EISSN 2979-9139

Official Journal of
TURKISH
THORACIC SOCIETY

Thoracic Research & Practice

Formerly Turkish Thoracic Journal

galenos Publishing House

Volume 27 Issue 1 January 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: January 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 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.

Thoracic Research & Practice
Official Journal of the Turkish Thoracic Society

CONTENTS

Thorac Res Pract. 2026;27(1):1-2 **DOI:** 10.4274/ThoracResPract.2026.2026.230126

Editorial

Beyond the First Impact Factor: Editorial Renewal and Direction in Thoracic Research and Practice

Metin Akgün[1,2], Begüm Ergan[1,3]

[1]Editor-in-Chief, Thoracic Research and Practice
[2]Department of Pulmonary Medicine, Ağrı İbrahim Çeçen University Faculty of Medicine, Ağrı, Türkiye
[3]Department of Pulmonary Medicine, Dokuz Eylül University Faculty of Medicine, İzmir, Türkiye

Cite this article as: Akgün M, Ergan B. Beyond the first impact factor: editorial renewal and direction in thoracic research and practice. *Thorac Res Pract.* 2026;27(1):1-2

KEYWORDS

Bibliometrics, ethics; health policy, peace, pulmonary medicine, Turkish Thoracic Society

Received: 22.01.2026

Accepted: 23.01.2026 ***Publication Date:*** 30.01.2026

In contemporary academic publishing, journal performance is increasingly discussed through numerical indicators such as impact factors, citation counts, and quartile rankings. While these metrics offer useful measures of visibility, they provide only a partial view of a journal's intellectual direction, editorial rigor, and contribution to clinical practice. For specialty journals affiliated with professional societies, the challenge extends beyond achieving visibility to ensuring that published content maintains sustained scientific relevance and practical value.

Against this background, *Thoracic Research and Practice* represents an example of gradual and deliberate editorial evolution rather than abrupt transformation. The journal's origins date to 2001, when *Türk Toraks Dergisi* and *Turkish Respiratory Journal* independently served the respiratory medicine community. Their merger in 2008 under the title *Turkish Thoracic Journal* marked the transition to English-language publication, while the renaming to *Thoracic Research and Practice* in 2023 reflected an intention to broaden scope and strengthen international orientation.[1]

The journal's first inclusion in *Journal Citation Reports* in 2024 constituted an important indexing milestone. With an initial Journal Impact Factor of 0.6 and a Journal Citation Indicator of 0.20, *Thoracic Research and Practice* entered the global citation landscape of respiratory system journals.[2] These figures are best interpreted as retrospective indicators of accumulated editorial effort rather than as determinants of future strategy. Metrics confirm presence, but they do not define editorial purpose.

Within this context, the editorial board renewal at the end of 2025 represents an opportunity to clarify, rather than redefine, editorial priorities. Emphasis has been placed on maintaining clinical relevance, methodological transparency, and consistency in editorial standards across different submission types. The intention is not rapid metric optimization, but the preservation of a coherent editorial approach capable of supporting meaningful international engagement.

Recent content published in the journal reflects this orientation, with studies grounded in real-world clinical practice, analyses addressing longer-term respiratory outcomes, and surveys highlighting variability in respiratory care across healthcare settings. Such contributions may not always generate immediate citation gains, but they contribute to cumulative clinical knowledge and practice-based evidence.

As the official journal of the Turkish Thoracic Society, *Thoracic Research and Practice* occupies a balanced position. It serves as a platform for society-led scientific initiatives while adhering to the norms of international scholarly publishing. This balance requires editorial independence together with institutional responsibility. Within this framework, guidelines, task force reports, and consensus documents remain integral to the journal's mission, provided they meet the same transparent methodological and editorial criteria applied to original research.

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

The *Turkish Thoracic Society Declaration on Peace and Health: The Importance of Breathing in a World Without War* illustrates this approach.[3] By addressing the health consequences of armed conflict, particularly those affecting respiratory health, environmental conditions, and vulnerable populations, the declaration situates peace within a public health framework. Its publication reflects the journal's role as an institutional medium through which scientific evidence, ethical responsibility, and public health concerns can be articulated within established editorial standards.

Looking ahead, the journal's strategic direction emphasizes consolidation rather than expansion for its own sake. International engagement is pursued as a qualitative process shaped by content quality, editorial clarity, and sustained scientific dialogue. Any future growth in bibliometric indicators is expected to follow as a consequence of these efforts rather than as their primary objective.

In an era where editorial success is increasingly quantified, maintaining focus on editorial coherence, clinical relevance, and methodological integrity remains a deliberate and measured stance. The current phase of *Thoracic Research and Practice* is defined less by the attainment of its first impact factor than by the opportunity to continue developing a stable and credible editorial identity within respiratory medicine.

REFERENCES

1. Kılınç O, Oğuzülgen İK. It's time to move forward: thoracic research and practice! *Thorac Res Pract.* 2023;24(1):5. **[Crossref]**
2. Journal Citation Reports—Respiratory System. Clarivate Analytics; 2024. Accessed January 23, 2026. **[Crossref]**
3. Çetinkaya PD, Töreyin ZN, Yüce Özdemir H, et al. Turkish Thoracic Society Declaration on Peace and Health the Importance of Breathing in a World Without War. *Thorac Res Pract.* 2026;27(1):68-74. **[Crossref]**

Thorac Res Pract. 2026;27(1):3-10 **DOI:** 10.4274/ThoracResPract.2025.2025-6-1

Original Article

Excessive Short-acting Beta-agonists Prescriptions in COPD Treated with Triple Inhaler Therapy: A Possible Marker of Frequent Exacerbations. A Retrospective Cohort Study

Bruno Sposato[1], Leonardo Gianluca Lacerenza[2], Elisa Petrucci[2], Alberto Ricci[3], Alberto Cresti[4], Pasquale Baratta[4], Andrea Serafini[5], Claudio Micheletto[6], Maurizio Di Tomassi[7], Antonio Perrella[1], Valerio Alonzi[1], Sara Croce[1], Marco Scalese[8]

[1]Department of Pneumology, Local Health Unit "Sud-Est", "Misericordia" Hospital, Grosseto, Italy
[2]Department of Pharmaceutical Medicine, Local Health Unit "Sud-Est", "Misericordia" Hospital, Grosseto, Italy
[3]Division of Pneumology, Department of Clinical and Molecular Medicine, Sapienza University of Rome, Rome, Sant'Andrea University Hospital (AOU Sant'Andrea), Rome, Italy
[4]Department of Cardiology, Local Health Unit "Sud-Est", "Misericordia" Hospital, Grosseto, Italy
[5]Department of Medical Management, Local Health Unit "Sud-Est", "Misericordia" Hospital, Grosseto, Italy
[6]Department of Pneumology, Azienda USL Toscana Sud-Est, San Andrea Hospital, Massa Marittima (GR), Italy
[7]Institute of Clinical Physiology, Italian National Research Council, Pisa, Italy
[8]Pulmonary Unit, Integrated University Hospital of Verona, Verona, Italy

Cite this article as: Sposato B, Lacerenza LG, Petrucci E, et al. Excessive short-acting beta-agonists prescriptions in COPD treated with triple inhaler therapy: a possible marker of frequent exacerbations. A retrospective cohort study. *Thorac Res Pract.* 2026;27(1):3-10

ABSTRACT

OBJECTIVE: Short-acting β2-agonists (SABA) are used both in asthma and in chronic obstructive pulmonary disease (COPD); SABA use appears to be associated with an increased risk of exacerbations. We evaluated whether COPD patients receiving regular treatment with single-inhaler triple therapy (SITT) used SABA and whether they experienced more exacerbations.

MATERIAL AND METHODS: Our single-center cohort study retrospectively included COPD patients who had been on SITT for 12 months and who were prescribed >7 inhaled corticosteroids/long-acting β2-agonists/long-acting muscarinic antagonist packages. Patients were divided into three groups based on the number of SABA boxes they received during the SITT year: no SABA (0 boxes/year), 1–2 boxes/year, and ≥3 boxes/year. Oral corticosteroids (OC) and antibiotic packs during the SITT year were considered outcomes for the SABA groups.

RESULTS: Five thousand one hundred and seven subjects were recruited, and 1,444 (28.3%) had at least one SABA prescription. Adherence to SITT treatment was similar across the three SABA groups: 10.7±2.8, 10.6±2.8, and 10.9±3.9 packages/year in the 0, 1–2, and ≥3 SABA groups, respectively. The number of OC/antibiotic packages increased progressively across SABA groups from 0 to 1-2 and ≥3 ($P < 0.0001$). When we applied logistic models, we also observed a progressively higher risk of taking OC and antibiotics among subjects who had taken 1-2 packs of SABA [odds ratio (OR): 2.299 (1.878-2.813) and 2.034 (1.621-2.551), respectively; $P < 0.0001$], and among those who had taken ≥3 packs of SABA [OR: 3.472 (2.871-4.200) and 2.714 (2.192-3.362), respectively; $P < 0.0001$].

CONCLUSION: A significant number of subjects were prescribed SABA despite SITT therapy. A relationship between SABA packages and the number of exacerbations, assessed by OC/antibiotic prescriptions, was observed. Excessive SABA use or prescription may indicate frequent exacerbations in patients with COPD despite receiving maximal inhaled therapy.

KEYWORDS: COPD, triple therapy, SABA, exacerbations, oral corticosteroids, antibiotics, real-life

Received: 17.06.2025 ***Revision Requested:*** 07.08.2025 ***Last Revision Received:*** 17.08.2025

Accepted: 15.11.2025 ***Epub:*** 15.01.2026 ***Publication Date:*** 30.01.2026

Corresponding author: Bruno Sposato, MD, e-mail: bru.sposato@gmail.com, bruno.sposato@uslsudest.toscana.it

INTRODUCTION

Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines suggest using either short-acting β2-agonists (SABA) or short-acting muscarinic antagonists as reliever medications in chronic obstructive pulmonary disease (COPD) patients who are already receiving the recommended pharmacological maintenance treatment, which includes long-acting muscarinic antagonists (LAMAs), long-acting β2-agonists (LABAs), and inhaled corticosteroids (ICS), either in combination with LABAs or as LAMAs+LABAs.[1]

However, less is known about the effects of frequent SABA use in COPD. SABA therapy is classically considered a cornerstone in the management of asthma. Global Strategy for Asthma Management and Prevention guidelines[2] recommend using it only as a medication on demand and not as a regular therapy. The systematic use of salbutamol may have a pro-inflammatory effect, potentially increasing the risk of exacerbations, as noted by some researchers.[3-5]

High SABA use appears to be associated with worse outcomes, even in COPD. In fact, an association between a high level of SABA use and a low health status, as measured by the COPD assessment test, has been described.[6] Other researchers found that patients using high doses of SABA had more severe airflow limitation, were more symptomatic, and had worse health status compared with patients taking lower doses of SABA.[7] Furthermore, some authors reported that the level of SABA use during the first month of the study predicted exacerbation risk during the subsequent 10-month period.[8] This was also confirmed by another recent study, which observed that high SABA use is relatively common and associated with a higher risk of exacerbations and all-cause mortality.[9] However, these studies analyzed the effects of SABAs by considering all COPD patients together, regardless of disease severity (groups A, B, and E according to GOLD guidelines) or adherence to treatment. It is not clear whether patients in group E, who are frequent exacerbators, may have the highest SABA use. Furthermore, it is conceivable that high use of SABA may be associated with an elevated risk of exacerbations in this group of patients, despite regular treatment, including triple therapy. GOLD guidelines recommend ICS/LABA/LAMA therapy for subjects who have frequent moderate-to-severe exacerbations despite receiving LABA/LAMA therapy or who have blood eosinophils >300 cells/μL.[1] Triple therapy, especially when taken with a single-inhaler triple therapy (SITT), has been demonstrated to be effective in improving lung function, symptoms, health status, and in reducing moderate/severe COPD exacerbations compared to ICS, LABA, or LAMA monotherapies and LAMA/LABA and ICS/LABA combinations,[10-14] and even when compared to triple treatment with multiple devices.[15-17]

Main Points

- Excessive use of short-acting β2-agonists (SABA) was associated to chronic obstructive pulmonary disease (COPD) exacerbations.
- There was a relationship between SABA packages and number of COPD exacerbations.
- Excessive SABA prescriptions may be used as a marker of possible disease exacerbation.

Given the uncertainties on this matter, we evaluated, in a group of patients continuously treated with SITT, the prevalence of SABA use and whether high SABA use was associated with an increased risk of COPD exacerbations compared with low SABA use, despite optimal adherence to SITT.

MATERIAL AND METHODS

Study Design

We retrospectively extracted from the pharmaceutical prescriptions archive database patients who had received triple therapy from a single dispenser (SITT) with fluticasone furoate/vilanterol/umeclidinium [dry powder inhaler (DPI)] or beclometasone dipropionate/formoterol fumarate/glycopyrronium [metered dose inhaler (MDI)] for 1 year. Among them, we considered only those prescribed more than seven triple-therapy per year. Packages of systemic corticosteroids [Anatomical Therapeutic Chemical Classification (ATC) code: H02], antibiotics (ATC code: J01), and salbutamol (SABA; ATC code: R03AC02) prescribed during the year of SITT were considered for each patient, identified using tax codes. Only MDI formulations of SABA were considered. In Italy, a medical prescription is required to obtain SABA at pharmacies. The number of SABA packages prescribed for each individual during the year of SITT therapy was used to divide patients into three groups: 0, 1-2, and ≥3 packs/year. In each group, the number of oral corticosteroids (OC) and antibiotic boxes (commonly used for COPD exacerbations) prescribed during the year of treatment with SITT was evaluated and compared across the three groups. In this study, COPD exacerbation was defined as the prescription of at least one package of OC, with or without concomitant antibiotics.

Setting and Participants

The USL SUDEST-Tuscany pharmaceutical database for the period 2019-2023 was consulted. As already said, we extracted from this database all patients on SITT therapy who got more than 7 boxes of ICS/LABA/LAMA during a year of triple treatment.

Prescribing SITT allowed us to identify only those subjects affected by COPD. In Italy, a pulmonary specialist is authorized and required to prescribe SITT with a therapeutic plan after establishing a correct COPD diagnosis. In Italy, SITT is considered only for patients diagnosed with COPD who have had at least two moderate-to-severe exacerbations in the previous year despite LAMA/LABA or ICS/LABA therapy (group E, according to GOLD guidelines). Therefore, our selection from the database of patients treated only with SITT allowed us to identify those with a definite COPD diagnosis classified as group E according to GOLD guidelines.[1] For each patient, a 12-month period of treatment was considered from the month they started through the twelfth month thereafter.

