulating working hours: frequent shift changes, short breaks, fluid and electrolyte support (water, buttermilk, etc.);

5. Expanding training programs: recognizing the risk, methods of coping with heat, and psychological support mechanisms (e.g., psychologist support for irritability and behavioral changes);

6. Improving rest areas: cooling, fluid intake, air circulation and air conditioning support;

7. Customizing health surveillance: implementing intensified and risk-based health monitoring protocols, and providing appropriate work clothing and PPE.

Inter-institutional coordination and monitoring mechanisms

Strengthening inter-institutional coordination is necessary for the early diagnosis and prevention of health problems in working life.

Recommendations:

1. Inspections of workplaces where health problems are reported should be increased.

2. The emergency notification and subsequent work stoppage mechanism should be clarified at the ministry level;

3. The activities carried out by subcontracting firms should be evaluated according to the main workplace hazard class and risk analyses, and health surveillance should be carried out accordingly.

4. Health surveillance should be provided through workplace nurses in environments where it is difficult to have a physician present at all times;

5. Dietitian consultation should be added.

6. The number of labor inspectors should be increased;

7. Specialist physicians in occupational health should be assigned to conduct inspections.

Informal employment and vulnerable groups

Informal employment poses a serious risk, especially for women, children, the elderly, and migrant workers.

Recommendations:

1. Family physicians should be informed about the risks associated with extreme heat and encouraged to be vigilant in reporting potential cases.

2. Training and awareness programs should be integrated into primary healthcare services.

3. Obstacles to unionization should be removed to facilitate workers' access to mechanisms for seeking their rights.

4. OHS materials and training for migrant workers should be prepared in their native languages.

Training

Sustaining measures to improve working conditions in extreme heat requires not only technical regulations but also a systematic, comprehensive training policy.

Recommendation:

Training should be planned at both the workplace and societal-awareness levels; formal, non-formal, and continuous training approaches should be implemented together.

The suggestions developed within this scope are summarized below.

Education policy and planning

Recommendations:

1. Educational content should be tailored to the nature of the work, the working environment, regional conditions, workplace and sector characteristics, geographical differences, and vulnerable groups.

2. In the mid- and long-term, training on heat stress for outdoor workers should be mandatory.

3. This training should be integrated into the Ministry of National Education curriculum, OHS laws, and emergency plans.

4. Considering the changing and increasingly severe environmental conditions due to climate change, educational curricula should be updated at regular intervals.

5. In sectors experiencing intense exposure to heat waves, such as agriculture, general public education should be conducted periodically, in view of the high risk faced by informal workers and small-scale businesses.

6. Personnel from organizations within the Ministry of Health should be assigned to these training programs; furthermore,

these topics should be integrated into widespread educational programs such as driving courses and military training.

7. Accessible telephone hotlines should be established for use during heat waves; personnel working on these hotlines should be specially trained.

8. In addition to group training in workplaces, individual counseling opportunities should be offered to employees.

Target groups for training

For training activities to be effective, the target audience should be addressed through a multi-layered structure. Accordingly, the following groups have been designated as priority training targets: politicians, managers in public institutions, managers of the Ministry of Health and the Ministry of Labour and Social Security, managers of municipalities and provincial special administrations, governors and district governors, employers and employees, trade unions and union representatives, professional organizations, relevant departments of universities, employer organizations, and civil society organizations.

Recommendation:

In training for these groups, the level of information and the depth of content should be differentiated. Policy development and supervision should be prioritized at the manager level, whereas protection and implementation skills should be prioritized at the employee level.

Training content and methodology

Recommendations:

The scope and methodology components of the trainings should include;

1. General information: climate change, physiological effects of heat stress, legislation and rights;

2. Specific information: regional, sectoral and vulnerable group-specific risks and precautions;

3. Protection and prevention: heat management in the workplace, hydration, break times, appropriate clothing selection and equipment use;

4. Action plans: preparation of specific action plans regarding excessive heat exposure in workplaces and integration of these plans into the training;

5. First aid and emergency response: practical training on steps to be taken in cases such as heatstroke, fainting, and dehydration;

6. Interactive training methods: case-based learning, simulations,

7. Practical sessions supported by audiovisual materials;

8. Support of face-to-face training with digital platforms;

9. Creation of a national training portal for standardization and accessibility of trainings;

10. Sector-specific training guides: preparation of separate guides for different sectors such as agriculture, construction, mining, and transportation;

11. Trainer training: organization of "train-the-trainer" programs to keep trainers' knowledge and skills up-to-date.

Emergency Interventions

Definition and classification of emergencies

Emergencies caused by extreme heat can manifest as both acute physiological responses and exacerbations of chronic diseases. According to evaluations conducted during the workshop, emergencies that may be encountered in extreme heat can be listed under the following subheadings: heatstroke (sunstroke); heart attack, stroke, hypertensive attack, hypoglycemia, gastroenteritis, and first- and second-degree burns; and asthma attack and exacerbation of chronic obstructive pulmonary disease (COPD).

In this context, symptoms of heatstroke can be listed as: weakness, exhaustion, fatigue, headache, dizziness, increased body temperature (≥38 °C), tachycardia, excessive sweating or inability to sweat, muscle cramps, difficulty breathing, confusion or loss of consciousness.

Recommendations:

1. First-aid steps include moving the person to a cool, shaded area, loosening or removing clothing, providing a cool shower or cooling the body with a damp towel or cloth, and placing the person on their back with their arms and legs elevated.

2. If the person is conscious and has no nausea, give them either a solution of one teaspoon of baking soda and one teaspoon of salt in one liter of water or soda. If they are unconscious, place them in the recovery (coma) position and call 112. In case of respiratory arrest, perform cardiopulmonary resuscitation and, if available, use an automated external defibrillator.

3. For those with existing chronic conditions, such as hypertension or hypoglycemia, provide first aid according to the appropriate protocol and call 112 immediately.

Vulnerable groups more susceptible to heatstroke and related complications include patients with chronic conditions (e.g., heart failure, vascular diseases, hypertension, diabetes, cancer, kidney diseases, cystic fibrosis, asthma, COPD); pregnant women; individuals aged 65 and over; obese or asthenic individuals; individuals with insufficient fluid intake; those on diets or with nutritional deficiencies; migrant workers (internal and external migrants); undocumented agricultural workers; and child laborers.

Preventive and protective measures in emergency situations

a. Engineering Measures

Recommendation:

Installing local or general ventilation systems (HVAC), insulating heat sources, using radiation barriers to reduce radiant heat, implementing shading solutions (white/reflective roofs,

awnings, mobile awnings), and installing adiabatic cooling and misting systems.

b. Administrative Measures

Recommendations:

1. Work-rest periods should be regulated (e.g., 15 minutes of rest in the shade after 45 minutes of work).

2. Implementing shift rotation by scheduling work during cooler hours is another option.

3. Implementing a heat acclimatization program that gradually increases working hours over 10–14 days can help new employees.

4. Providing fluid and electrolyte support (ensuring consumption of 200–250 mL of water or ayran every 15–20 minutes) is of great importance.