The study was approved by Area Vasta Sudest Ethical Committee (C.E.A.S.V.E.), Azienda Ospedaliera Universitaria Senese, and Azienda USL Toscana Sud-Est (Protocol TRIPLECOPD, ID:

19196; determination: N° 358, 16/02/2021) on the basis that it complied with the Declaration of Helsinki and that the protocol followed existing Good Clinical Practice guidelines.

Inclusion Criteria

We included patients who got more than 7 boxes of SITT during a year of observation, who were >40 years old and who never underwent any changes to other triple therapies during the year considered. SITT prescriptions of >7 boxes/year confirmed the diagnosis of COPD (in Italy, until the end of 2023, the prescription of SITT was exclusive to patients with COPD diagnosed by a pulmonologist with spirometry and clinical features suggestive of the disease).

The study period was 2019-2023. For each patient, a 12-month treatment period was considered, from the month they started therapy through the twelfth month thereafter.

Exclusion Criteria

Patients aged and with fewer than 7 boxes/year of SITT were excluded from the study. Changes to other triple therapies during the year of treatment were another exclusion criterion. The association of Montelukast with SITT was also considered one more reason to exclude patients as it is a drug prescribed for asthma which could have altered the results.

Variables and Measurements

Age, gender, comorbidities associated with COPD, and the number of SITT packages, OC packages, and antibiotic packages during the year of triple treatment were considered for each patient. The systemic corticosteroids considered in this study, which can be used in COPD exacerbations, were betamethasone, dexamethasone, methylprednisolone, prednisolone, prednisone, and deflazacort. Antibiotics, such as penicillins with an extended spectrum of activity, combinations of penicillins with beta-lactamase inhibitors, cephalosporins, macrolides, fluoroquinolones, combinations of sulphonamides and trimethoprim, and some aminoglycosides, were considered for this research because they can also be used to treat COPD exacerbations.

The dispensing of >7 packages of maintenance inhaler medications (SITT) was the cut-off used to identify proper adherence to treatment. Individuals taking fewer than 7 ICS/LABA/LAMA packs per year were excluded from the study because they were either poorly adherent or had their therapy changed (i.e., switched to other inhaler treatments) during the year under consideration.

Prescriptions for drugs indicated for other diseases were also reviewed for each patient to identify comorbidities affecting the individuals examined in this study. Comorbidities were identified by searching for medications taken by each patient for various COPD-related diseases during triple treatment. The criterion for identifying comorbidities was the use of at least three boxes per year of each medication taken for non-COPD diseases. The following ATC codes of drugs used together with the triple A02 were identified: A07, A10, A11/A12/H05/M05, B01/C01/C03C/C03D/C03E/C08D, B03, B03X, C02/C03A/C03B/C07/C08/C09, C10, G04, H03, L01/L02/L03/N02, L04/M01/M04AC01/P01, M04, N03/N04, N05/N06/N07, S01. Such codes identified the following conditions: gastroesophageal reflux/dyspepsia, intestinal disorders, diabetes, osteoporosis, and cardiovascular diseases, including heart failure, cardiac arrhythmias, and coronary artery disease. cerebrovascular diseases (considered all together), anemia, renal failure, hypertension, dyslipidemia, prostatic hypertrophy, thyroid disorders, oncology pathologies, autoimmune disorders, hyperuricemia, neurological disorders, psychiatric disorders, and glaucoma.

Unavailable Variables

Patients' lung function and other clinical data (smoking status, body mass index, blood eosinophils) were unavailable because subjects came from different parts of southeastern Tuscany and the data had been archived in various local databases that were not readily accessible. However, for a triple therapy prescription, as previously stated, a diagnosis of symptomatic moderate-to-severe COPD with a history of frequent and/or severe exacerbations (according to GOLD guidelines)[1] was required. Such a prescription had to be made exclusively by a pulmonologist (not by a GP) and accompanied by a written treatment plan, as indicated by the Italian Drug Agency. In addition, the diagnosis of COPD had to be made by spirometry, as recommended by GOLD guidelines.[1] However, the absence of the aforementioned data did not affect the objectives of our study.

Statistical Analysis

Chi-square tests were used to compare categorical variables. The Kruskal-Wallis test was used to compare continuous variables among the different groups.

Multivariate analysis was also performed. Linear and logistic regression models were applied to test whether there was a relationship between the number of SABA packages and the number of OC/antibiotics prescriptions. All models adjusted for sex, age, adherence to treatment with SITT (number of packages per year), and comorbidities. P values < 0.05 were considered statistically significant.

RESULTS

A total of 5,107 patients were enrolled in SITT therapy for 1 year. Among these, 1,444 (28.3%) had at least one SABA prescription; in particular, those who had 1-2 boxes were 651 (12.8%), while 793 (15.5%) had ≥3 (Figure 1). Patient characteristics are described in Table 1. Females tend to have more SABA prescriptions than males (626/1,844 - 34% vs. 818/3,263-25%; P = 0.0001), while older subjects used fewer packages of short-acting bronchodilators. The number of comorbidities appeared to be significantly higher in subjects with a higher number of SABA box prescriptions (Table 1). Over 60% of patients have ≥4 comorbidities. Adherence to SITT (number of triple-therapy boxes per year) did not differ among the three groups (means ± standard deviation: 10.7±2.8, 10.6±2.8, and 10.9±3.9 for subjects using 0, 1-2, and ≥3 SABA packages per year, respectively; P = 0.310; Table 1). Instead, the number of OC/antibiotic packages per year increased progressively with the number of SABA packs prescribed during the year of SITT treatment (Figure 2A and 2B).

Table 1. Baseline demographics and clinical characteristics of COPD patients subdivided into 3 groups on the basis of SABA prescriptions

	SABA packages/year			*P*
	0	**1-2**	**≥3**	
Sex				
Males	2445 (66.7%)	387 (59.4%)	431 (54.4%)	0.0001
Females	1218 (33.3%)	264 (40.6%)	362 (45.6%)	
Age				
≤70 years	964 (26.3%)	184 (28.3%)	301 (38%)	0.0001
71-79 years	1404 (38.3%)	242 (37.2%)	310 (39.1%)	
≥80 years	1295 (35.4%)	225 (34.6%)	182 (23%)	
Comorbidities				
0-1	437 (11.9%)	65 (10%)	100 (12.6%)	0.040
2-3	956 (26.1%)	144 (22.1%)	214 (27%)	
≥4	2270 (62%)	442 (67.9%)	479 (60.4%)	
SITT packages/year	10.7±2.8	10.6±2.8	10.9±3.9	0.310

COPD: chronic obstructive pulmonary disease, SABA: short-acting β2-agonists, SITT: single-inhaler triple therapy

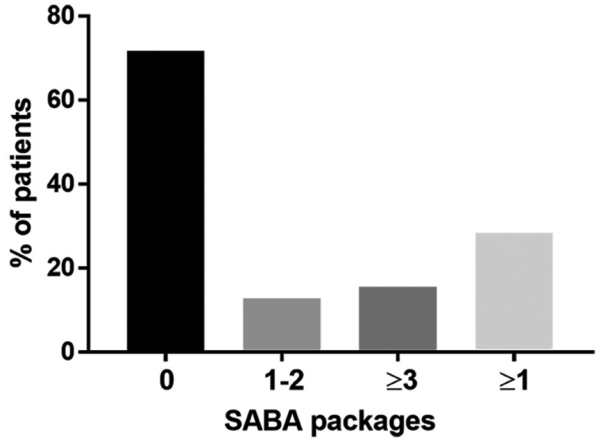

Figure 1. Prevalence of subjects who had SABA prescriptions

SABA: short-acting β2-agonists

When we applied multivariate analysis (linear regression models) adjusted for all confounding variables, we found that the groups with 1-2 and ≥3 SABA prescriptions per year received higher numbers of OC (β: 0.905±0.238 and β: 1.698±0.222, respectively) and of antibiotic packages (β: 1.256±0.287 and β: 1.853±0.269, respectively; $P < 0.0001$) than subjects who had never taken SABA during the year of SITT therapy (Figure 3A and 3B).

Even when we applied logistic models, we observed a progressively higher risk of taking OCs and antibiotics among subjects who had taken 1-2 packets of SABA [odds ratios (OR): 2.299 (1.878-2.813) and 2.034 (1.621-2.551), respectively; $P < 0.0001$] and ≥3 packets of SABA [ORs: 3.472 (2.871-4.200) and 2.714 (2.192-3.362), respectively; $P < 0.0001$] (Figure 4A and 4B).

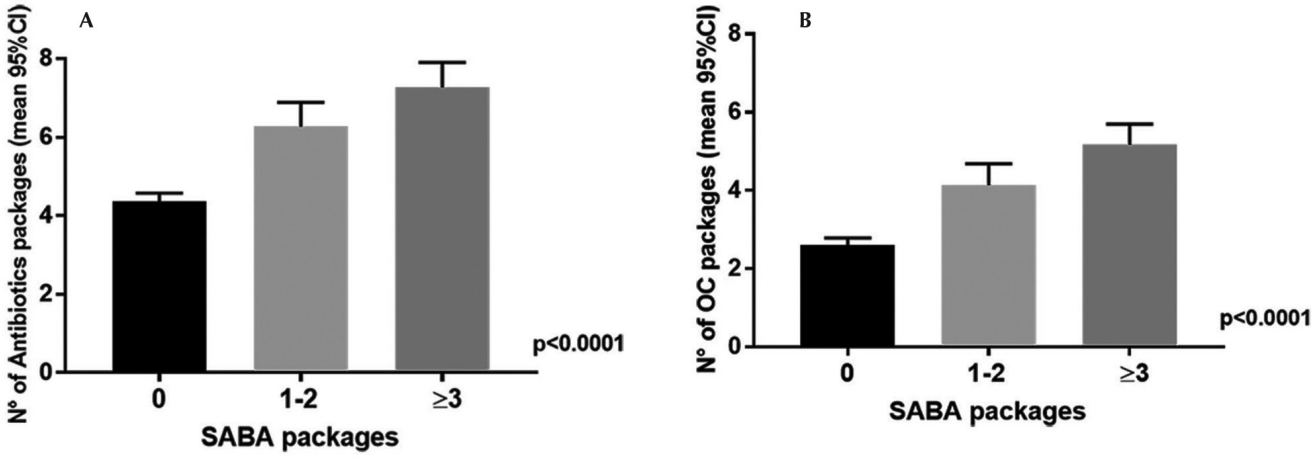

Figure 2. Number of oral corticosteroid and antibiotic packages used in the different SABA groups

SABA: short-acting β2-agonists, OC: oral corticosteroids, CI: confidence interval

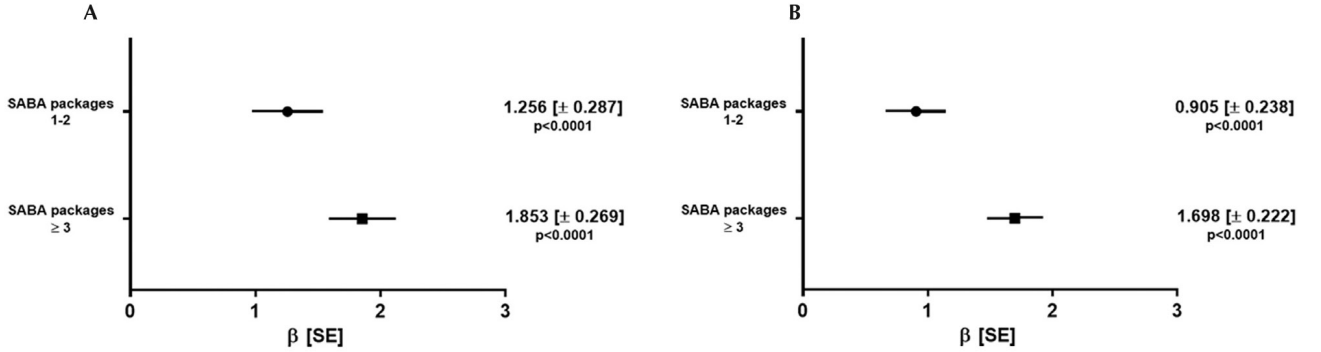

Figure 3. Number of additional OC (A) and antibiotic (B) boxes prescribed in the groups of subjects who had 1-2 or >3 boxes of SABA compared to those who never took SABA, calculated with a linear regression model adjusted for sex, age, comorbidities and number of SITT boxes used in the year of follow-up considered

SABA: short-acting β2-agonists, OC: oral corticosteroids, SITT: single-inhaler triple therapy, SE: standard error

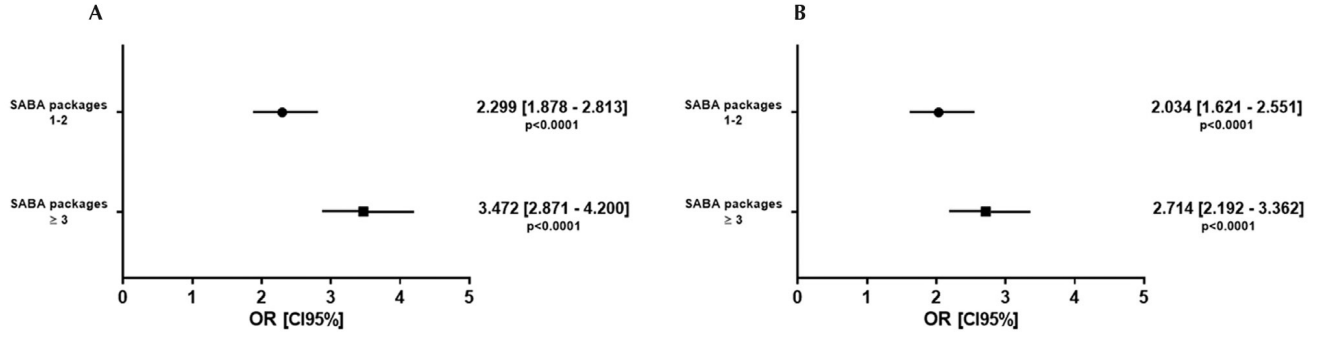

Figure 4. Risk (odds ratios) of using OC (A) and antibiotics (B) in subjects who took 1-2 or >3 boxes of SABA compared to those who never had any SABA prescription calculated with logistic models adjusted for sex, age, comorbidities and number of boxes of SITT taken during the year of follow-up considered

SABA: short-acting β2-agonists, OC: oral corticosteroids, SITT: single-inhaler triple therapy, CI: confidence interval, OR: odds ratios

DISCUSSION

This study highlighted that, in a very large COPD population in group E, despite receiving maximized SITT treatment, about 28% of patients had at least one SABA prescription. Furthermore, we observed an association between the number of SABA packages prescribed and COPD exacerbations in patients regularly treated with SITT, as evaluated by OC/antibiotic use. In fact, as SABA prescriptions increased, OC/antibiotic prescriptions also showed a progressive increase. Our study examines only patients with optimal adherence to triple treatment (approximately 11 packs/year of SITT), unlike other studies[8,9] that considered all COPD patients treated with mono-, dual, and triple therapies regardless of their level of adherence. This highlights that despite maximum and regular SITT, many individuals still use SABA because they remain symptomatic and are therefore susceptible to exacerbations. Moreover, patients receiving SITT (group E according to GOLD guidelines) have substantially more severe COPD than patients in GOLD groups A and B and are therefore more difficult to treat.[18] In our analysis, we considered OC/antibiotic prescriptions as outcomes of exacerbations. When they are at least moderate in severity, they should be treated with OCs and/or antibiotics according to guidelines.[1] Therefore, the outcome of OC/antibiotics prescriptions analyzed in our study during the year of SITT should be considered to correspond to COPD exacerbations.

A recent study found that 30.1% of patients used SABAs during the 12-month baseline period,[9] which is comparable to what we observed. Our study only included patients belonging to group E according to GOLD guidelines,[1] i.e., those with more severe COPD characterized by frequent exacerbations (more precisely, a distinct COPD phenotype). In fact, according to another study, these subjects use more SABA than patients with less severe disease do, namely those belonging to groups A and B according to GOLD recommendations.[9] Excessive use of SABA indicates that patients, despite regular treatment with ICS/LABA/LAMA (SITT), remain symptomatic and are therefore at increased risk of disease exacerbations. Subjects who use SABA excessively may be patients with very severe COPD, for which even regular triple therapy fails to control the disease, suggesting that high SABA use is a marker of more severe disease. A recent study highlighted that patients with high SABA use had more severe airflow limitation, were more symptomatic, and had worse health status than patients with lower SABA use.[7] Therefore, these patients should receive additional treatments, such as roflumilast and/or azithromycin, as recommended by guidelines.[1] Dupilumab should also be considered because recent evidence has shown significant efficacy in reducing exacerbations in COPD patients receiving triple therapy who have blood eosinophils <300 cells/microliter.[19] Conversely, using high doses of ICS does not appear to be effective. A recent meta-analysis has shown that high doses of ICS, compared

with medium doses, do not appear to further reduce COPD exacerbations and mortality.[20] Also, rehabilitation interventions and the best possible management of comorbidities must always be considered; in particular, management of comorbidities remains an important determinant of health in COPD.[1,21]

Our study demonstrated a strong association between the number of SABA prescriptions and OC/antibiotics packages dispensed. This suggests that increased SABA use is associated with a higher number of COPD exacerbations. This appears to be in line with observations by other authors.[7-9] As further confirmation, a systematic review reported that reducing the number of rescue puffs per day proportionally decreased the rate of moderate-to-severe COPD exacerbations.[22] High SABA use often indicates COPD deterioration, characterized by declining lung function and worsening symptoms such as dyspnoea, sore throat, and cough, which usually precede disease exacerbations.[23-25] Another randomized clinical trial showed that increased SABA use related to symptom deterioration is common and contributes to negative outcomes in patients with COPD.[26] Thus, the use of SABA in COPD, even with optimal continuous treatment, may predict a subsequent exacerbation.

Another possible explanation for excessive SABA use is the presence of marked bronchial hyperreactivity (BHR) in some patients with COPD. Airway hyperresponsiveness is a hallmark of obstructive airway disorders, such as asthma and COPD.[27] Bronchospasm events in COPD may be related to the presence of BHR. The reported prevalence of BHR in COPD patients was consistently higher than in the general population, reaching values of 63% in men, and 87% when assessed with the methacholine challenge test using a dose of 25 mg/mL as a cut-off to identify it.[28] Airway hyperresponsiveness significantly increases the frequency of exacerbations in COPD, regardless of airway eosinophilia.[29,30] SITT, which also includes ICS treatment, may not effectively reduce airway hyperresponsiveness in all patients. In fact, ICS therapy does not always appear to reduce BHR in patients with COPD.[29] High SABA use has been associated not only with increased bronchial responsiveness but also with the development of tolerance to the bronchodilating effect of β2-agonists.[31] A reduced efficacy of LABAs in patients with COPD who are high SABA users has also been reported.[7] Prolonged or repeated use of β2-agonists leads to the loss of some of their effects, a pervasive phenomenon termed tachyphylaxis, refractoriness, or desensitization.[32] Regular β2-agonist use can induce tolerance to their bronchoprotective effects and reduce bronchodilator responsiveness to β2-agonists.[33] Moreover, a study observed that salmeterol might be subject to tachyphylaxis because the duration of action and the peak effect decreased over time.[34] Prolonged use of LABAs may therefore lead to receptor tachyphylaxis and reduced therapeutic efficacy of these drugs, which may consequently justify greater use of SABA in COPD patients even if treated with SITT. It is also possible that excessive and continuous SABA use could be responsible for down-regulation of β2-agonist receptors (tachyphylaxis), thus leading to a consequential and progressive reduction in the efficacy of LABAs when taken with triple therapy, thereby predisposing the patient to exacerbations.

Another aspect to consider is that comorbidities can influence disease exacerbations.[35] In fact, real-world data confirm that, in patients with COPD, they also affect their frequency and their COPD-related health care resource utilization.[36] According to recent studies, COPD patients receiving SITT were older and had more comorbidities, especially cardiovascular diseases, than patients receiving LABA/LAMA or other treatments.[37] In our study, subjects with more SABA packages had more comorbidities, confirming that these comorbidities could be associated with greater SABA use and, consequently, with an increased number of COPD exacerbations. As we have already said, in addition to inhaled treatment, it is necessary to optimize all available therapies for the various comorbidities associated with COPD.

Another aspect to consider is that some patients on triple therapy may make critical errors during various phases of inhalation therapy, such as performing inhalations incorrectly. Some studies report that, even with triple therapy in both MDI and DPI, 10-15% of patients make critical errors with the device.[38,39] This could compromise the effectiveness of the therapy, thus pushing them to use SABA more often and thereby increasing the risk of disease exacerbations.

In summary, excessive SABA prescriptions could be a marker of poor treatment efficacy even in patients receiving regular SITT. These subjects should be identified and re-evaluated, with improved management of comorbidities, optimization of inhalation technique to the greatest extent possible, and addition of therapies to achieve optimal disease control. It is also necessary to investigate whether asthma coexists with COPD, as this can occur in between 4.2% and 66% of individuals identified as having COPD.[40] This might allow us to consider additional therapies currently used solely for the treatment of asthma.

Another issue to consider is the differential influence of gender on aspects of the disease, with possible repercussions for differences in treatment efficacy between males and females. Notably, in our study we observed a higher frequency of SABA prescriptions in female patients compared with males, a pattern already observed in asthma.[5] The explanations for this could be multifactorial: differences in types of airway inflammation; a higher prevalence of obesity in females; greater BHR, which may lead to greater symptom perception; and greater anxiety and mood disturbances[5] when females are compared with males. Other explanations are also possible, but this aspect should be investigated further by designing an appropriate study to analyze it.

Study Limitations

The major limitation of our study is that clinical data, such as lung function measurements, smoking status, and symptoms, were not available in our dataset; the strength of our analysis is the large number of subjects enrolled and their strong adherence to SITT. It should also be noted that the drug prescriptions referenced in our study may not always correspond to actual use of the drug in question. For example, not all prescribed SABAs may have been used by patients. In fact, physicians often prescribe SABAs for possible emergencies, which, in many

cases, do not occur. In other circumstances, patients want to feel secure that the drug will be available if needed. Therefore, SABA prescriptions may not always correspond to their actual use. However, our analysis considered drug prescriptions rather than actual drug use because the study was conducted using a database.

OCs administered may also not correspond to the actual number of exacerbations. Some OC packages may cover multiple exacerbation treatments and therefore lead to underestimation of their intake. On the other hand, it is possible that the prescription of OCs and antibiotics does not always correspond to their use. This may not accurately capture all true exacerbation events, potentially leading to misclassification.

However, we believe that the large number of patients studied may mitigate such biases. Furthermore, our study evaluated drug prescriptions dispensed directly by pharmacies, that is, drug packages taken from the pharmacy by them. Therefore, we believe that prescribed drugs are frequently used.

Some SABA prescriptions may have been given during exacerbations in our study. However, as we have previously stated, our study only considered prescriptions for SABA in MDI formulations. It should also be noted that in Italy, COPD exacerbations are commonly treated with OCs combined with antibiotics, and in some cases with the addition of SABA, primarily via nebulized formulations rather than using MDI devices. Furthermore, two similar studies confirm that COPD patients use SABA primarily to control symptoms despite ongoing maintenance therapies, among those with more severe disease.[7,9]

CONCLUSION

In a significant number of COPD patients belonging to group E SABA was prescribed despite SITT therapy. This means that, despite such treatment, they still did not demonstrate complete disease control. Furthermore, there is a significant relationship between SABA use and OC and antibiotic prescriptions commonly used during disease exacerbations. In other words, an increase in SABA use is associated with an increased risk of COPD exacerbations.

Therefore, excessive SABA use could be a marker of frequent exacerbations in patients with COPD despite maximal inhaled therapy. Other therapies, in addition to SITT, should be considered for them.

Ethics

Ethics Committee Approval: The study was approved by Area Vasta Sudest Ethical Committee (C.E.A.S.V.E.), Azienda Ospedaliera Universitaria Senese, and Azienda USL Toscana Sud-Est (Protocol TRIPLECOPD, ID: 19196; determination: N° 358, 16/02/2021) on the basis that it complied with the Declaration of Helsinki and that the protocol followed existing Good Clinical Practice guidelines.

Informed Consent: Due to the retrospective nature of the study and the exceptionally large sample size, reaching each individual participant was deemed unfeasible. All data were processed in a pseudonymized form to ensure the highest level of privacy protection in accordance with the Italian laws. No patients were identifiable at any time during the entire study period. Informed consent was obtained only in those patients in whom it was possible to acquire it.

Footnotes

Authorship Contributions

Surgical and Medical Practices: B.S., Concept: B.S., Design: B.S.,L.G.L., E.P., A.S., Data Collection or Processing: B.S.,L.G.L., E.P., A.S., Analysis or Interpretation: B.S., L.G.L., E.P., A.R., A.C., P.B., A.S., C.M., M.D.T., A.P., V.A., S.C., M.S., Literature Search: B.S., A.R., A.C., P.B., M.D.T., A.P., V.A., S.C., M.S., Writing: B.S., A.R., A.C., P.B., M.D.T., A.P., V.A., S.C., M.S.

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 (GOLD). Global Strategy for the Diagnosis, Management and Prevention of COPD. 2023. Accessed June 06, 2024. [Crossref]
2. Global Strategy for Asthma Management and Prevention (GINA) 2023. Accessed June 06, 2024. [Crossref]
3. Gauvreau GM, Jordana M, Watson RM, Cockroft DW, O'Byrne PM. Effect of regular inhaled albuterol on allergen-induced late responses and sputum eosinophils in asthmatic subjects. *Am J Respir Crit Care Med*. 1997;156(6):1738-1745. [Crossref]
4. Ritchie AI, Singanayagam A, Wiater E, Edwards MR, Montminy M, Johnston SL. β_2-agonists enhance asthma-relevant inflammatory mediators in human airway epithelial cells. *Am J Respir Cell Mol Biol*. 2018;58(1):128-132. [Crossref]
5. Sposato B, Petrucci E, Lacerenza GL, et al. Sex differences in excessive oral corticosteroid exposure in poor adherent adult asthmatics overusing short-acting β-2 agonists. *Minerva Med*. 2023;114(5):642-651. [Crossref]
6. Gondalia R, Bender BG, Theye B, Stempel DA. Higher short-acting beta-agonist use is associated with greater COPD burden. *Respir Med*. 2019;158:110-113. [Crossref]
7. Maltais F, Naya IP, Vogelmeier CF, et al. Salbutamol use in relation to maintenance bronchodilator efficacy in COPD: a prospective subgroup analysis of the EMAX trial. *Respir Res*. 2020;21(1):280. [Crossref]
8. Jenkins CR, Postma DS, Anzueto AR, et al. Reliever salbutamol use as a measure of exacerbation risk in chronic obstructive pulmonary disease. *BMC Pulm Med*. 2015;15:97. [Crossref]
9. Janson C, Wiklund F, Telg G, Stratelis G, Sandelowsky H. High use of short-acting β_2-agonists in COPD is associated with an increased risk of exacerbations and mortality. *ERJ Open Res*. 2023;9(3):00722-2022. [Crossref]
10. Zheng Y, Zhu J, Liu Y, et al. Triple therapy in the management of chronic obstructive pulmonary disease: systematic review and meta-analysis. *BMJ*. 2018;363:k4388. [Crossref]
11. Lipson DA, Barnhart F, Brealey N, et al. Once-daily single-inhaler triple versus dual therapy in patients with COPD. *N Engl J Med*. 2018;378(18):1671-1680. [Crossref]
12. Singh D, Papi A, Corradi M, et al. Single inhaler triple therapy versus inhaled corticosteroid plus long-acting β2-agonist therapy

for chronic obstructive pulmonary disease (TRILOGY): a double-blind, parallel group, randomised controlled trial. *Lancet*. 2016;388(10048):963-973. **[Crossref]**