5. Heat-stress training should be organized.

6. Implementing a "buddy system" among employees is important for early recognition of heat-related illness.

7. WBGT should be measured regularly, with the date, time, and environmental conditions recorded appropriately.

8. The number of first aid certified personnel be increased in accordance with regulations.

9. In line with the Presidential Directive dated 27.04.2024, unregistered agricultural workers should be registered through the E-Metip system. Supervision and monitoring should be provided by the governorships and the Ministry of Agriculture. The working hours of high-risk occupational groups should be regulated during periods of extreme heat.

c. PPE

While PPE plays an important role in protecting workers in extreme heat conditions, it is the last step in the prevention hierarchy and cannot replace engineering and administrative measures. Clothing selection is of great importance at this point: whereas normal work clothes do not affect the heat correction factor, thin cotton smocks can increase body temperature by approximately 2 °C, and vapor-barrier clothing can cause an increase of up to 11 °C.

Recommendations:

1. Cooling vests and vapor-permeable lightweight protectors should be preferred.

2. Clothing made of light-colored, sweat-absorbing, ultraviolet-protective fabrics should be used.

3. Wide-brimmed hats and neck protectors are also effective for sun protection.

d. Technical and Innovative Approaches

Since it is often impossible to completely eliminate the sources of danger in industrial processes, collective protection measures should be prioritized, especially in high-temperature enclosed systems such as metal-melting furnaces or in open-field work, such as agriculture and construction.

Recommendations:

1. Shade, automation, and remote-control systems can be considered in this context.

PPE applications should be continued to support these collective measures. Studies have shown that cooling vests reduce body temperature by up to 1.3 °C in employees, and water and sodium supplementation is effective against heatstroke.[9,10]

2. Building on conventional monitoring and preventive strategies as described above, advanced technological solutions may further enhance heat stress management. Integrating WBGT measurements with thermal imaging systems and artificial intelligence–supported algorithms enables real-time risk prediction based on workers' body temperature and environmental conditions, allowing dynamic adjustment of rest periods.

Fluid intake can be monitored through smart water dispensers, enabling automated warnings for insufficient hydration.

3. In workplaces with a high risk of extreme heat exposure, it is essential to have technological cooling devices and emergency response equipment readily available and to ensure access to first aid tools and communication systems.

Since each workplace involves different risks depending on the type of activity, environmental conditions, and job descriptions, it is crucial to prepare and regularly update customized emergency plans.

DISCUSSION

The primary outcome of this workshop is a clear, consensus-based framework which highlights the need for urgent adaptation of current OHS mechanisms in Türkiye to manage the escalating threat of extreme heat. The key messages derived from the working groups emphasize four critical action areas: 1) the explicit integration of extreme weather risks into primary legislation (OHS Law No. 6331); 2) the establishment of secondary regulations mandating WBGT-driven work-rest schedules and acclimatization protocols; 3) the prioritization of structural engineering and administrative controls over PPE; and 4) the implementation of targeted protection mechanisms for informal, seasonal, and migrant workers.

This workshop report demonstrates that extreme heat is no longer an "exceptional weather event" but a permanent, predictable occupational hazard for working life. The WHO European Regional Office reports that between 2000 and 2019, more than 175,000 deaths each year in the European Region were related to extreme heat, which accounts for approximately one-third of global heat-related deaths.[1] This shows that the effects of climate change on health are no longer marginal, but rather a fundamental public health issue. Joint WHO and World Meteorological Organization reports published in recent years highlight that approximately 2.4 billion workers, or more

than 70% of the global workforce, are exposed to heat stress at work; and that occupational heat-related incidents occur every year, resulting in millions of injuries and thousands of deaths.[11] These findings reveal that the TTS's initiative to address working conditions during extreme heat as a distinct policy area is highly appropriate. The ILO report entitled "Working on a Warmer Planet" predicts that by 2030, more than 2% of total global working hours will be lost solely due to temperature.[12] This loss will be concentrated in labor-intensive sectors such as agriculture, construction, and other outdoor work. Considering both Türkiye's climate projections and employment structure, it is clear that similar losses are likely in our country as well.

Studies in industrialized countries show that the risk of traumatic occupational accidents increases significantly with rising temperatures. In their systematic review, Spector et al.[13] revealed that the common finding of numerous studies conducted in different sectors is that working in hot environments increases the incidence of events such as falls, loss of equipment control, and vehicle accidents. Park et al.[14], who analyzed 11.6 million worker compensation files in California between 2001 and 2018, showed that the risk of occupational accidents increased by approximately 10–15% on days when the daily maximum temperature rose to the 32–37 °C range, and this rate increased even further on extremely hot days. Reports from the California Environmental Health Agency also show that occupational heat-related illness and death rates doubled between 2000 and 2017.[15] These findings support the observation, as emphasized during the workshop, that "heat stress is a risk factor that increases occupational accidents not only in outdoor work but also in indoor environments". The lack of a national coding system for monitoring heat-related occupational injuries in Türkiye reduces the visibility of the problem; however, it does not imply that the risk is absent. Systematic studies of heat stress-related accidents are necessary, especially in sectors with long working hours, low wages, and widespread informal employment.

The literature has increasingly focused on heat stress, which predisposes not only to acute events but also to long-term consequences such as chronic kidney disease. Pioneering studies on sugarcane workers in El Salvador found a significant deterioration in end-of-shift kidney function, associated with intense physical workload, high ambient temperature, and inadequate fluid intake; this condition was linked to a condition called "Mesoamerican nephropathy".[16] Wesseling et al.[17], in their studies involving different occupational groups, reported a strong association between high heat, dehydration, and increased serum uric acid levels and kidney dysfunction in workers in fields such as agriculture, ports, and construction. In a more recent review, Elinder[18] emphasized that heat-related kidney damage is not limited to Central America; similar clinical patterns have been reported in groups performing heavy physical work in hot climates, and this should be considered a new type of "occupational nephropathy". These data represent a potential warning for workers in Türkiye, particularly those in the agriculture, construction, mining, and logistics sectors. Since current databases do not specifically track heat-induced renal dysfunction, this area should be a priority for future epidemiological studies.

Analyses by the ILO show that heat-related work hours lost in the agricultural and construction sectors in low- and middle-income countries have significant impacts on economic productivity.[12,19] These losses are important not only macroeconomic indicators but also income inequality and social justice; the highest losses are seen in those already working in precarious and low-wage jobs. Review studies on Mesoamerican nephropathy reveal that kidney damage is most common in low-wage, non-unionized, and socially unprotected male workers; the disease is a "climate justice" issue arising at the intersection of climate change and social inequality.[18] These findings explain why the dimensions of "informal labor, migrant labor, and child labor," which were frequently emphasized at the workshop, should be central.

In Türkiye, a large proportion of agricultural workers are seasonal or migrant and work informally. The steps proposed in the workshop, such as registration via E-Metip, OHS training in the native language, and integration with the family physician system, are consistent with the "temperature-sensitive social protection systems" approach advocated by the ILO and WHO.[11,12,19] In our country, there is no specific law or regulation regarding working conditions in extreme heat. However, the climate crisis, with increasing temperatures and extreme weather events necessitates legal measures regarding OHS. In countries with higher temperatures, such as the United Arab Emirates and Qatar, legislation includes quite comprehensive

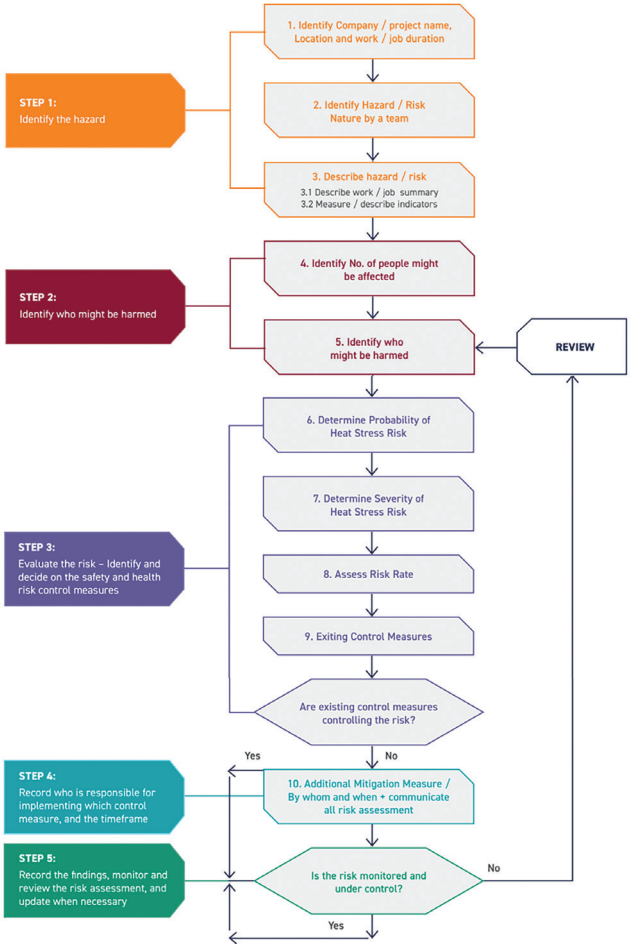

Figure 1. Risk assessment flowchart used in Qatar[20]

(Reproduced from ILO publication with permission)

regulations regarding working conditions in terms of OHS. Figure 1 illustrates the risk assessment flowchart used in Qatar.[20] A risk assessment flowchart can also be prepared in our country, taking regional conditions into account.

Similarly, in Dubai, the temperature index is calculated by taking the humidity level into account, and a risk assessment is carried out accordingly (Figure 2).[21]

In Qatar, work-rest schedules based on WBGT measurement results, work intensity and whether the employee is acclimatized, (starting with short working hours for new employees or those returning from a long break and gradually increasing the duration-approximately 15 days) are shown in Figure 3.[20]

A study of 307 workers in Australia[22] highlighted the association between rising temperatures and several types of injuries, including lifting injuries, hand injuries, and loss of control of electrical tools. Climate change causes far more workplace accidents than official records indicate, and research shows that the effects of climate change are concentrated among workers in open areas and among low-income workers living in housing without air conditioning. Studies show that heat impairs concentration, increases workplace accidents, and leads to an additional 20,000 such accidents every year in California alone.[14]

Heat islands in cities also contribute to extreme temperatures. In our country, unplanned urbanization exacerbates this effect. Heat island formation can increase temperatures by up to 10 °C at night.

According to ILO data, agricultural and construction workers, numbering 940 million, are the most affected groups, with particularly elevated risks among informal workers. Risk has also increased among workers in environmental services, fishing, waste collection, cargo distribution, tourism, sports, road maintenance and repair, and power-line workers, firefighters, search-and-rescue and emergency-response workers, those employed in mining and manufacturing, and individuals engaged in military activities.

Some countries have begun to recognize heat stress as an occupational risk and to explicitly reflect it in their legislation. In Qatar, a 2021 Ministerial Decree required a risk assessment based on WBGT and stipulated that work be stopped when WBGT exceeds 32.1 °C.[20] In addition, an outdoor work ban was imposed between 10:00 and 15:30 from June 1 to

Relative Humidity

Temperature, °C	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%
42°	48	50	52	55	57	59	62	64	66	68	71	73	75	77	80	82
41°	46	48	51	53	55	57	59	61	64	66	68	70	72	74	76	79
40°	45	47	49	51	53	55	57	59	61	63	65	67	69	71	73	75
39°	43	45	47	49	51	53	55	57	59	61	63	65	66	68	70	72
38°	42	44	45	47	49	51	53	55	56	59	60	62	64	66	67	69
37°	40	42	44	45	47	49	51	52	54	56	58	59	61	63	65	66
36°	39	40	42	44	45	47	49	50	52	54	55	57	59	60	62	63
35°	37	39	40	42	44	45	47	48	50	51	53	54	56	58	59	61
34°	36	37	39	40	42	43	45	46	48	49	51	52	54	55	57	58
33°	34	36	37	39	40	41	43	44	46	47	48	50	51	53	54	55
32°	33	34	36	37	38	40	41	42	44	45	46	48	49	50	52	53
31°	32	33	34	35	37	38	39	40	42	43	44	45	47	48	49	50
30°	30	32	33	34	35	36	37	39	40	41	42	43	45	46	47	48
29°	29	30	31	32	33	35	36	37	38	39	40	41	42	43	45	46
28°	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
27°	27	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41
26°	26	26	27	28	29	30	31	32	33	34	34	35	36	37	38	39
25°	25	25	26	27	27	28	29	30	31	32	33	34	34	35	35	37
24°	24	24	24	25	26	27	28	28	29	30	31	32	33	33	34	35
23°	23	23	23	24	25	25	26	27	28	28	29	30	31	32	32	33
22°	22	22	22	22	23	24	25	25	26	27	27	28	29	30	30	31

Legend:

Up to 29°C	No discomfort
From 30 to 34°C	Slight discomfort sensation
From 35 to 39°C	Strong discomfort. Caution: limit the heaviest physical activities
From 40 to 45°C	Strong indisposition sensation. Danger: Avoid efforts
From 46 to 53°C	Serious danger: stop all physical activities
Over 54°C	Death danger: imminent heat stroke

Figure 2. Temperature index calculated, taking humidity into account, reproduced with permission from Technical Guideline DM-HSD-GU38-MHSW2, Safety Section, Health and Safety Department, Dubai Municipality, Dubai, United Arab Emirates

(Reproduced from ILO publication with permission)

Minutes of break within an hour of work shift	WBGT for type of work load (°C)			
	Light work	Moderate-intensity work	High-intensity work	Very high-intensity work
Acclimatized workers				
0 – 15	≤ 30.8	≤ 28.2	---	---
15 – 30	30.9 - 31.2	28.3 - 29.0	26.7 - 27.6	---
30 – 45	31.3 - 31.8	29.1 - 30.1	27.7 - 28.8	26.6 - 27.9
45 - 60	> 31.8	> 30.1	> 28.8	> 27.9
Non-acclimatized workers				
0 - 15	≤ 28.1	≤ 25.0	---	---
15 - 30	28.2 - 28.7	25.1 - 26.0	23.1 - 24.2	---
30 - 45	28.8 - 29.3	26.1 - 27.2	24.3 - 25.7	23.0 - 24.6
45 - 60	> 29.3	> 27.2	> 25.7	> 24.6

--- : No Threshold Limit Values provided. Detailed analysis and/or physiological monitoring should be used.