13. Papi A, Vestbo J, Fabbri L, et al. Extrafine inhaled triple therapy versus dual bronchodilator therapy in chronic obstructive pulmonary disease (TRIBUTE): a double-blind, parallel group, randomised controlled trial. *Lancet*. 2018;391(10125):1076-1084. **[Crossref]**
14. Calzetta L, Ritondo BL, de Marco P, Cazzola M, Rogliani P. Evaluating triple ICS/LABA/LAMA therapies for COPD patients: a network meta-analysis of ETHOS, KRONOS, IMPACT, and TRILOGY studies. *Expert Rev Respir Med*. 2021;15(1):143-152. **[Crossref]**
15. Halpin DMG, Kendall R, Shukla S, et al. Cost-effectiveness of single- versus multiple-inhaler triple therapy in a UK COPD population: the INTREPID trial. *Int J Chron Obstruct Pulmon Dis*. 2022;17:2745-2755. **[Crossref]**
16. Lin L, Liu C, Cheng W, et al. Comparison of treatment persistence, adherence, and risk of exacerbation in patients with COPD treated with single-inhaler *versus* multiple-inhaler triple therapy: a prospective observational study in China. *Front Pharmacol*. 2023;14:1147985. **[Crossref]**
17. Sposato B, Ricci A, Lacerenza LG, et al. Triple therapy in COPD in real life: is it better to use single or multiple inhalers? *J Clin Med*. 2024;13(20):6191. **[Crossref]**
18. Roche N, Antoniadis A, Hess D, et al. Are there specific clinical characteristics associated with physician's treatment choices in COPD? *Respir Res*. 2019;20(1):189. **[Crossref]**
19. Bhatt SP, Rabe KF, Hanania NA, et al. Dupilumab for COPD with blood eosinophil evidence of type 2 inflammation. *N Engl J Med*. 2024;390(24):2274-2283. **[Crossref]**
20. Archontakis Barakakis P, Tran T, You JY, et al. High versus medium dose of inhaled corticosteroid in chronic obstructive lung disease: a systematic review and meta-analysis. *Int J Chron Obstruct Pulmon Dis*. 2023;18:469-482. **[Crossref]**
21. Raghavan S, Hatipoğlu U, Aboussouan LS. Goals of chronic obstructive pulmonary disease management: a focused review for clinicians. *Curr Opin Pulm Med*. 2025;31(2):156-164. **[Crossref]**
22. Punekar YS, Sharma S, Pahwa A, Takyar J, Naya I, Jones PW. Rescue medication use as a patient-reported outcome in COPD: a systematic review and regression analysis. *Respir Res*. 2017;18(1):86. **[Crossref]**
23. Watz H, Tetzlaff K, Magnussen H, et al. Spirometric changes during exacerbations of COPD: a post hoc analysis of the WISDOM trial. *Respir Res*. 2018;19(1):251. **[Crossref]**
24. Seemungal TA, Donaldson GC, Bhowmik A, Jeffries DJ, Wedzicha JA. Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2000;161(5):1608-1613. **[Crossref]**
25. Calverley P, Pauwels Dagger R, Löfdahl CG, et al. Relationship between respiratory symptoms and medical treatment in exacerbations of COPD. *Eur Respir J*. 2005;26(3):406-413. **[Crossref]**
26. Ferguson GT, Skärby T, Nordenmark LH, et al. Unreported and overlooked: a post hoc analysis of COPD symptom-related attacks from the RISE study. *Int J Chron Obstruct Pulmon Dis*. 2020;15:3123-3134. **[Crossref]**
27. Jude JA, Panettieri RA Jr. Bronchomotor tone imbalance evokes airway hyperresponsiveness. *Expert Rev Respir Med*. 2024;18(11):835-841. **[Crossref]**
28. Borak J, Lefkowitz RY. Bronchial hyperresponsiveness. *Occup Med (Lond)*. 2016;66(2):95-105. **[Crossref]**
29. Kume H, Hojo M, Hashimoto N. Eosinophil inflammation and hyperresponsiveness in the airways as phenotypes of COPD, and usefulness of inhaled glucocorticosteroids. *Front Pharmacol*. 2019;10:765. **[Crossref]**
30. Zanini A, Cherubino F, Zampogna E, Croce S, Pignatti P, Spanevello A. Bronchial hyperresponsiveness, airway inflammation, and reversibility in patients with chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis*. 2015;10:1155-1161. **[Crossref]**
31. Beasley R, Pearce N, Crane J, Burgess C. Beta-agonists: what is the evidence that their use increases the risk of asthma morbidity and mortality? *J Allergy Clin Immunol*. 1999;104(2 Pt 2):S18-S30. **[Crossref]**
32. Cazzola M, Page CP, Calzetta L, Matera MG. Pharmacology and therapeutics of bronchodilators. *Pharmacol Rev*. 2012;64(3):450-504. **[Crossref]**
33. Cheung D, Timmers MC, Zwinderman AH, Bel EH, Dijkman JH, Sterk PJ. Long-term effects of a long-acting beta 2-adrenoceptor agonist, salmeterol, on airway hyperresponsiveness in patients with mild asthma. *N Engl J Med*. 1992;327(17):1198-1203. **[Crossref]**
34. Tsagaraki V, Amfilochiou A, Markantonis SL. Evidence of tachyphylaxis associated with salmeterol treatment of chronic obstructive pulmonary disease patients. *Int J Clin Pract*. 2006;60(4):415-421. **[Crossref]**
35. Martinez CH, Mannino DM, Divo MJ. Defining COPD-related comorbidities, 2004-2014. *Chronic Obstr Pulm Dis*. 2014;1(1):51-63. **[Crossref]**
36. Krishnan JK, Martinez FJ, Altman P, et al. Multimorbidities in COPD are associated with increased exacerbations and health care resource utilization in real-world patients from a U.S. database. *Chronic Obstr Pulm Dis*. 2024;11(5):472-481. **[Crossref]**
37. Dondi L, Ronconi G, Calabria S, et al. Correction: clinical characteristics, use and switch of drugs for obstructive airway diseases among patients with COPD experiencing an exacerbation: a retrospective analysis of Italian administrative healthcare data. *BMC Pulm Med*. 2024;24(1):606. **[Crossref]**
38. Brusselle G, Himpe U, Fievez P, et al. Evolving to a single inhaler extrafine LABA/LAMA/ICS - inhalation technique and adherence at the heart of COPD patient care (TRIVOLVE). *Respir Med*. 2023;218:107368. **[Crossref]**
39. Kwok WC, Ma TF, Tsui CK, Ho JCM, Tam TCC. Prospective randomized study on switching triple inhaler therapy in COPD from multiple inhaler devices to a single inhaler device in a chinese population. *Chronic Obstr Pulm Dis*. 2025;12(1):52-60. **[Crossref]**
40. Uchida A, Sakaue K, Inoue H. Epidemiology of asthma-chronic obstructive pulmonary disease overlap (ACO). *Allergol Int*. 2018;67(2):165-171. **[Crossref]**

Thorac Res Pract. 2026;27(1):11-20 **DOI:** 10.4274/ThoracResPract.2025.2025-6-6

Original Article

Expiratory Muscle Strength, Peak Oxygen Consumption and Hyperinflation Predicts Severe Exacerbation in Chronic Obstructive Pulmonary Disease Patients

Mahima Mayee Tripathy[1], Aqsa Mujaddadi[1], Obaidullah Ahmed[2], Deepak Talwar[2]

[1]Centre for Physiotherapy and Rehabilitation Sciences, Jamia Millia Islamia, New Delhi, India
[2]Metro Centre for Respiratory Diseases, Metro Hospitals & Heart Institute, Noida, India

Cite this article as: Tripathy MM, Mujaddadi A, Ahmed O, Talwar D. Expiratory muscle strength, peak oxygen consumption and hyperinflation predicts severe exacerbation in chronic obstructive pulmonary disease patients. *Thorac Res Pract.* 2026;27(1):11-20

ABSTRACT

OBJECTIVE: To explore the predictive ability of physiological and clinical parameters, including respiratory muscle strength, peak oxygen consumption, exercise capacity assessed by the six-minute walk distance (6MWD), pulmonary function, and arterial blood gas for identifying patients with chronic obstructive pulmonary disease (COPD) who are at risk of frequent severe acute exacerbations.

MATERIAL AND METHODS: This retrospective, observational study analyzed data from 265 patients who were hospitalized for severe exacerbations between January 1st, 2018 to February 28th, 2024. Patients were classified as infrequent or frequent exacerbators based on the annual frequency of severe exacerbations. Binary logistic regression models were used to identify independent predictors, adjusting for clinically relevant covariates.

RESULTS: In adjusted multivariate analysis, maximal expiratory pressure [odds ratio (OR): 0.989; 95% confidence interval (CI): 0.980–0.998; $P = 0.014$], 6MWD (OR: 0.997; 95% CI: 0.994–1.000; $P = 0.028$), 6MWD% (OR: 0.985; 95% CI: 0.970–0.999; $P = 0.041$), peak oxygen consumption (OR: 0.874; 95% CI: 0.776–0.986; $P = 0.028$), residual volume (OR: 1.006; 95% CI: 1.001–1.011; $P = 0.017$), and functional residual capacity (OR: 1.008; 95% CI: 1.001–1.014; $P = 0.028$) emerged as significant predictors of frequent severe exacerbations.

CONCLUSION: Expiratory muscle weakness, reduced peak oxygen consumption, diminished exercise capacity, and pulmonary hyperinflation are independent predictors of frequent severe acute exacerbations in patients with COPD. Incorporating these parameters into routine assessments may enhance risk stratification and goal-directed therapies, and potentially reduce hospitalization rates.

KEYWORDS: Chronic obstructive pulmonary disease, flare-up, maximal respiratory pressures, oxygen consumption, walk test, functional residual capacity

Received: 26.06.2025 ***Revision Requested:*** 18.08.2025 ***Last Revision Received:*** 16.09.2025

Accepted: 15.11.2025 ***Publication Date:*** 30.01.2026

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a progressive lung condition marked by airflow limitation and respiratory symptoms such as dyspnea and cough.[1] Globally, COPD affected approximately 480 million people in 2020, with projections reaching 592 million by 2050.[2] In India, the overall prevalence is estimated at 7.4%,[3] with state-wise variations as 10% in Delhi,[4] 6.19% in Kerala.[5] Acute exacerbations of COPD (AECOPD), characterized by a sudden worsening of respiratory symptoms such as dyspnea, cough, and sputum production within a short period, typically less than two weeks, significantly contribute to hospitalizations and the healthcare burden. The Global Initiative for Chronic Obstructive Lung Disease (GOLD) classifies exacerbations by severity; with severe events requiring emergency care or hospitalization.[6] In low- and middle-income countries, including India, 20.1% of patients with COPD experience a severe exacerbation annually, and the associated healthcare cost per severe exacerbation is substantial.[7]

Corresponding author: Aqsa Mujaddadi, MPT, e-mail: aqsamuj786@gmail.com

Exacerbation frequency varies widely among COPD patients, with more frequent exacerbations associated with more rapid decline in lung function and increased mortality. While multiple factors such as older age, smoking, low body mass index (BMI), cold temperatures, air pollution, poor quality of life, comorbidities, prior exacerbations, and elevated eosinophils have been previously linked to frequent and severe exacerbations,[8] the predictive role of physiological and clinical parameters remains inadequately explored. Therefore, a significant gap remains in studies that concurrently evaluate a comprehensive set of clinical and rehabilitative parameters, namely respiratory muscle strength, peak oxygen consumption (VO_2 peak), exercise capacity, arterial blood gas (ABG), and pulmonary function parameters beyond spirometry, in relation to severe acute exacerbation frequency, while adequately adjusting for multiple established clinical covariates in COPD patients. The objective of this research was to evaluate whether physiological and clinical parameters of COPD could serve as predictors of frequent severe acute exacerbations. We hypothesized that reduced respiratory muscle strength, reduced VO_2 peak, reduced exercise capacity, impaired pulmonary function, and abnormal ABG values would be significant predictors of frequent severe acute exacerbations in patients with COPD.

MATERIAL AND METHODS

Study Design and Subjects

The observational retrospective study was conducted at Metro Centre for Respiratory Diseases, Metro Hospitals & Heart Institute, Noida, India. The medical records and discharge summaries of all COPD patients admitted to hospital between January 1st, 2018 and February 28th, 2024 were reviewed, and 265 patients with a diagnosis of severe AECOPD were analyzed. Efforts were made to ensure data accuracy by cross-checking clinical variables against multiple sources in the medical records, where available, to minimize selection bias. We included patients aged 40–80 years who were clinically identified as having COPD, confirmed by post-bronchodilator spirometry showing forced expiratory volume in 1 second/forced volume capacity (FEV_1/FVC) <0.70 and grade II, III, or IV airflow limitation according to the GOLD criteria. COPD grades are classified as moderate (GOLD stage II): FEV_1 ≥50% predicted but <80% predicted; severe (GOLD stage III): FEV_1 ≥30% predicted but <50% predicted; and very severe (GOLD stage IV): FEV_1 <30% predicted. Severe AECOPD cases in which patients were hospitalized for aggravated respiratory status of less than two weeks' duration were included in this study. It included patients with comorbidities but excluded those with inflammatory diseases (such as rheumatoid arthritis, systemic lupus erythematosus, fibromyalgia, and inflammatory bowel disease), previously diagnosed asthma or asthma-COPD overlap syndrome, myopathy, and those hospitalized for unstable angina or acute myocardial infarction.

Main Points

- The present retrospective trial highlights that expiratory muscle weakness, impaired peak oxygen uptake, impaired exercise tolerance, and pulmonary hyperinflation may independently predict severe exacerbations in chronic obstructive pulmonary disease patients.
- These markers enhance early identification of high-risk individuals and improve clinical risk stratification.
- Incorporating these predictors into routine assessments may help reduce exacerbation-related hospital admissions.

Patients were further categorized based on the frequency of severe COPD exacerbations. Severe AECOPD events were defined as exacerbations leading to hospitalization in intensive care units or medical wards. Group A, infrequent exacerbators included patients who were not rehospitalized for severe AECOPD within one year after the index event. Group B, frequent exacerbators, included patients who were rehospitalized for severe AECOPD one or more times within one year after an index event.[8] Cases with incomplete or missing data for key variables were excluded from the final analysis. Complete-case analysis was adopted to ensure the robustness of statistical comparisons.

Sample Size

A priori sample size of 102 participants was calculated using G*Power version 3.1.9.7, based on an expected odds ratio (OR) of 2.9 for maximal inspiratory muscle strength (PImax) reported in a previous study.[9] This calculation assumed a significance level (α) of 0.05 and a statistical power of 0.95, with an additional 20% to account for potential incomplete data. However, given the retrospective nature of the study and the availability of data from 265 eligible patients, the full dataset was utilized to enhance the statistical power, improve the precision of estimates, and minimize the risk of type II error.

Data Collection

The primary end points, PImax and maximum expiratory muscle strength (PEmax), were assessed at the time of hospital discharge, when patients were clinically stable, with the objective of evaluating their ability to predict frequent severe exacerbations of COPD. The secondary end points were exercise capacity [six-minute walk distance (6MWD), percent predicted 6MWD (6MWD%), six-minute walk work (6MWW), VO_2 peak], static lung volumes for pulmonary function [tidal volume (TV), inspiratory reserve volume (IRV), expiratory reserve volume (ERV), residual volume (RV), inspiratory capacity (IC), functional residual capacity (FRC), vital capacity (VC), total lung capacity (TLC), and dynamic lung volumes such as RV/TLC%, diffusing capacity of the lung for carbon monoxide (DLCO), FEV_1, FVC, FEV_1/FVC%, maximum inspiratory flow (MIF), maximal expiratory flow at 25%, 50% and 75% of FVC (MEF25%, MEF75%, MEF50%), peak inspiratory flow (PIF), peak expiratory flow (PEF)] measured were assessed at the time of hospital discharge, when patients were clinically stable and ABG [(pH, partial pressure of carbon dioxide ($PaCO_2$), partial pressure of oxygen (PaO_2), bicarbonate (HCO_3^-), oxygen saturation (SpO_2)] were obtained at hospital admission, during the acute exacerbation phase. The dependent outcome was frequent severe AECOPD episodes.

Respiratory muscle strength was assessed at the mouth using the MicroRPM manometer (Care Fusion, Hoechberg, Germany) to measure PImax and PEmax. According to Black and Hyatt's assessment,[10] the Muller maneuver is performed at RV for PImax, and the Valsalva maneuver is performed at TLC for PEmax. The best of at least three efforts is obtained following American Thoracic Society (ATS) guidelines.[10] Exercise capacity was measured via the 6MWT (Spiropalm, COSMED, Rome, Italy) with a pulse oximeter (Nonin Medical Inc., Plymouth, MN, USA), and mobile exercise testing (Vyntus™ WALK, Vyaire Medical, Hoechberg, Germany) was used to record the total distance walked according to ATS guidelines.[11] The 6MWW was calculated as 6MWD x body weight,[12] and VO_2 peak was estimated using the formula 4.948+0.023x6MWD.[13] Pulmonary function testing, including body plethysmography, was performed using the MasterScreen™ PFT (JAEGER, CareFusion, Hoechberg, Germany). ATS guidelines were followed to record lung volumes and capacities. ABG analysis was conducted using the modified Allen test.[14]

Statistical Analysis

Data were analyzed using the Statistical Package for the Social Sciences (SPSS) version 27. Data are presented as mean ± standard deviation. Demographic details, clinical characteristics, comorbidities, medications, and smoking history were compared between the two groups of severe AECOPD patients, defined by exacerbation frequency; continuous variables were compared using the Mann–Whitney U test, and categorical variables were compared using the chi-square test. The chi-square test is used to compare categorical variables, which are presented as frequencies (n) and percentages (%). Univariate binary logistic regression analysis using the enter method was used to compute OR, 95% confidence intervals (CI), and *P* values. The OR, which represents the association between an exposure and an outcome, is computed to assess the predictive ability of each variable. A *P* value <0.25 in univariate analysis indicates a significant predictor of frequent severe AECOPD.[15] Significant predictors from the univariate analysis are included in multivariate logistic regression models (enter method), adjusted for clinically relevant covariates [BMI, gender, FEV_1, non-inflammatory musculoskeletal disorders (osteoarthritis, osteoporosis, osteopenia), use of bronchodilators, and corticosteroids (inhaled and/or systemic)]. The variables which showed significance in the univariate analysis and association with exacerbations previously are chosen as covariates.[16-22] A *P* value <0.05 in the multivariate analysis indicates a statistically significant predictor.[15] Multicollinearity analysis is conducted to identify any collinearity among each predictor variable and covariates. Variance inflation factors (VIFs) are computed, and a VIF <3 suggests the absence of multicollinearity.[15]

RESULTS

Baseline Characteristics of the Participants

Throughout the study, medical records of 304 subjects were reviewed. Of these, 289 patients (95%) hospitalized for severe AECOPD met the eligibility criteria and were included in the study. Two groups were formed from the patient population; group A infrequent exacerbations comprised 162 patients (56.1%), and group B frequent exacerbations comprised 127 patients (43.9%). However, 18 patients (11.1%) in group A and 6 (4.7%) in group B had incomplete data due to missing variables and were subsequently excluded from the final analysis. Consequently, 265 patients remained for the final analysis.

The baseline profiles of the patients, including demographic and clinical characteristics, are summarized in (Table 1). Data analysis demonstrated statistically significant differences in body weight and BMI between groups categorized by frequency of severe AECOPD ($P < 0.05$) (Table 1). No other baseline variables showed statistically significant differences between the groups. The comparative analysis of key physiological and clinical parameters, namely respiratory muscle strength, exercise capacity, pulmonary function tests, and ABG values, between the frequent and infrequent COPD exacerbator groups is presented in (Table 2). Statistically significant differences ($P < 0.05$) were observed in several variables, including PEmax, 6MWD, 6MWD%pred, 6MWW, VO_2 peak, FEV_1%, FEV_1/FVC%, PEF%pred, MIF (L/s), MEF75%pred, MEF50%pred, RV, TLC%pred, RV/TLC%, and FRC%pred. Overall, respiratory muscle strength parameters (PImax and PEmax), exercise capacity indicators (6MWD, 6MWD% predicted, 6MWW, and VO_2 peak), and pulmonary function parameters (FEV_1%, FEV_1/FVC% ratio, PEF%pred, MIF (L/s), MEF75%, MEF50%, TV%pred, TLC%pred, RV/TLC%, and DLCO) were notably better in the infrequent exacerbator group than in the frequent exacerbator group.

Univariate Binary Logistic Regression Analysis

In the present research, a binary logistic regression model identified predictors of frequent, severe AECOPD. Univariate analysis revealed predictive ability for several variables, including PImax (OR: 0.991; 95% CI: 0.980-1.003; $P = 0.162$), PEmax (OR: 0.989; 95% CI: 0.981-0.996; $P = 0.004$), 6MWD (OR: 0.997; 95% CI: 0.994-0.999; $P = 0.006$), 6MWD% (OR: 0.984; 95% CI: 0.971-0.997; $P = 0.014$), 6MWW (OR: 1.000; 95% CI: 1.000-1.000; $P < 0.001$), VO_2 peak (OR: 0.861; 95% CI: 0.773-0.959; $P = 0.006$), FEV_1/FVC % (OR: 0.971; 95% CI: 0.953-0.990; $P = 0.003$), PEF% predicted (OR: 0.988; 95% CI: 0.975-1.001; $P = 0.061$), maximal expiratory flow at 75% of FVC (MEF75% predicted) (OR: 0.990; 95% CI: 0.978-1.003; $P = 0.120$), RV (OR: 1.007; 95% CI: 1.002-1.011; $P = 0.002$), and FRC (OR: 1.008; 95% CI: 1.002-1.014; $P = 0.006$) as presented (Table 3).

Multivariate Binary Logistic Regression Analysis

A multicollinearity analysis was implemented to assess the relationship between each significant predictor variable and the clinical covariates. All variables had VIF <3, confirming the absence of significant multicollinearity and thereby reinforcing the strength of the regression analysis findings. Multivariate binary logistic regression models were constructed for each significant predictor identified in the univariate analysis and were adjusted for clinically relevant covariates, which were selected based on prior literature and univariate screening, including BMI, FEV_1, gender, use of bronchodilators and corticosteroids, and presence of musculoskeletal disorders. In the adjusted model, PEmax (OR: 0.989; 95% CI: 0.980–0.998; $P = 0.014$), 6MWD (OR: 0.997; 95% CI: 0.994–1.000;

Table 1. Baseline characteristics of patients

Variables	Infrequent exacerbator group (n = 144)	Frequent exacerbator group (n = 121)	*P* value
Demographics			
Age (years)	66.31±7.07	67.09±7.41	0.340
Body weight (kg)	63.77±15.31	57.32±12.19	**<0.001***
Height (cm)	164.19±7.99	161.90±8.77	0.175
BMI (kg/m^2)	23.90±5.98	21.85±5.00	**0.003***
Gender, n (%)			
Male	115 (79.9%)	88 (72.7%)	0.172
Female	29 (20.1%)	33 (27.3%)	
Smoking status, n (%)			
Non-smoker	36 (25%)	31 (25.6%)	0.993
Active smoker	19 (13.2%)	16 (13.2%)	
Former smoker	89 (61.8%)	74 (61.2%)	
Symptoms, n (%)			
Dyspnea	143 (99.3%)	121 (100%)	0.358
Cough	111 (77.1%)	100 (82.6%)	0.263
Sputum	107 (74.3%)	97 (80.8%)	0.208
Medications, n (%)			
Use of bronchodilator	138 (95.8%)	120 (99.2%)	0.091
Use of corticosteroid	74 (51.4%)	51 (42.1%)	0.133
Use of combination drug	132 (91.7%)	109 (90.1%)	0.654
Comorbidities, n (%)			
Musculoskeletal disorder	63 (43.8%)	65 (53.7%)	0.106
Metabolic disorder	79 (54.9%)	73 (60.3%)	0.370
Pneumonia	26 (18.1%)	27 (22.3%)	0.388
Pulmonary tuberculosis	22 (15.3%)	22 (18.2%)	0.527
Sleep disorder	21 (14.6%)	22 (18.2%)	0.429
Psychological disorder	6 (4.2%)	5 (4.1%)	0.989
Inflammatory marker			
CRP (mg/dL)	73.04±84.76	78.27±96.29	0.878

*Significance considered at $P < 0.05$ (bold). Data presented in mean ± standard deviation
BMI: body mass index, CRP: C-reactive protein

P = 0.028), 6MWD% (OR: 0.985; 95% CI: 0.970–0.999; *P* = 0.041), VO_2 peak (OR: 0.874; 95% CI: 0.776–0.986; *P* = 0.028), RV (OR: 1.006; 95% CI: 1.001–1.011; *P* = 0.017), and FRC (OR: 1.008; 95% CI: 1.001–1.014; *P* = 0.028) emerged as significant independent predictors of frequent severe AECOPD, while 6MWW (OR: 1.000; 95% CI: 1.000–1.000; *P* = 0.026) showed statistical significance, however, the odds ratio of 1.000 indicates an absence of clinically meaningful association with exacerbations in both univariate and multivariate analyses as shown in (Table 4).

DISCUSSION

The primary findings of the univariate analysis revealed that: (i) respiratory muscle strength (inspiratory and expiratory muscle strength; PImax, PEmax), overall exercise capacity measures (6MWD, 6MWD%, VO_2 peak), and pulmonary function parameters (FEV_1/FVC%, PEF%pred, MEF75%pred, RV%pred, and FRC%pred) were significantly associated with frequent severe AECOPD; and (ii) in the multivariate analysis, expiratory muscle strength (PEmax), overall exercise capacity measures (6MWD, 6MWD%, VO_2 peak), and pulmonary hyperinflation (RV%pred and FRC%pred) remained significant independent predictors even after adjusting for various clinical covariates. The significance of this study is rooted in its emphasis on the interplay between clinical and physiological parameters as potential predictors of frequent severe exacerbations in COPD. This approach underscores the value of incorporating a comprehensive assessment of physiological dysfunction, moving beyond traditional risk factors. The integration of these diverse markers enhances our ability to stratify exacerbation risk based on key functional impairments.

Table 2. Comparison of clinical characteristics such as arterial blood gas, respiratory muscle strength, exercise capacity and pulmonary function parameters

Variables	Infrequent exacerbator group (n = 144)	Frequent exacerbator group (n = 121)	*P* value
Arterial blood gas			
pH	7.39±0.07	7.40±0.06	0.666
$PaCO_2$ (mmHg)	45.62±15.47	45.39±13.70	0.773
PaO_2 (mmHg)	84.75±36.68	84.58±32.07	0.914
HCO_3^- (mEq/L)	26.60±6.37	27.29±5.69	0.328
SpO_2 (%)	93.51±7.82	93.85±6.23	0.791
Respiratory muscle strength			
PImax (cm H_2O)	65.94±21.33	62.42±19.19	0.217
PEmax (cm H_2O)	110.06±34.70	98.26±29.60	**0.003***
Exercise capacity			
6MWD (meters)	291.79±128.00	253.64±84.95	**0.006***
6MWD% (% predicted)	58.09±20.68	52.27±16.27	**0.015***
6MWW (kg-meter)	18973.79±11816.80	14531.42±5877.23	**<0.001***
VO_2 peak (mL/kg/min)	11.65±2.94	10.78±1.95	**0.006***
Pulmonary function tests			
FVC (% predicted)	72.28±15.70	74.72±19.31	0.471
FEV_1 (% predicted)	43.29±16.11	39.79±16.97	**0.027***
FEV_1/FVC% (%)	47.62±13.08	42.60±13.15	**0.002***
PIF (L/s)	3.27±1.27	3.38±1.32	0.522
PEF (% predicted)	47.99±19.24	43.45±19.45	**0.016***
MIF (L/s)	1.20±3.76	1.04±3.86	**0.031***
MEF75% (% predicted)	22.65±19.71	18.61±21.78	**0.001***
MEF50% (% predicted)	15.24±14.76	15.09±29.70	**0.006***
MEF25% (% predicted)	13.97±10.62	13.97±15.79	0.269
TV (% predicted)	154.45±61.53	147.24±54.65	0.357
IRV (L)	0.77±0.37	1.14±3.50	0.552
ERV (% predicted)	91.76±39.35	92.44±35.12	0.799
IC (% predicted)	61.88±17.35	62.21±17.96	0.962
VCmax (% predicted)	70.64±14.46	73.12±20.47	0.675
RV (% predicted)	165.82±56.02	191.22±69.94	**<0.001***
TLC (% predicted)	149.77±31.56	112.37±25.07	**<0.001***
RV/TLC% (%)	66.72±21.79	66.39±9.73	**0.049***
FRC (% predicted)	142.46±42.70	158.60±47.56	**<0.001***
DLCO (% predicted)	51.92±17.71	50.26±19.06	0.215

Data presented in mean ± standard deviation. *Significance considered at $P < 0.05$

PImax: maximum inspiratory pressure, PEmax: maximum expiratory pressure, 6MWD: six-minute walk distance, 6MWD%: percent predicted six-minute walk distance, 6MWW: six-minute walk work, VO_2 peak: peak oxygen uptake, FVC: forced vital capacity, FEV_1: forced expiratory volume in 1 second, PIF: peak inspiratory flow, PEF: peak expiratory flow, MIF: maximal inspiratory flow, MEF75%, MEF50%, MEF25%: maximum expiratory flow at 75%, 50%, 25% of FVC, respectively, TV: tidal volume, IRV: inspiratory reserve volume, ERV: expiratory reserve volume, IC: inspiratory capacity, VCmax: maximum vital capacity, RV: residual volume, TLC: total lung capacity, FRC: functional residual capacity, RV/TLC%: residual volume/total lung capacity ratio, DLCO: diffusing capacity of the lungs for carbon monoxide

In present study, although both inspiratory and expiratory muscle weakness were significantly associated with frequent severe exacerbations in the univariate analyses, only expiratory muscle weakness retained significance as an independent predictor of frequent severe exacerbations in COPD after adjustment for clinical covariates. Thus, our study suggests that greater expiratory muscle strength is associated with lower odds of frequent, severe exacerbations in COPD. Though a prior study,[23] linked % predicted PImax with risk of exacerbation using a time-course Cox proportional hazards model, their work did not focus specifically on the frequency of severe exacerbations. In contrast, our findings offer a distinct and clinically meaningful perspective by identifying PEmax as a stronger and more practical predictor of frequent, severe, exacerbation-related hospitalizations. This