Figure 3. Work break recommendations according to WBGT[20]

(Reproduced from ILO publication with permission)

WBGT: Wet-Bulb Globe Temperature

September 15 each year, and rest areas and cooling stations were mandated. Similarly, in the United Arab Emirates, a "lunch break work ban" is implemented between 12:30 and 15:00 during the summer, and the compliance rate is reported to reach 99%.[23]

These examples show that the recommendations in the workshop report, such as "not working between 10:30 and 15:30 for vulnerable groups at 32 °C and above" and "incorporating WBGT-based work-rest schedules into legislation," are not merely theoretical demands but are consistent with good practices already implemented in some countries.

When the reviewed literature and the workshop findings are evaluated together, the following conclusions can be drawn:

- Extreme heat is one of the most important climate-related causes of death and workforce loss, according to WHO and ILO data.
- Heat stress increases both acute traumatic injuries and long-term health problems such as chronic kidney disease.
- The highest risk is concentrated in sectors with open spaces and heavy workloads, such as agriculture, construction, mining, logistics, and environmental services, and among informal and migrant workers.
- Examples from some countries show that work bans based on WBGT protocols, acclimatization protocols, and mandatory rest and cooling areas are feasible in practice.

Therefore, in Türkiye, as suggested in the workshop report:

- Adding a clear provision specific to extreme weather conditions to Law No. 6331,
- Arranging work-rest periods in accordance with temperature and WBGT thresholds in the Working Hours Regulation related to the Labour Law,
- Establishing special protection and monitoring mechanisms for unregistered and migrant workers,
- All these recommendations should be regarded as scientifically-based policy steps consistent with national health data and the international literature.

Study Limitations

While this workshop report provides a comprehensive and interdisciplinary framework, several limitations should be noted. First, participant selection was based on purposive sampling. Although it aimed for broad representation across the health, labor, and legal sectors, representatives of some specific industrial sectors or local employer associations may have been underrepresented. Second, the consensus process relied on open qualitative discussions rather than a formal, mathematically structured consensus method (such as the Delphi technique). While the final recommendations reflect a strong agreement among attendees, the outputs were shaped by the dynamics of open group discussions. The proposed WBGT thresholds and work-rest schedules are currently based on international standards (e.g., ISO 7243) and on practices in other countries. Future empirical studies are needed to validate these specific thresholds within Türkiye's local climatic and demographic context.

CONCLUSION

The climate crisis is one of the most important issues threatening our health worldwide. Besides measures to prevent the crisis from deepening, adaptation and mitigation are crucial. Workers' health is threatened by the climate crisis, rising temperatures, and extreme weather events. To address this issue, our country needs legal regulations, training for workers and employers, and structural adjustments. In the workshop we held, which included representatives from all relevant fields of expertise, we discussed the issue in all its aspects and developed recommendations. We believe that the topics discussed and

the recommendations made at our workshop will guide future steps in this field within our country.

Acknowledgments

We thank all invited speakers, institutional representatives, and participants who contributed to the workshop. Specific gratitude is extended to: Dr. Duygu Acar, Michail Kandarakis, Dr. Mümine Yüksel, Prof. Dr. Ahmet Uğur Demir, Assoc. Prof. Dr. Cavit Işık Yavuz, Mr. Kadir Kadiroğulları, Dr. Buket Gülhan, Hüseyin Kekeç, Hulusi Türkyılmaz, Esat Ayçiçek, Erdem Mızrak, Barış İyiaydın, Nuran Gülenç, and Assist. Prof. Dr. Canan Demir. We also acknowledge the organizing team: Assoc. Prof. Dr. Haluk Çalışır, Assist. Prof. Dr. Çağla Hüseynova, Assist. Prof. Dr. Merve Erçelik Koncak, Dr. Damla Karadeniz, Assoc. Prof. Dr. Nilüfer Aykaç, and Prof. Dr. Ali Kocabaş.

Footnotes

Authorship Contributions

Concept: S.G., S.Ç.K., S.S.O., Design: S.G., S.S.O., Data Collection or Processing: S.G., S.Ç.K., S.S.O., M.K., İ.C.D.Ü., Analysis or Interpretation: S.G., S.Ç.K., S.S.O., M.K., İ.C.D.Ü., Literature Search: S.G., S.Ç.K., M.K., Writing: S.G., S.Ç.K., S.S.O.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

Appendix 1: https://d2v96fxpocvxx.cloudfront.net/bd1986e1-0bc1-4f4d-af66-3a184850a065/content-images/7335677d-6d13-4975-b682-c7f7c04cf9d9.pdf

REFERENCES

1. World Health Organization Regional Office for Europe. Statement – Heat claims more than 175 000 lives annually in the WHO European Region, with numbers set to soar. WHO; 1 Aug 2024. Accessed July 3, 2026. [Crossref]
2. World Health Organization. Climate change and heat: fact sheet. WHO; 2023. Accessed July 3, 2026. [Crossref]
3. Masselot P, Mistry MN, Rao S, et al. Estimating future heat-related and cold-related mortality under climate change, demographic and adaptation scenarios in 854 European cities. *Nat Med*. 2025;31(4):1294-1302. [Crossref]
4. General Directorate of Water Management (SYGM). Reports from the Project on the Impact of Climate Change on Water Resources. Republic of Turkey Ministry of Forestry and Water Affairs, Ankara; 2016. [Crossref]
5. Ercan A. Assessment of the occupational health and safety in construction sector in Turkey. *J Polytech*. 2010;13(1):49-53. [Crossref]
6. Carter S, Field E, Oppermann E, Brearley M. The impact of perceived heat stress symptoms on work-related tasks and social factors: a cross-sectional survey of Australia's Monsoonal North. *Appl Ergon*. 2020;82:102918. [Crossref]
7. Babalık FC. Ergonomics for Engineers. Bursa: Dora Publications. 2016. [Crossref]
8. Turkish Standards Institution (TSE). TS EN ISO 7243: Ergonomics of the thermal environment-Assessment of heat stress using the WBGT (wet bulb globe temperature) index (ISO 7243:2017). Ankara: Turkish Standards Institution; 2017. Accessed July 3, 2026. [Crossref]
9. Nerbass FB, Pecoits-Filho R, Clark WF, et al. Occupational heat stress and kidney health: from farms to factories. *Kidney Int Rep*. 2017;2(6):998-1008. [Crossref]
10. Üner MH, Yılmaz İ. Effects of heat stress and fluid intake on occupational health and safety. *Ohs Academy*. 2020;3:129-134. [Crossref]
11. World Health Organization, World Meteorological Organization. Climate change and workplace heat stress: technical report and guidance. Geneva: World Health Organization and World Meteorological Organization; 2025. Accessed July 20, 2026. [Crossref]
12. International Labour Office. Working on a warmer planet: the impact of heat stress on labour productivity and decent work [Internet]. Geneva: International Labour Office; 2019. Accessed July 20, 2026. [Crossref]
13. Spector JT, Masuda YJ, Wolff NH, Calkins M, Seixas N. Heat exposure and occupational injuries: review of the literature and implications. *Curr Environ Health Rep*. 2019;6(4):286-296. [Crossref]
14. Park RJ, Pankratz N, Behrer AP. Temperature, workplace safety, and labor market inequality [Internet]. Bonn: IZA Institute of Labor Economics; 2021. IZA Discussion Paper No. 14560. Accessed July 20, 2026. [Crossref]
15. Office of Environmental Health Hazard Assessment. Occupational heat-related illness. In: Indicators of Climate Change in California. 4th ed. Sacramento, CA: California Environmental Protection Agency; 2022. Accessed July 20, 2026. [Crossref]
16. García-Trabanino R, Jarquín E, Wesseling C, et al. Heat stress, dehydration, and kidney function in sugarcane cutters in El Salvador--A cross-shift study of workers at risk of Mesoamerican nephropathy. *Environ Res*. 2015:142:746-55. [Crossref]
17. Wesseling C, Aragón A, González M, et al. Heat stress, hydration and uric acid: a cross-sectional study in workers of three occupations in a hotspot of Mesoamerican nephropathy in Nicaragua. *BMJ Open*. 2016;6(12):e011034. [Crossref]
18. Elinder CG. Heat-induced kidney disease: understanding the impact. *J Intern Med*. 2025;297(1):101-112. [Crossref]
19. Heat at work: Implications forsafety and health. International Labour Organization. Accessed July 20, 2026. [Crossref]
20. International Labour Organization. Heat stress legislation in Qatar: a guide for employers [Internet]. Geneva: International Labour Organization; 2021. Accessed July 20, 2026. [Crossref]
21. Technical Guidelines forManagement of Heat Stress at Work. Dubai Municipality Health and Safety Department. Accessed July 20, 2026. [Crossref]
22. Varghese BM, Hansen AL, Williams S, et al. Determinants of heat-related injuries in Australian workplaces: Perceptions of health and safety professionals. *Sci Total Environ*. 2020;718:137138. [Crossref]
23. United Arab Emirates Ministry of Human Resources and Emiratisation. Heat-stress legislation and the Midday Break Policy for the protection of outdoor workers. Accessed July 20, 2026. [Crossref]