Table 3. Univariate binary logistic regression analyses

Variables	OR	95% CI	*P* value
Predictor variables			
Arterial blood gas analyses			
pH	2.151	0.069-6.868	0.662
$PaCO_2$ (mmHg)	0.999	0.983–1.016	0.896
PaO_2 (mmHg)	1.000	0.993–1.007	0.968
HCO_3^- (mEq/L)	1.019	0.979–1.061	0.355
SpO_2 (%)	1.007	0.973–1.042	0.699
Respiratory muscle strength			
PImax (cm H_2O)	0.991	0.980–1.003	**0.162***
PEmax (cm H_2O)	0.989	0.981–0.996	**0.004***
Exercise capacity			
6MWD (meters)	0.997	0.994–0.999	**0.006***
6MWD (% predicted)	0.984	0.971–0.997	**0.014***
6MWW (kg meters)	1.000	1.000–1.000	**<0.001***
VO_2 peak (mL/kg/min)	0.861	0.773–0.959	**0.006***
Pulmonary function tests			
FVC (L)	1.008	0.994–1.022	0.259
FEV_1/FVC (%)	0.971	0.953–0.990	**0.003***
PIF (L/s)	1.066	0.884–1.285	0.506
PEF (% predicted)	0.988	0.975–1.001	**0.061***
MIF (L/s)	0.989	0.925–1.057	0.736
MEF75 (% predicted)	0.990	0.978–1.003	**0.120***
MEF50 (% predicted)	1.000	0.989–1.010	0.955
MEF25 (% predicted)	1.000	0.982–1.018	0.999
TV (% predicted)	0.998	0.994–1.002	0.317
IRV (L)	1.353	0.760–2.409	0.305
ERV (% predicted)	1.000	0.994–1.007	0.884
IC (% predicted)	1.001	0.987–1.015	0.879
VCmax (% predicted)	1.008	0.994–1.022	0.252
RV (% predicted)	1.007	1.002–1.011	**0.002***
TLC (% predicted)	0.999	0.996–1.001	0.308
RV/TLC (%)	0.999	0.985–1.013	0.879
FRC (% predicted)	1.008	1.002–1.014	**0.006***
DLCO (% predicted)	0.995	0.982–1.008	0.464
Covariates			
Age (years)	1.015	0.982–1.050	0.378
BMI (kg/m^2)	0.933	0.889–0.978	**0.004***
CRP (mg/dL)	1.001	0.998–1.003	0.637
FEV_1 (%)	0.987	0.972–1.002	**0.088***
Gender			
Female (ref)			
Male	0.672	0.380–1.190	**0.173***
Smoking status			
Non–smoker (ref)			
Active smoker	0.978	0.431–2.221	0.957
Former smoker	0.966	0.546–1.709	0.904
Medications (ref: no use)			
Use of Bronchodilator	5.217	0.619–43.952	**0.129***
Use of Corticosteroid	0.689	0.424–1.121	**0.134***
Use of Combination drug	0.826	0.357–1.912	0.655
Comorbidities (ref: absent)			
Musculoskeletal disorder	1.492	0.918–2.426	**0.106***
Metabolic disorder	1.251	0.766–2.043	0.370
Pneumonia disorder	1.304	0.713–2.382	0.389
Pulmonary tuberculosis disorder	1.232	0.645–2.355	0.527
Sleep disorder	1.302	0.677–2.503	0.429
Psychological disorder	0.991	0.295–3.332	0.989

*Statistically significant considered at $P < 0.25$ (bold)

OR: odds ratio, CI: confidence interval, pH: potential of hydrogen, $PaCO_2$: partial pressure of carbon dioxide, PaO_2: partial pressure of oxygen, HCO_3^-: bicarbonate, SpO_2: peripheral capillary oxygen saturation, PImax: maximal inspiratory pressure, PEmax: maximal expiratory pressure, 6MWD: six-minute walk distance, 6MWW: six-minute walk work, VO_2 peak: peak oxygen uptake, FVC: forced vital capacity, FEV_1: forced expiratory volume in 1 second, PIF: peak inspiratory flow, PEF: peak expiratory flow, MIF: maximal inspiratory flow, MEF75%, 50%, 25%: maximal expiratory flow at 75%, 50%, and 25% of FVC respectively, TV: tidal volume, IRV: inspiratory reserve volume, ERV: expiratory reserve volume, IC: inspiratory capacity, VCmax: maximal voluntary capacity, RV: residual volume, TLC: total lung capacity, RV/TLC%: percentage of residual volume to total lung capacity ratio, FRC: functional residual capacity, DLCO: diffusing capacity of the lung for carbon monoxide, BMI: body mass index, CRP: C-reactive protein

discrepancy may stem from differences in demographics, environment, study region, statistical methods, and the severity or frequency of exacerbations considered. Our study included older, hospitalized patients with advanced COPD and multiple comorbidities, who were mostly former smokers, unlike the study,[23] which involved a population with a milder disease stage and a mix of moderate and severe exacerbations.

The pathophysiological basis for our findings is well-supported. Expiratory muscles are essential for effective coughing and secretion clearance but are highly susceptible to dysfunction due to chronic mechanical loading, increased airway resistance, and reduced lung elastic recoil.[24] These factors lead to progressive muscle fatigue, reduced PEmax, impaired secretion clearance, and heightened airway inflammation, all of which increase the likelihood of severe exacerbations.[25] These effects are further exacerbated in older patients due to age-related declines in muscle strength, lung compliance, and chest wall mobility.[26] Although PImax did not emerge as a strong predictor in multivariate analysis, it was associated with the frequency of severe exacerbations in univariate analysis. Hence, in clinical practice, this suggests that greater expiratory muscle strength, particularly PEmax, may contribute to reducing the frequency of severe exacerbations and may be considered in rehabilitative strategies for COPD patients. Our findings support the clinical value of incorporating respiratory muscle assessment into COPD risk stratification and highlight the potential benefits of targeted expiratory muscle training to reduce the frequency of severe exacerbations and hospitalizations.[27]

Table 4. Multivariate binary logistic regression analyses (adjusted model)

Predictor variables	OR	95% CI	*P* value
Model 1			
PImax (cm H_2O)	0.993	0.980–1.006	0.304
BMI (kg/m²)	0.923	0.876–0.972	**0.002***
FEV_1 (%)	0.990	0.974–1.006	0.203
Gender	0.501	0.263–0.952	**0.035***
Use of bronchodilator	7.791	0.843–71.972	0.070
Use of corticosteroid	0.537	0.314–0.918	**0.023***
Musculoskeletal disorder	1.247	0.748–2.082	0.397
Model 2			
PEmax (cm H_2O)	0.989	0.980–0.998	**0.014***
BMI (kg/m²)	0.929	0.882–0.979	**0.006***
FEV_1 (%)	0.988	0.972–1.004	0.143
Gender	0.572	0.297–1.101	0.095
Use of bronchodilator	8.206	0.889–75.717	0.063
Use of corticosteroid	0.494	0.287–0.851	**0.011***
Musculoskeletal disorder	1.229	0.733–1.061	0.434
Model 3			
6MWD (m)	0.997	0.994–1.000	**0.028***
BMI (kg/m²)	0.920	0.872–0.969	**0.002***
FEV_1 (%)	0.991	0.975–1.007	0.279
Gender	0.594	0.307–1.152	0.124
Use of bronchodilator	6.655	0.731–60.558	0.093
Use of corticosteroid	0.600	0.355–1.017	0.058
Musculoskeletal disorder	1.338	0.796–2.250	0.272
Model 4			
6MWD% (%)	0.985	0.970–0.999	**0.041***
BMI (kg/m²)	0.993	0.977–1.009	**<0.001***
FEV_1 (%)	0.914	0.867–0.963	0.371
Gender	0.561	0.291–1.081	0.084
Use of bronchodilator	6.936	0.756–63.628	0.087
Use of corticosteroid	0.619	0.365–1.050	0.075
Musculoskeletal disorder	1.296	0.771–2.177	0.328
Model 5			
6MWW (kg-m)	1.000	1.000–1.000	**0.026***
BMI (kg/m²)	0.946	0.895–0.999	**0.046***
FEV_1 (%)	0.991	0.975–1.007	0.271
Gender	0.660	0.333–1.306	0.233
Use of bronchodilator	6.818	0.746–62.307	0.089
Use of corticosteroid	0.604	0.356–1.024	0.061
Musculoskeletal disorder	1.298	0.744–2.178	0.323
Model 6			
VO_2 peak (mL/kg/min)	0.874	0.776–0.986	**0.028***
BMI (kg/m²)	0.920	0.872–0.969	**0.002***
FEV_1 (%)	0.991	0.975–1.007	0.279
Gender	0.594	0.307–1.152	0.124
Use of bronchodilator	6.655	0.731–60.558	0.093
Use of corticosteroid	0.600	0.355–1.017	0.058
Musculoskeletal disorder	1.338	0.796–2.250	0.272

Table 4. Multivariate binary logistic regression analyses (adjusted model)

Predictor variables	OR	95% CI	*P* value
Model 7			
FEV_1/FVC (%)	0.970	0.940–1.000	0.053
BMI (kg/m²)	0.931	0.883–0.982	**0.008***
FEV_1 (%)	1.006	0.983–1.030	0.612
Gender	0.452	0.236–0.866	**0.017***
Use of bronchodilator	7.344	0.787–68.485	0.080
Use of corticosteroid	0.586	0.347–0.990	**0.046***
Musculoskeletal disorder	1.177	0.701–1.975	0.537
Model 8			
PEF (% predicted)	0.997	0.976–1.019	0.782
BMI (kg/m²)	0.921	0.874–0.970	**0.002***
FEV_1 (%)	0.991	0.967–1.017	0.498
Gender	0.489	0.256–0.935	**0.030***
Use of bronchodilator	7.780	0.846–71.581	0.070
Use of corticosteroid	0.573	0.341–0.965	**0.036***
Musculoskeletal disorder	1.238	0.742–2.065	0.413
Model 9			
MEF75% (% predicted)	0.998	0.979–1.018	0.862
BMI (kg/m²)	0.919	0.872–0.968	**0.001***
FEV_1 (%)	0.990	0.966–1.014	0.401
Gender	0.454	0.237–0.868	**0.001***
Use of bronchodilator	8.239	0.886–76.615	0.064
Use of corticosteroid	0.558	0.330–0.941	**0.029***
Musculoskeletal disorder	1.260	0.753–2.109	0.379
Model 10			
RV (% predicted)	1.006	1.001–1.011	**0.017***
BMI (kg/m²)	0.920	0.874–0.970	**0.002***
FEV_1 (%)	0.998	0.981–1.016	0.836
Gender	0.494	0.259–0.944	**0.033***
Use of bronchodilator	9.760	0.992–95.982	0.051
Use of corticosteroid	0.613	0.362–1.039	0.069
Musculoskeletal disorder	1.143	0.680–1.922	0.615
Model 11			
FRC (% predicted)	1.008	1.001–1.014	**0.028***
BMI (kg/m²)	0.920	0.873–0.969	**0.002***
FEV_1 (%)	0.997	0.980–1.014	0.725
Gender	0.474	0.248–0.905	**0.024***
Use of bronchodilator	10.359	1.058–101.408	**0.045***
Use of corticosteroid	0.598	0.354–1.011	0.055
Musculoskeletal disorder	1.143	0.680–1.921	0.614

*Statistically significant considered at $P < 0.05$ (bold)

OR: adjusted odds ratio, CI: confidence interval, PImax: maximal inspiratory pressure, PEmax: maximal expiratory pressure, 6MWD: six-minute walk distance, 6MWD%: six-minute walk distance percent predicted, 6MWW: six-minute walk work, VO_2 peak: peak oxygen uptake, FEV_1/FVC%: ratio of forced expiratory volume in 1 second to forced vital capacity in percent, FEV_1: forced expiratory volume in 1 second, PEF: peak expiratory flow, MEF75%: maximal expiratory flow at 75% of FVC, RV: residual volume, FRC: functional residual capacity, BMI: body mass index

This study highlights peak VO_2 and exercise capacity as prognostic indicators of frequent severe exacerbations in COPD. We found that 6MWD, 6MWD%, and VO_2 peak were significant independent predictors of exacerbation risk, whereas 6MWW showed no association. Higher 6MWD values were linked to lower odds of exacerbations, consistent with previous research showing that reduced 6MWD predicts an increased risk of hospitalization and mortality.[28,29] Our study further confirms that even a single severe exacerbation can cause lasting reductions in 6MWD. To the best of our knowledge, this is the first study to explore the predictive value of variables derived from the six-minute walk test, such as 6MWW and VO_2 peak. VO_2 peak was a meaningful predictor of exacerbation risk than 6MWW; higher VO_2 peak was associated with lower exacerbation risk, underscoring the importance of oxygen uptake efficiency in assessing patients with COPD. Since actual VO_2 peak is determined by direct measurement, this study suggests that indirectly estimated VO_2 peak may also serve as a practical surrogate, given its demonstrated predictive value, facilitating its application in clinical rehabilitation and professional practice.

In the present study, FEV_1/FVC%, PEF%pred, MEF75%, RV%pred, and FRC%pred were significantly associated with frequent severe exacerbations in the univariate analysis. However, markers of air trapping, elevated RV%pred and FRC%pred -remained as significant independent predictors of frequent severe exacerbations in COPD after adjustment for clinical covariates. Previous studies have proposed FEV_1/FVC% and PEF as clinical markers, but their clinical utility has been limited by the lack of reported OR.[30,31] A study[32] addressed this, showing that FEV_1/FVC% (OR: 0.994) and PEF (OR: 0.891) were associated with exacerbation risk in univariate analysis, although PEF lost significance in multivariate analysis. Similarly, our study is consistent with previous findings,[32] showing that FEV_1/FVC% (OR: 0.97) and PEF (OR: 0.98) were significant only in univariate models.

Together, these findings emphasize the limited predictive power of spirometric indices and reinforce the need for alternative markers, such as pulmonary hyperinflation. Hyperinflation was defined as RV >120% predicted and/or FRC >120% predicted, according to established guidelines.[33-35] Notably, RV%pred and FRC%pred remained significant predictors in our study despite bronchodilator and corticosteroid treatment, reinforcing the independent role of hyperinflation. Our findings align with prior research showing that a 10% increase in RV/TLC% is associated with a 36% rise in exacerbation risk, particularly in the absence of triple inhaler therapy.[33] While earlier studies mainly used Cox proportional hazards models, static hyperinflation, measured by IC/TLC% and RV/TLC%, has consistently predicted mortality, exercise capacity, and quality of life.[34,35] By quantifying these associations using logistic regression, our study provides further evidence that lung-volume measures reflecting hyperinflation are more robust and consistent predictors of frequent severe exacerbations in COPD.