Thorac Res Pract. 2026;27(4):261-263 **DOI:** 10.4274/ThoracResPract.2026.2026-2-6

Letter to the Editor

Secondary Pulmonary Alveolar Proteinosis Following Brief Toxic Gas Inhalation: A Case Report

Ali Usman Arif[1], Mehwish Aslam[1], Moiz Salahuddin[1], Muhammad Irfan[1], Muhammad Usman Ghani[1], Akbar Shoukat Ali[1], Hania Naveed[2], Tehreem Khan[1]

[1]Section of Pulmonary and Critical Care, Department of Medicine, Aga Khan University Hospital, Karachi, Pakistan
[2]Department of Pathology and Laboratory Medicine, Aga Khan University Hospital, Karachi, Pakistan

Cite this article as: Arif AU, Aslam M, Salahuddin M, et al. Secondary pulmonary alveolar proteinosis following brief toxic gas inhalation: a case report. *Thorac Res Pract.* 2026;27(4):261-263

KEYWORDS

Pulmonary alveolar proteinosis, pulmonary diseases, toxic gas inhalation

Received: 19.02.2026 ***Revision Requested:*** 12.03.2026 ***Last Revision Received:*** 02.04.2026

Accepted: 16.04.2026 ***Epub:*** 05.06.2026 ***Publication Date:*** 24.07.2026

DEAR EDITOR,

Pulmonary alveolar proteinosis (PAP) is a syndrome characterized by impaired clearance of surfactant, leading to its accumulation within the alveoli and subsequent restrictive lung disease.[1-3] We present an instructive case of secondary PAP triggered by brief inhalation of a toxic gas to underscore the diagnostic importance of a detailed exposure history.

A 30-year-old Afghan male, an occasional smoker (2–3 pack-years) with no significant past medical history, presented with a one-year history of progressive dry cough and exertional dyspnoea (modified Medical Research Council grade 4). He reported a brief, 15-minute exposure to an unidentified irritant gas in a confined space at symptom onset. The exposure occurred in a domestic setting, while the patient was cleaning with a commercial cleaning agent in a poorly ventilated bathroom. The specific chemical composition of the agent was not identifiable; however, the acute onset of respiratory symptoms within minutes suggested a highly irritating gas, which could be chlorine- or ammonia-based. The patient had no prior history of respiratory or cardiac disease. He was not taking any regular medications and had no known allergies. An echocardiogram performed during the initial workup showed normal left ventricular systolic function and no evidence of pulmonary hypertension. Prior to the toxic gas exposure, he had no respiratory symptoms and had never undergone chest imaging. On admission, he was hypoxemic (SpO_2 86% on room air), requiring high-flow nasal oxygen (40 L/min, FiO_2 0.5). Auscultation revealed bilateral fine inspiratory crackles. The chest radiograph demonstrated diffuse bilateral alveolar opacities (Figure 1). A comprehensive etiological evaluation was performed to exclude other causes of secondary PAP. Complete blood count and peripheral smear showed no evidence of hematological malignancy. Serum human immunodeficiency virus serology and tuberculosis screening (sputum microscopy and GeneXpert) were negative. The patient had no history of occupational exposure to silica, metal dusts, or other inorganic particulates. Immunoglobulin levels and lymphocyte subset analysis were within normal limits, ruling out an underlying immunodeficiency. Serum anti-granulocyte-macrophage colony-stimulating factor antibodies were tested and found to be negative, ruling out autoimmune PAP. High-resolution computed tomography (HRCT) showed extensive bilateral ground-glass opacities with smooth interlobular septal thickening, the classic "crazy-paving" pattern of PAP (Figure 2). Bronchoscopy yielded an opaque, milky bronchoalveolar lavage (BAL) fluid (Figure 3). Cytology showed abundant acellular granular material that was strongly periodic acid-Schiff (PAS)-positive and diastase-resistant. Transbronchial biopsy confirmed alveolar filling with PAS-positive lipoproteinaceous material, establishing the diagnosis of PAP (Figure 4). He underwent sequential therapeutic

Corresponding author: Muhammad Irfan, MD, e-mail: muhammad.irfan@aku.edu

whole lung lavage (WLL) under general anaesthesia. After approximately 25 litres of warmed saline were instilled into both lungs, oxygenation improved markedly, and he was weaned from high-flow oxygen within 48 hours. A follow-up chest radiograph at two weeks showed significant resolution (Figure 5). At three months, he was asymptomatic on room air.