Study Limitations

This research employed robust statistical methods, including both univariate and multivariate analyses. Furthermore, the study confirmed the absence of multicollinearity among predictors and covariates, thereby strengthening the statistical validity and robustness of the model. These findings support the development of multifactorial predictive models that include not only clinical variables but also physiological indicators, facilitating more targeted and individualized management strategies. This study has several limitations. First, a retrospective study design may introduce inherent bias in data collection and analysis. Second, it was a single-center study. Third, no specific cut-off values for predictors were established, limiting clinical applicability.

CONCLUSION

The findings of the present study extend the concept that susceptibility to extrapulmonary manifestations is related to frequent severe AECOPD. Our results indicate that lower PEmax, reduced VO_2 peak, diminished exercise capacity (6MWD, 6MWD%), and greater pulmonary hyperinflation (RV and FRC) are independent predictors of frequent severe AECOPD. However, variables such as PImax, FEV_1/FVC%, PEF%pred, and MEF75%pred were predictive only in univariate analyses. These findings suggest that incorporating them into routine assessment may prompt recognition of high-risk patients and the stratification of individuals with elevated risk of frequent severe exacerbations, potentially leading to reduced hospitalization rates. Future multicenter, prospective, and longitudinal studies are recommended to confirm these findings, to better understand causal relationships, extrapulmonary influences, and to improve risk stratification in COPD.

Ethics

Ethics Committee Approval: 1) Metro Ethical Review Board, Metro Hospitals and Heart Institute, Noida, India (ECR/335/Inst/UP/2013/RR-20); approval number: 93/MERB/2024; date of approval: 23rd January, 2025. 2) Institutional Ethics Committee, Jamia Millia Islamia EC/NEW/INST/2022/DL/0170); approval number: 3/10/523/JMI/IEC/2024; date of approval: 12th March, 2025.

Informed Consent: The study involved a retrospective analysis of existing medical records; therefore, the requirement for informed consent was waived. Patient confidentiality was strictly maintained, and all data were anonymized prior to analysis.

Acknowledgments

The authors would like to express their sincere gratitude to the staffs of the Pulmonary Rehabilitation Department and the Medical Records Department at Metro Hospitals and Heart Institute, Noida, Uttar Pradesh, for their invaluable administrative support and assistance in facilitating the smooth execution of the research process and retrieving data of patients. The authors also gratefully acknowledge the Centre for Physiotherapy and Rehabilitation Sciences, Jamia Millia Islamia, New Delhi, for their academic support and institutional guidance throughout the study.

Footnotes

The patient's data was anonymized and study was listed under Clinical Trial Registry-India (CTRI), CTRI/2025/01/079833.

Authorship Contributions

Concept: M.M.T., A.M., O.A., D.T., Design: M.M.T., A.M., O.A., D.T., Data Collection or Processing: M.M.T., A.M., O.A., D.T., Analysis or Interpretation: M.M.T., A.M., Literature Search: M.M.T., A.M., Writing: M.M.T., A.M., O.A., D.T.

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. Agustí A, Celli BR, Criner GJ, et al. Global Initiative for Chronic Obstructive Lung Disease 2023 report: GOLD executive summary. *Eur Respir J.* 2023;61(4):2300239. [Crossref]
2. Boers E, Barrett M, Su JG, et al. Global burden of chronic obstructive pulmonary disease through 2050. *JAMA Netw Open.* 2023;6(12):e2346598. [Crossref]
3. Daniel RA, Aggarwal P, Kalaivani M, Gupta SK. Prevalence of chronic obstructive pulmonary disease in India: a systematic review and meta-analysis. *Lung India.* 2021;38(6):506-513. [Crossref]
4. Sinha B; Vibha; Singla R, Chowdhury R. An epidemiological profile of chronic obstructive pulmonary disease: a community-based study in Delhi. *J Postgrad Med.* 2017;63(1):29-35. [Crossref]
5. Viswanathan K, Rakesh PS, Balakrishnan S, Shanavas A, Dharman V. Prevalence of chronic respiratory diseases from a rural area in Kerala, southern India. *Indian J Tuberc.* 2018;65(1):48-51. [Crossref]
6. Celli BR, Fabbri LM, Aaron SD, et al. An updated definition and severity classification of chronic obstructive pulmonary disease exacerbations: the Rome proposal. *Am J Respir Crit Care Med.* 2021;204(11):1251-1258. [Crossref]
7. Athanazio RA, Villada LB, Avdeev SN, et al. Rate of severe exacerbations, healthcare resource utilisation and clinical outcomes in patients with COPD in low-income and middle-income countries: results from the EXACOS international study. *BMJ Open Respir Res.* 2024;11(1):e002101. [Crossref]
8. Husebø GR, Bakke PS, Aanerud M, et al. Predictors of exacerbations in chronic obstructive pulmonary disease--results from the Bergen COPD cohort study. *PLoS One.* 2014;9(10):e109721. [Crossref]
9. Vilaró J, Ramirez-Sarmiento A, Martínez-Llorens JM, et al. Global muscle dysfunction as a risk factor of readmission to hospital due to COPD exacerbations. *Respir Med.* 2010;104(12):1896-1902. [Crossref]
10. American Thoracic Society. ATS/ERS statement on respiratory muscle testing. *Am J Respir Crit Care Med.* 2002;166(4):518-624. [Crossref]
11. 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]
12. Carter R, Holiday DB, Nwasuruba C, Stocks J, Grothues C, Tiep B. 6-minute walk work for assessment of functional capacity in patients with COPD. *Chest.* 2003;123(5):1408-1415. [Crossref]
13. Ross RM, Murthy JN, Wollak ID, Jackson AS. The six minute walk test accurately estimates mean peak oxygen uptake. *BMC Pulm Med.* 2010;10:31. [Crossref]
14. Castro D, Patil SM, Zubair M, Keenaghan M. Arterial Blood Gas. 2024 Jan 8. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025. [Crossref]
15. Grant SW, Hickey GL, Head SJ. Statistical primer: multivariable regression considerations and pitfalls. *Eur J Cardiothorac Surg.* 2019;55(2):179-185. [Crossref]
16. Lee J, Jung HM, Kim SK, et al. Factors associated with chronic obstructive pulmonary disease exacerbation, based on big data analysis. *Sci Rep.* 2019;9(1):6679. [Crossref]
17. Wu Z, Yang D, Ge Z, Yan M, Wu N, Liu Y. Body mass index of patients with chronic obstructive pulmonary disease is associated with pulmonary function and exacerbations: a retrospective real world research. *J Thorac Dis.* 2018;10(8):5086-5099. [Crossref]
18. Westerik JA, Metting EI, van Boven JF, Tiersma W, Kocks JW, Schermer TR. Associations between chronic comorbidity and exacerbation risk in primary care patients with COPD. *Respir Res.* 2017;18(1):31. [Crossref]
19. Kim J, Kim K, Kim Y, et al. The association between inhaled long-acting bronchodilators and less in-hospital care in newly-diagnosed COPD patients. *Respir Med.* 2014;108(1):153-161. [Crossref]
20. Lindenauer PK, Pekow PS, Lahti MC, Lee Y, Benjamin EM, Rothberg MB. Association of corticosteroid dose and route of administration with risk of treatment failure in acute exacerbation of chronic obstructive pulmonary disease. *JAMA.* 2010;303(23):2359-2367. [Crossref]
21. Cazzola M, Hanania NA. The role of combination therapy with corticosteroids and long-acting β2-agonists in the prevention of exacerbations in COPD. *Int J Chron Obstruct Pulmon Dis.* 2006;1(4):345-354. [Crossref]
22. Su B, Liu T, Fan H, et al. Inflammatory markers and the risk of chronic obstructive pulmonary disease: a systematic review and meta-analysis. *PloS One.* 2016;11(4):e0150586. [Crossref]
23. Furukawa Y, Miyamoto A, Asai K, et al. Respiratory muscle strength as a predictor of exacerbations in patients with chronic obstructive pulmonary disease. *Respirology.* 2025;30(5):408-416. [Crossref]
24. Ramírez-Sarmiento A, Orozco-Levi M, Barreiro E, et al. Expiratory muscle endurance in chronic obstructive pulmonary disease. *Thorax.* 2002;57(2):132-136. [Crossref]
25. Gosselink R, Troosters T, Decramer M. Distribution of muscle weakness in patients with stable chronic obstructive pulmonary disease. *J Cardiopulm Rehabil.* 2000;20(6):353-360. [Crossref]
26. O'Donnell DE, Sanii R, Anthonisen NR, Younes M. Expiratory resistive loading in patients with severe chronic air-flow limitation. An evaluation of ventilatory mechanics and compensatory responses. *Am Rev Respir Dis.* 1987;136(1):102-107. [Crossref]
27. Kim J, Sapienza CM. Implications of expiratory muscle strength training for rehabilitation of the elderly: Tutorial. *J Rehabil Res Dev.* 2005;42(2):211-224. [Crossref]
28. Spruit MA, Polkey MI, Celli B, et al. Predicting outcomes from 6-minute walk distance in chronic obstructive pulmonary disease. *J Am Med Dir Assoc.* 2012;13(3):291-297. [Crossref]
29. Andrianopoulos V, Wouters EF, Pinto-Plata VM, et al. Prognostic value of variables derived from the six-minute walk test in patients with COPD: results from the ECLIPSE study. *Respir Med.* 2015;109(9):1138-1146. [Crossref]
30. Jang JG, Kim Y, Shin SH, et al. The role of FEV_1/FVC in the prediction of acute exacerbation of COPD. *Respir Med.* 2024;234:107780. [Crossref]

31. Cen J, Ma H, Chen Z, Weng L, Deng Z. Monitoring peak expiratory flow could predict COPD exacerbations: A prospective observational study. *Respir Med*. 2019;148:43-48. **[Crossref]**

32. Song Q, Lin L, Cheng W, et al. Clinical–functional characteristics and risk of exacerbation and mortality among more symptomatic patients with chronic obstructive pulmonary disease: a retrospective cohort study. *BMJ Open*. 2023;13(3):e065625. **[Crossref]**

33. Kim Y, Kim SH, Rhee CK, et al. Air trapping and the risk of COPD exacerbation: analysis from prospective KOCOSS Cohort. *Front Med*. 2022;9:835069. **[Crossref]**

34. Shin TR, Oh YM, Park JH, et al. The prognostic value of residual volume/total lung capacity in patients with chronic obstructive pulmonary disease. *J Korean Med Sci*. 2015;30(10):1459-1465. **[Crossref]**

35. Nishimura K, Yasui M, Nishimura T, Oga T. Airflow limitation or static hyperinflation: which is more closely related to dyspnea with activities of daily living in patients with COPD? *Respir Res*. 2011;12(1):135. **[Crossref]**

Thorac Res Pract. 2026;27(1):21-29 **DOI:** 10.4274/ThoracResPract.2025.2025-7-9

Original Article

How Much Do We Know About Acute Kidney Injury Following Pneumonectomy? A Retrospective Study

Çiğdem Yıldırım Güçlü[1], Süheyla Karadağ Erkoç[1], Bengi Şafak[1], Yusuf Kahya[2], Süleyman Gökalp Güneş[2], Başak Ceyda Meco[1,3]

[1]Department of Anaesthesiology and Intensive Care Medicine, Ankara University Faculty of Medicine, Ankara, Türkiye
[2]Department of Thoracic Surgery, Ankara University Faculty of Medicine, Ankara, Türkiye
[3]Ankara University Brain Research Center (AUBAUM), Ankara University Faculty of Medicine, Ankara, Türkiye

Cite this article as: Yıldırım Güçlü Ç, Karadağ Erkoç S, Şafak B, Kahya Y, Güneş SG, Meco BC. How much do we know about acute kidney injury following pneumonectomy? A retrospective study. *Thorac Res Pract.* 2026;27(1):21-29

ABSTRACT

OBJECTIVE: Acute kidney injury (AKI) is a significant postoperative complication of thoracic surgery, but data on AKI after pneumonectomy remain scarce. This study aimed to determine the incidence, risk factors, and short-term outcomes of AKI, as defined by Kidney Disease Improving Global Outcomes 2012 criteria, occurring within one week after pneumonectomy.

MATERIAL AND METHODS: This retrospective single-center cohort included adults who underwent elective pneumonectomy between 2008–2018. Patients with preoperative chronic kidney disease or AKI, or with missing data, were excluded. Demographic, perioperative, and postoperative data were collected from hospital records. AKI was identified based on postoperative creatinine values measured within one week. Univariate and multivariate logistic regression analyses were performed to identify independent risk factors.

RESULTS: Of the 308 patients, 166 met the inclusion criteria. The incidence of AKI was 12% (19 stage 1, 1 stage 2); none required renal replacement therapy. In multivariate analysis, increased body mass index [odds ratio (OR): 1.10, 95% confidence interval (CI): 1.01–1.21, $P = 0.038$]; acetylsalicylic acid use (OR: 10.56, 95% CI: 1.58–70.60, $P = 0.015$); higher intraoperative fluid volume (OR: 1.00, 95% CI: 1.00–1.00, $P = 0.036$); and length of stay (OR: 1.07, 95% CI: 1.01–1.13, $P = 0.016$) were associated with increased AKI risk, while nonsteroidal anti-inflammatory drug use was independently protective (OR: 0.03, 95% CI: 0.00–0.13, $P < 0.001$), as was diuretic use (OR: 0.06, 95% CI: 0.01–0.50, $P = 0.009$). AKI was associated with longer hospitalization but not with increased mortality.

CONCLUSION: Reducing the incidence of AKI may improve patient outcomes, and AKI should be considered a key quality indicator in thoracic surgery. Identifying and understanding the risk factors for AKI may provide the foundation for predictive models and guide strategies to prevent this complication.

KEYWORDS: Acute kidney injury, pneumonectomy, anesthesia

Received: 22.07.2025 ***Revision Requested:*** 23.09.2025 ***Last Revision Received:*** 30.09.2025

Accepted: 01.12.2025 ***Publication Date:*** 30.01.2026

INTRODUCTION

Acute kidney injury (AKI) is a common and serious complication associated with substantial morbidity and mortality after major surgical procedures.[1] Thoracic surgery carries the third-highest risk of postoperative AKI, after cardiac and general surgery.[2] Patient characteristics, medical history, and surgical and anesthetic factors were identified as risk factors for AKI after thoracic surgery.[3,4] The incidence of AKI was 12.1% after bilobectomy or pneumonectomy and was associated with increased 30-day mortality and longer hospital stay.[3] Furthermore, AKI was associated with significantly higher rates of tracheal reintubation and postoperative mechanical ventilation following lung resection surgery.[5] The identification of patient and procedural risk factors may improve long-term outcomes by allowing for targeted perioperative management.