This case exemplifies secondary PAP arising from an acute, non-occupational inhalational injury. Inhalational triggers of secondary PAP are diverse and include inorganic dusts such as silica, aluminum, titanium, and indium tin oxide, as well as acutely irritating gases like chlorine and nitrogen dioxide.[2,3]

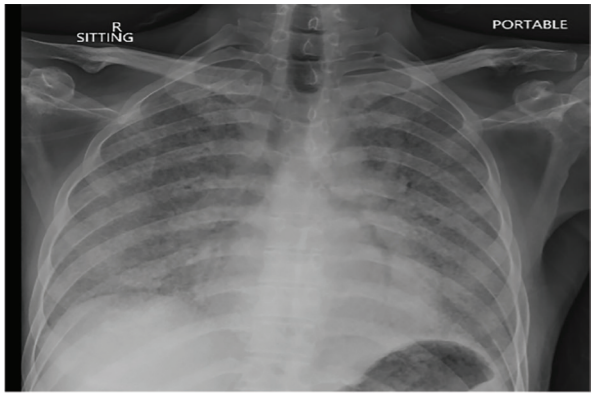

Figure 1. Initial anterior-posterior chest radiograph on presentation showing bilateral, diffuse, symmetrical alveolar opacities

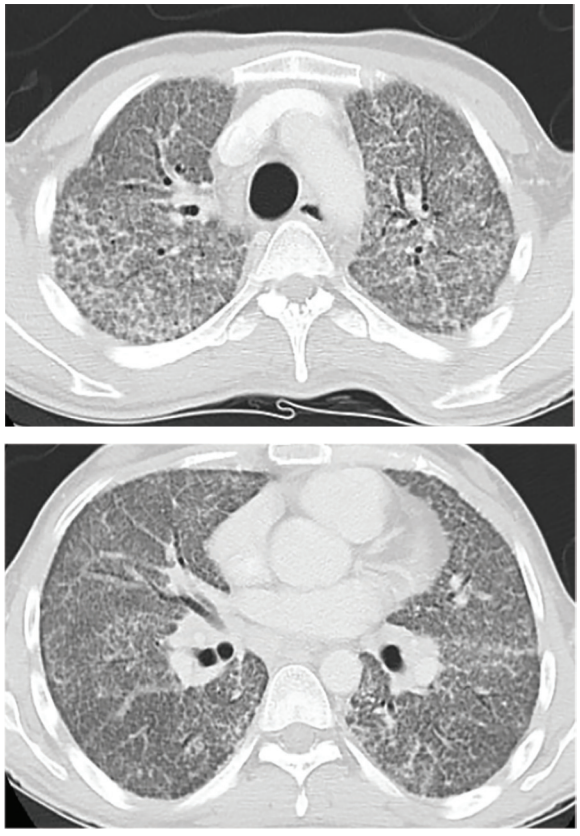

Figure 2. Axial high-resolution computed tomography scan of the thorax demonstrating the classic "crazy paving" pattern, characterized by extensive ground glass opacification with superimposed smooth interlobular septal thickening pattern indicative of PAP

PAP: Pulmonary alveolar proteinosis

The underlying mechanisms vary by agent, but commonly involve direct alveolar epithelial injury, impaired clearance, or a combination of these mechanisms. Our patient's clear temporal link with exposure to a toxic gas strongly supports this mechanism. HRCT is pivotal; the "crazy-paving" pattern, while not pathognomonic, is a highly characteristic sign that should prompt immediate suspicion of PAP.[4] In such a context, BAL cytology demonstrating PAS-positive, lipid-laden acellular debris is often sufficient for diagnosis, as in our patient.[3] Histopathological confirmation remains the diagnostic gold standard. For symptomatic, hypoxemic patients, WLL is the established first-line therapy. It is a highly effective mechanical method of removing the accumulated alveolar material, leading to sustained physiological and clinical improvements in the majority of patients, as we observed.[5] The prognosis of secondary PAP is intrinsically linked to the underlying cause, and secondary cases can follow a more variable course than autoimmune PAP.[1] A limitation of this report is the absence of post-treatment HRCT imaging; however, the marked clinical improvement and chest radiograph resolution support the efficacy of WLL.

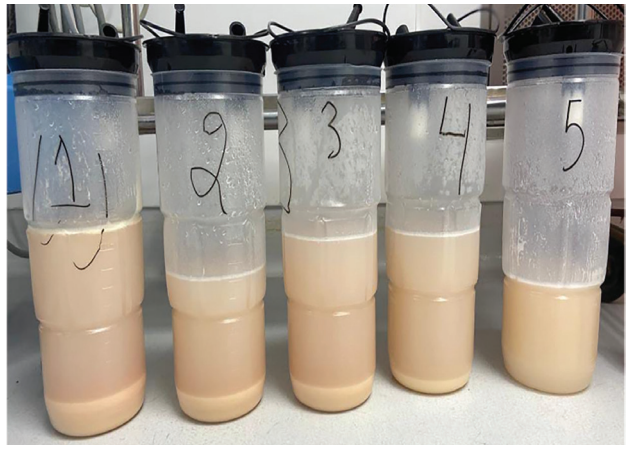

Figure 3. Fiberoptic bronchoscopy showing opaque and milky bronchoalveolar lavage fluid characteristic of PAP

PAP: Pulmonary alveolar proteinosis

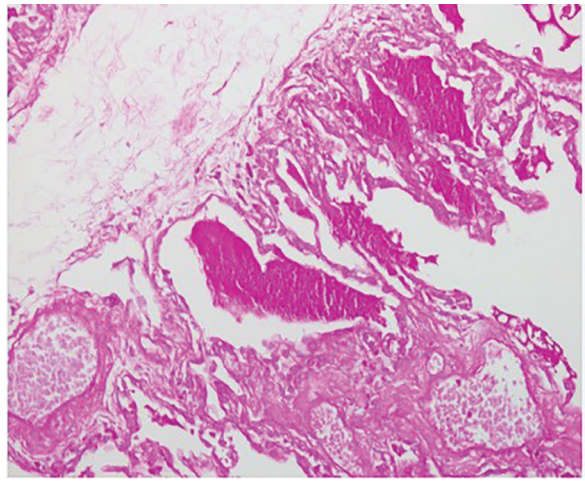

Figure 4. Histopathology of transbronchial lung biopsy [periodic acid-Schiff (PAS) stain, original magnification ×40]. Alveolar spaces are filled with granular, PAS-positive lipoproteinaceous material, confirming the diagnosis of pulmonary alveolar proteinosis. Alveolar architecture is preserved

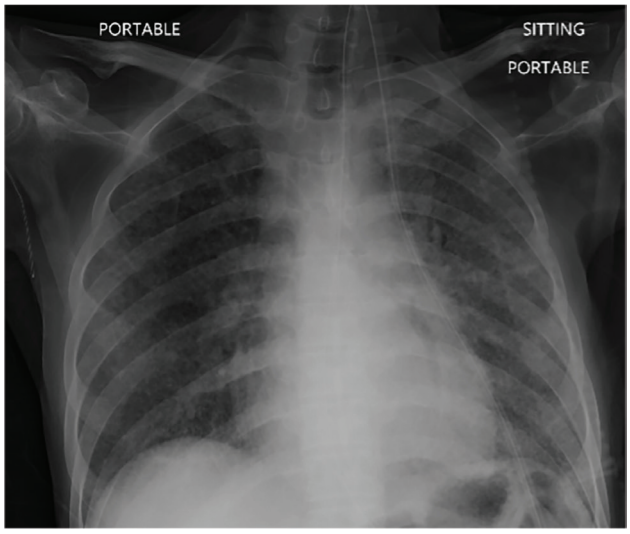

Figure 5. Follow-up chest radiograph obtained two weeks after sequential whole lung lavage, demonstrating marked resolution of the previously noted bilateral alveolar opacities

Acknowledgements

We thank the patient for consenting to share their medical case and for contributing to the advancement of medical knowledge through this publication.

Footnotes

Informed Consent: Informed consent was obtained from the patient for anonymized information to be published in this article.

Declaration on the use of Artificial Intelligence (AI):

The authors declared that OpenAI's ChatGPT was used only for language editing, grammatical correction, and sentence refinement. The content, data analysis, interpretation, and conclusions remain the sole responsibility of the authors.

Authorship Contributions

Concept: A.U.A., M.A., M.S., M.I., M.U.G., H.N., T.K., Design: A.U.A., M.A., M.S., M.I., M.U.G., H.N., T.K., Data Collection or Processing: A.U.A., M.A., M.S., M.I., A.S.A., H.N., T.K., Analysis or Interpretation: A.U.A., M.A., M.S., M.I., M.U.G., H.N., T.K., Literature Search: A.U.A., M.A., M.S., M.I., M.U.G., A.S.A., H.N., T.K., Writing: A.U.A., M.A., M.S., M.I., A.S.A., H.N., T.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Trapnell BC, Nakata K, Bonella F, et al. Pulmonary alveolar proteinosis. *Nat Rev Dis Primers*. 2019;5(1):16. **[Crossref]**
2. Borie R, Danel C, Debray MP, et al. Pulmonary alveolar proteinosis. *Eur Respir Rev*. 2011;20(120):98-107. **[Crossref]**
3. Bhatnagar M, Slight R, Prasad AB, Stanton AE. Prolonged air leak after IPC insertion: an unusual complication. *Respir Med Case Rep*. 2020;31:101257. **[Crossref]**
4. Frazier AA, Franks TJ, Cooke EO, Mohammed TL, Pugatch RD, Galvin JR. From the archives of the AFIP: pulmonary alveolar proteinosis. *Radiographics*. 2008;28(3):883-899; quiz 915. **[Crossref]**
5. Campo I, Mariani F, Rodi G, et al. Assessment and management of pulmonary alveolar proteinosis in a reference center. *Orphanet J Rare Dis*. 2013;8:40. **[Crossref]**

Thorac Res Pract. 2026;27(4):264-265 **DOI:** 10.4274/ThoracResPract.2026.2026-6-3

Letter to the Editor

Letter to the Editor: Free Large Language Model-based Chatbots Can Help to Align Statistical Tests with the Study Design and Avoid Pseudo-replication

Konradin Metze[1], Cauê Sousa da Silva[1], Irene Lorand-Metze[2], Amilcar Castro de Mattos[1,3]

[1]Department of Pathology, State University of Campinas Faculty of Medical Sciences, São Paulo, Brazil
[2]Department of Internal Medicine, State University of Campinas Faculty of Medical Sciences, São Paulo, Brazil
[3]Laboratory of Pathology, Pontifical Catholic University of Campinas, Campinas, São Paulo, Brazil

Cite this article as: Metze K, da Silva CS, Lorand-Metze I, de Mattos AC. Letter to the Editor: Free large language model-based chatbots can help to align statistical tests with the study design and avoid pseudo-replication. *Thorac Res Pract.* 2026;27(4):264-265

KEYWORDS

Clinical problems, eHealth, artificial intelligence, biostatistics, large language models

Received: 07.06.2026

Accepted: 19.06.2026 ***Publication Date:*** 24.07.2026

DEAR EDITOR,

In a recent contribution,[1] the authors compared the quality of replies provided by different large language-based chatbots (LLMs) to questions related to diagnosis, treatment, and prognosis of pulmonary thromboembolism. The study design was based on 10 questions, all submitted to four different chatbots and their replies were scored by specialists. This study is composed of blocks of 4 dependent answers always related to one question. Therefore, the Kruskal-Wallis test, which has been created for independent data, is not adequate. When incorrectly applied for nested data, this test will provoke pseudo-replication.[2] The test power will be distorted, thus provoking unreliable results. Adequate solutions would be tests for repeated measures, such as mixed models, Friedman tests with post-hoc tests and alpha-error corrections. Eventually, ANOVA tests for dependent data, when assumptions were met.

Here we tried to find out whether freely available LLM-based chatbots could help to select the adequate statistical tests. The published abstract was sent to 14 different chatbots with the following prompt:

“Please analyze this text and look for major problems, such as methodological errors, inconsistencies, or contradictions. Pay special attention to the statistical tests and examine whether they are well aligned with the study design. Please suggest corrections. Should we change a test by another one?“

The following 12 algorithms recognized the nested study design and suggested substitution by the above-mentioned adequate test. appointing it as the most important error: ChatGPT Free, Claude Sonnet 4.6, Consensus Corpus All, Copilot Smart, Deep Seek Fast DeepThink, Gemini 3.5 Flash, Grok Fast, Julius 1.2 Lite, Meta AI Instant, Notebook LM Free, Perplexity Learn Tutor. Furthermore, other questions regarding test power, data distribution, counting procedures, post-hoc tests, alpha-error correction, and inter-rater variability were also correctly discussed.

The remaining two bots (Bohrium Expert and Mistral Le Chat Balanced) did not recognize the nested study design and therefore did not indicate correction of the most important error.

Corresponding author: Konradin Metze, MD, e-mail: kmetze@fcm.unicamp.br

Since replies of different LLMs to the same prompt or repeated submissions may lead to diverging results,[3-5] we always recommend to consult several LLMs and eventually, to discuss conflicting answers with the chatbots until a final solution can be reached.[6]

Footnotes

Authorship Contributions

Concept: K.M., A.C.D.M., Design: K.M., C.S.D.S., Data Collection or Processing: K.M., C.S.D.S., A.C.D.M., Literature Search: K.M., C.S.D.S., I.L-M., A.C.D.M., Analysis or Interpretation: K.M., C.S.D.S., Writing: K.M., I.L-M., A.C.D.M.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Karakoyun ÖF, Koyuncuoğlu HE, Sağnıç ÖH, Özdemir ME, Gölcük Y, Yıldırım B. AI in patient care: evaluating large language model performance against evidence-based guidelines for pulmonary embolism. *Thorac Res Pract*. 2026;27(1):38-46. **[Crossref]**
2. Lazic SE. The problem of pseudoreplication in neuroscientific studies: is it affecting your analysis? *BMC Neurosci*. 2010;11:5. **[Crossref]**
3. Metze K, Morandin-Reis RC, Lorand-Metze I, Florindo JB. Bibliographic research with ChatGPT may be misleading: the problem of hallucination. *J Pediatr Surg*. 2024;59(1):158. **[Crossref]**
4. Metze K, Morandin-Reis RC, Lorand-Metze I, Florindo JB. Bibliographic research with large language model ChatGPT-4: instability, hallucinations and sometimes alerts. Clinics (Sao Paulo). 2024;79:100409. **[Crossref]**
5. Metze K, Morandin-Reis RC, de Ávila Reis MF, da Silva Fago M, Florindo JB. Misinformation, false positives and delegation of tasks - large language models should not be used for the detection of retracted literature - A study of 21 Chatbots. *J Clin Anesth*. 2025;107:112032. **[Crossref]**
6. Metze K, Mattos AC, Lorand-Metze I. Generative artificial intelligence sycophancy and critical thinking - clues for introducing chatbots in the classroom. *Plast Reconstr Surg*. 2026 Mar 10. [Epub ahead of print]. **[Crossref]**

Thorac Res Pract. 2026;27(4):266-267 **DOI:** 10.4274/ThoracResPract.2026.2026-6-15

Letter to the Editor

Reply Letter to the Editor: Free Large Language Model-based Chatbots Can Help to Align Statistical Tests with the Study Design and Avoid Pseudo-replication

Ömer F. Karakoyun[1], Halil E. Koyuncuoğlu[1], Ömer H. Sağnıç[2], Mehmed E. Özdemir[3], Yalçın Gölcük[4], Birdal Yıldırım[4]

[1]Clinic of Emergency Medicine, Muğla Training and Research Hospital, Muğla, Türkiye
[2]Clinic of Emergency Medicine, Isparta City Hospital, Isparta, Türkiye
[3]Department of Artificial Intelligence, Muğla Sıtkı Koçman University, Graduate School of Natural and Applied Sciences, Muğla, Türkiye
[4]Department of Emergency Medicine, Muğla Sıtkı Koçman University Faculty of Medicine, Muğla, Türkiye

Cite this article as: Karakoyun ÖF, Koyuncuoğlu HE, Sağnıç ÖH, Özdemir ME, Gölcük Y, Yıldırım B. Reply Letter to the Editor: Free large language model-based chatbots can help to align statistical tests with the study design and avoid pseudo-replication. *Thorac Res Pract.* 2026;27(4):266-267

KEYWORDS

Friedman test, Kruskal-Wallis test, large language models, pulmonary embolism, clinical decision support

Received: 17.06.2026 ***Revision Requested:*** 19.06.2026 ***Last Revision Received:*** 23.06.2026

Accepted: 23.06.2026 ***Publication Date:*** 24.07.2026

DEAR EDITOR,

We thank the author(s) of the Letter to the Editor entitled "Letter to the Editor Re: Free Large Language Model-based Chatbots Can Help to Align Statistical Tests with the Study Design and Avoid Pseudo-replication" for their careful reading of our article, "AI in Patient Care: Evaluating Large Language Model Performance Against Evidence-Based Guidelines for Pulmonary Embolism," and for highlighting an important statistical issue.[1,2]

The main concern relates to the use of the Kruskal-Wallis test for comparing scores from four large language model (LLM)-based chatbots. We acknowledge that, because the same ten clinical questions were submitted to each model, the observations are better considered as blocked or repeated measures rather than independent groups. Therefore, a repeated-measures non-parametric approach is more appropriate for the overall comparison.[3,4]

To address this point, we reanalyzed the model scores using the Friedman test, treating each question as a block and the four LLMs as related conditions. The Friedman test showed no statistically significant difference among the four models [χ^2=4.50, P = 0.212]. Since the overall test was not significant, no post-hoc pairwise comparisons were performed.[5] This result is consistent with the main conclusion of our original analysis: although ChatGPT-4o achieved the highest total score and DeepSeek-V2 the lowest, no model demonstrated statistically significant overall superiority.

We agree that the statistical analysis section could have more clearly reflected the repeated-measures structure of the data. Nevertheless, the reanalysis did not change the principal interpretation of the study. The evaluated LLMs demonstrated model-specific strengths and limitations across different components of pulmonary embolism management, but none consistently outperformed the others across all domains.

Accordingly, the relevant Methods and Results statements in the original article should be corrected through an erratum, replacing the Kruskal-Wallis analysis with the Friedman test and reporting the updated result [$\chi^2(3)$=4.50, P = 0.212].

Corresponding author: Prof. Birdal Yıldırım, MD, e-mail: birdalgul@gmail.com

We appreciate the opportunity to clarify this issue and believe that this discussion contributes constructively to the methodological rigor of studies evaluating AI-based clinical decision-support tools.

Footnotes

Authorship Contributions

Concept: Ö.F.K., H.E.K., Ö.H.S., M.E.Ö., Y.G., B.Y., Design: Ö.F.K., H.E.K, Y.G., B.Y.,Data Collection or Processing: H.E.K., Ö.H.S., M.E.Ö., Y.G., Literature Search: Ö.F.K., H.E.K., Ö.H.S., Y.G. Analysis or

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Metze K, da Silva CS, Lorand-Metze I, de Mattos AC. Letter to the Editor: Free large language model-based chatbots can help to align statistical tests with the study design and avoid pseudo-replication. *Thorac Res Pract*.2026;27(4):264-265. **[Crossref]**
2. Karakoyun ÖF, Koyuncuoğlu HE, Sağnıç ÖH, Özdemir ME, Gölcük Y, Yıldırım B. AI in patient care: evaluating large language model performance against evidence-based guidelines for pulmonary embolism. *Thorac Res Pract*. 2026;27(1):38-46. **[Crossref]**
3. Kim HY. Statistical notes for clinical researchers: nonparametric methods for comparing three or more groups. *Restor Dent Endod*. 2014;39(4):329-332. **[Crossref]**
4. Sheldon MR, Fillyaw MJ, Thompson WD. The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs. *Physiother Res Int*. 1996;1(4):221-228. **[Crossref]**
5. Eisinga R, Heskes T, Pelzer B, Te Grotenhuis M. Exact p-values for pairwise comparison of Friedman rank sums, with application to comparing classifiers. *BMC Bioinformatics*. 2017;18:68. **[Crossref]**

Erratum

DOI: 10.4274/ThoracResPract.2025.2025-6-3

Karakoyun ÖF, Koyuncuoğlu HE, Sağnıç ÖH, Özdemir ME, Gölcük Y, Yıldırım B. AI in patient care: evaluating large language model performance against evidence-based guidelines for pulmonary embolism. *Thorac Res Pract.* 2026;27(1):38-46.

The errors were made inadvertently by the authors.

i. The P value is incorrect in the "Results" subsection of the Abstract on page 38.

No significant difference was found in total scores **(*P* = 0.390)**.

The *P* value has been corrected in the "Results" subsection of the Abstract on page 38.

No significant difference was found in total scores **(*P* = 0.212)**.

ii. Information regarding the statistical test used for the overall comparison is missing in the "Statistical Analysis" section on page 42.

The following sentence has been added to the "Statistical Analysis" section on page 42:

Overall differences among the four LLM-based chatbots across the ten case-based questions were evaluated using the Friedman test, considering each question as a repeated-measures block.

iii. The statistical test, test statistic, and P value are incorrect in the "Results" section on page 42.

Although ChatGPT-4o obtained the highest total score and DeepSeek-V2 obtained the lowest, no statistically significant differences were observed among the AI models in overall performance ($H = 3.013$, $P = 0.390$).

The statistical test, test statistic, and *P* value have been corrected in the "Results" section on page 42.

Although ChatGPT-4o obtained the highest total score and DeepSeek-V2 the lowest, the Friedman test showed no statistically significant difference among the four LLM-based chatbots in overall performance across the ten questions, $\chi^2 = 4.50$, $P = 0.212$.

The relevant sections of the article have been corrected accordingly.