Corresponding author: Bengi Şafak, MD, e-mail: bengisafak@hotmail.com

This study was presented as a poster at the 18th World Congress of Anaesthesiologists (2024).

The main objective of our work was to determine the incidence and risk factors for postoperative AKI within the first 7 days after pneumonectomy, defined according to the serum creatinine (sCr) criteria of the Kidney Disease Improving Global Outcomes (KDIGO) 2012 guideline.[6] The secondary aims included assessment of postoperative length of stay in the intensive care unit and in hospital, complications, and mortality at 30 days, 90 days, and 1 year. We present this article in accordance with the STROBE reporting checklist.

MATERIAL AND METHODS

This retrospective, observational, single-center cohort study was conducted at a university hospital. The Ethics Committee of Ankara University Faculty of Medicine approval number is I8-568-21, date: 15/10/2021, and the clinical trial number is NCT05672238. The requirement to obtain informed consent was waived because the study was retrospective. Patients aged over 18 years who underwent elective pneumonectomy for primary lung cancer at Ankara University İbni Sina Hospital between January 2008 and December 2018 were included. Patients with chronic kidney disease, preoperative AKI, missing data, or who underwent pneumonectomy for trauma or other indications were excluded.

Patient demographics, American Society of Anesthesiologist (ASA) status, comorbidities, and intraoperative variables [surgery and anesthesia duration, fluid and red blood cell (RBC) administration, hemodynamic events, urine output, and furosemide use] were collected. The Charlson comorbidity index (CCI) and the estimated 10-year survival were calculated using patients' preoperative information.[7] Postoperative data included fluid therapy during the first 24 hours, sCr levels, RBC transfusions, medications, and complications occurring within 7 days. All data were collected from patients' records. Non-AKI complications were classified using the Clavien-Dindo classification (CDC),[8] and mortality was assessed at 30 days, 90 days, and 1 year.

The baseline sCr was defined as the most recent measurement obtained within 7 days before surgery. Patients without a sCr measurement during this period were excluded. All available sCr values measured within the first postoperative week were evaluated. Blood sampling, performed as part of routine clinical follow-up, was not conducted at fixed intervals, resulting in variable sampling frequency among patients. AKI was defined and staged according to the KDIGO criteria.[6] Since urinary catheters were not routinely used at our center during the one-week postoperative period, AKI was defined based on sCr levels. The maximum AKI stage was calculated for patients who developed AKI.

Main Points

- Acute kidney injury (AKI) developed in 12% of patients following pneumonectomy in this single-center, 10-year retrospective cohort study.
- Multivariate analysis identified higher body mass index, greater intraoperative fluid administration, and preoperative acetylsalicylic acid use as independent risk factors for AKI development.
- AKI was associated with a significantly longer hospital stay but not with increased short-term mortality.
- Recognizing these risk factors underscores the need for individualized perioperative management and targeted preventive strategies in patients undergoing pneumonectomy.

Statistical Analysis

Statistical analyses were performed using SPSS version 11.5. Quantitative variables were summarized using mean ± standard deviation and median (minimum–maximum), while qualitative variables were presented as counts and percentages. The normality of quantitative variables was assessed using the Shapiro-Wilk test. Differences between two groups were assessed using the Student's t-test or Mann-Whitney U test, as appropriate. The chi-square test or Fisher's exact test was used to evaluate associations between qualitative variables. Univariate and multivariate logistic regression analyses were performed to identify independent risk factors for AKI development. Variables with a P value <0.10 in univariate analysis, as well as those deemed clinically relevant, were included in the multivariate model. Odds ratios (OR) and 95% confidence intervals (CI) were reported. A P value <0.05 was considered statistically significant.

RESULTS

From January 2008 to December 2018, 308 patients underwent pneumonectomy at our center. Seven patients with preoperative chronic kidney disease and 135 patients with missing data were excluded from the statistical analysis (Figure 1). AKI was observed in 20 patients (12%); 19 patients (95.0%) were classified as stage 1 and one patient (5.0%) was classified as stage 2. None of the patients required renal replacement therapy. At discharge, 6 (30.0%) patients still had AKI: 5 (83.3%) were classified as stage 1 and 1 (16.7%) as stage 2. Among the patients, hypertension (HT) was present in 49 patients (29.5%), diabetes mellitus (DM) in 24 patients (14.5%), obstructive lung disease in 23 patients (13.9%), atherosclerotic heart disease in 20 patients (12.0%), hypothyroidism in 6 patients (3.6%), and other comorbidities in 11 patients (6.6%). Patient demographics are summarized in Table 1.

Compared with patients without AKI, patients with AKI had a higher body mass index (BMI), more comorbidities, and a higher CCI score. AKI occurred in 20.4% of patients with HT and in 25% of patients with DM (Table 2). In univariate analysis, increased BMI (OR: 1.12, 95% CI: 1.03–1.21, P = 0.006), higher CCI score (OR: 1.62, 95% CI: 1.11–2.36, P = 0.013), presence of comorbidities (OR: 2.92, 95% CI: 1.01–8.45, P = 0.048), HT (OR: 2.74, 95% CI: 1.06–7.09, P = 0.037), DM (OR: 3.05, 95% CI: 1.04–8.94, P = 0.042), drug use (OR: 3.49, 95% CI: 1.11–10.94, P = 0.032), non-steroidal anti-inflammatory drug (NSAID) use (OR: 4.98, 95% CI: 1.09–22.69, P = 0.038), oral antidiabetic drug use (OR: 3.25, 95% CI: 1.10–9.59, P = 0.033), and acetylsalicylic acid use (OR: 5.83, 95% CI: 1.84–18.44, P = 0.003) were significantly associated with development of AKI. In the multivariate logistic regression analysis, increased BMI (OR: 1.10, 95% CI: 1.01–1.21, P = 0.038), acetylsalicylic acid use (OR: 10.56, 95% CI: 1.58–70.60, P = 0.015), and diuretic use (OR: 0.06, 95% CI: 0.01–0.50, P = 0.009) were found to be independent predictors of AKI. Although diuretic use was

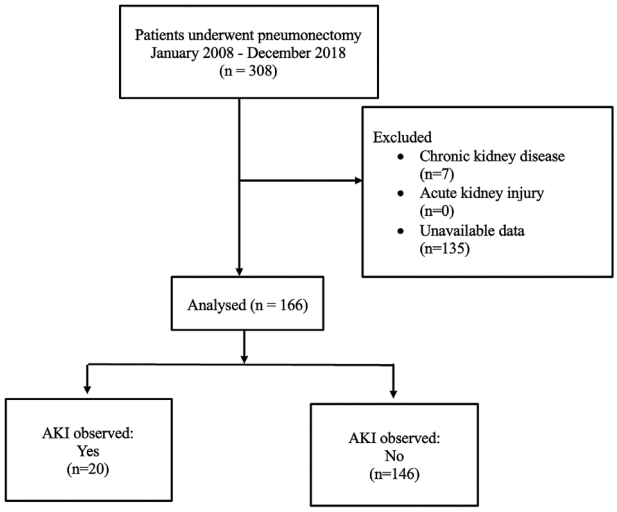

Figure 1. Flow chart

AKI: acute kidney injury

not significantly associated with AKI in univariate analysis, it emerged as an independent protective factor in the multivariate logistic regression model.

AKI occurred in seven (28%) of the patients who received colloids, compared with 13 (9.2%) of the patients who did not receive colloids; this difference was statistically significant (P = 0.015). Moreover, the mean fluid volume administered was 2525.00±1312.64 mL in patients with AKI and 1882.53±768.11 mL in patients without AKI, representing a statistically significant difference between the two groups (P = 0.006). Other intraoperative variables did not differ significantly between groups (Table 3). Univariate logistic regression analysis indicated that higher total intraoperative fluid volume was significantly associated with an increased risk of AKI [OR: 1.00 (95% CI: 1.00–1.00), P = 0.008], whereas surgical duration exhibited only a borderline association [OR: 1.01 (95% CI: 1.00–1.01), P = 0.092]. In the multivariate analysis, total fluid volume remained an independent predictor of AKI (OR: 1.00; 95% CI: 1.00–1.00; P = 0.036), whereas surgical duration was not statistically significant (P = 0.761).

In the postoperative period, AKI incidence was significantly higher in patients who received RBC transfusion, NSAIDs, or furosemide, as well as in those with complications and CDC grade ≥ 3. The length of hospital stay (LOS) was significantly longer among patients with AKI than among those without AKI. However, there was no significant difference in mortality rates between groups (Table 4). Univariate logistic regression analysis revealed that postoperative AKI was significantly associated with RBC transfusion during the first week (OR: 4.10, 95% CI: 1.18–14.21, P = 0.026), NSAID use (OR: 0.03, 95% CI: 0.01–0.13, P < 0.001), and LOS (OR: 1.06, 95% CI: 1.02–1.11, P = 0.001), while furosemide use (OR: 3.45, 95% CI: 0.89–13.33, P = 0.073), fluoroquinolone use (OR: 2.79, 95% CI: 0.85–9.13, P = 0.090), and CDC class ≥3 (OR: 0.95, 95% CI: 0.35–2.56, P = 0.922) were not significantly associated. In the multivariate logistic regression analysis, only postoperative NSAID use (OR: 0.03, 95% CI: 0.00–0.13, P < 0.001) and LOS (OR: 1.07, 95% CI: 1.01–1.13, P = 0.016) were identified as independent predictors.

Table 1. Demographic data of all patients

Parameters		
Age (years)	Mean ± SD	58.96±8.50
	Median (min-max)	59.00 (26.00–80.00)
Gender, n (%)	Female	14 (8.4)
	Male	152 (91.6)
Diagnosis, n (%)	Squamous epithelial cell carcinoma	107 (64.5)
	Adenocarcinoma	29 (17.5)
	Non-small cell lung cancer	15 (9.0)
	Pleomorphic carcinoma	5 (3.0)
	Carcinoid tumor	2 (1.2)
	Others	8 (4.8)
BMI (kg/m²)	Mean ± SD	26.53±5.18
	Median (min-max)	25.86 (15.28–48.33)
ASA class, n (%)	1	76 (45.8)
	2	78 (47.0)
	3	12 (7.2)
Comorbidity, n (%)	No	77 (46.4)
	Yes	89 (53.6)
CCI point	Mean ± SD	3.59±1.21
	Median (min-max)	4.00 (1.00–9.00)
Diuretic usage, n (%)	No	131 (78.9)
	Yes	35 (21.1)
AKI stage, n (%)	1	19 (95.0)
	2	1 (5.0)
Presence of AKI at discharge, n (%)	No	14 (70.0)
	Yes	6 (30.0)
AKI stage at discharge, n (%)	1	5 (83.3)
	2	1 (16.7)
Complication, n (%)	No	117 (70.5%)
	Yes	49 (29.5%)
	Wound infection	14 (8.4)
	Arrhythmia	12 (7.2)
	Pneumonia	11 (6.6)
	Bleeding	10 (6.0)
	Bronchopleural fistula	5 (3.0)
	Elevated cardiac enzymes	4 (2.4)
	Respiratory failure	2 (0.12)
	Empyema	2 (0.12)
	Chylothorax	2 (0.12)
	Pulmonary thromboembolism	1 (0.6)
	Pericarditis	1 (0.6)
	Recurrent nerve damage	1 (0.6)
	Esophageal rupture	1 (0.6)

AKI: acute kidney injury, ASA: American Society of Anesthesiologist, BMI: body mass index, CCI: Charlson comorbidity index, SD: standard deviation, min: minimum, max: maximum

Table 2. Evaluation of demographic characteristics and risk factors of patients with and without AKI

		AKI		*P* value
		No	Yes	
Gender, n (%)	Female	11 (78.6)	3 (21.4)	*P* = 0.38[b]
	Male	135 (88.8)	17 (11.2)	
Age (year)	Mean ± SD	58.72±8.47	60.75±8.75	*P* = 0.32[c]
	Median (min-max)	58.00 (26.00–80.00)	61.50 (42.00±78.00)	
BMI	Mean ± SD	26.17±5.02	29.25±5.77	***P* = 0.016**[d]
	Median (min-max)	25.51 (15.28–48.33)	30.69 (18.34–42.61)	
ASA, class, n (%)	1	71 (93.4)	5 (6.6)	*P* = 0.14[a]
	2	65 (88.3)	13 (16.7)	
	3	10 (83.3)	2 (16.7)	
Comorbidity n, (%)	No	72 (93.5)	5 (6.5)	***P* = 0.041**[a]
	Yes	74 (83.1)	15 (16.9)	
CCI point	Mean ± SD	3.50±1.19	4.25±1.16	***P* = 0.005**[d]
	Median (min-max)	3.00 (1.00–9.00)	4.00 (2.00–7.00)	
Hypertension, n (%)	No	107 (91.5)	10 (8.5)	***P* = 0.032**[a]
	Yes	39 (79.6)	10 (20.4)	
Diabetes mellitus, n (%)	No	128 (90.1)	14 (9.9)	***P* = 0.046**[b]
	Yes	18 (75.0)	6 (25.0)	
Drug usage, n (%)	No	68 (94.4)	4 (5.6)	***P* = 0.025**[a]
	Yes	78 (83.0)	16 (17.0)	
Diuretic usage, n (%)	No	114 (87.0)	17 (13.0)	*P* = 0.57[b]
	Yes	32 (91.4)	3 (8.6)	
NSAID, n (%)	No	141 (89.2)	17 (10.8)	*P* = 0.057[b]
	Yes	5 (62.5)	3 (37.5)	
Paracetamol, n (%)	No	145 (87.9)	20 (12.1)	*P* > 0.99[b]
	Yes	1 (100.0)	0 (0.0)	
ACEI, n (%)	No	128 (87.7)	18 (12.3)	*P* > 0.99[b]
	Yes	18 (90.0)	2 (10.0)	
Furosemide, n (%)	No	143 (87.7)	20 (12.3)	*P* > 0.99[b]
	Yes	3 (100.0)	0 (0.0)	
Oral antidiabetic, n (%)	No	129 (90.2)	14 (9.8)	***P* = 0.038**[b]
	Yes	17 (73.9)	6 (26.1)	
Acetylsalicylic acid, n (%)	No	136 (90.7)	14 (9.3)	***P* = 0.005**[b]
	Yes	10 (62.5)	6 (37.5)	

[a]: chi-square test, [b]: Fisher's exact test, [c]: Student's t-test, [d]: Mann-Whitney U test

ACEI: angiotensin-converting enzyme inhibitors, AKI: acute kidney injury, ASA: American Society of Anesthesiologist, BMI: body mass index, CCI: Charlson comorbidity index, SD: standard deviation, NSAID: non-steroidal anti-inflammatory drug, min: minimum, max: maximum

DISCUSSION

We assessed the incidence of AKI and its risk factors following pneumonectomy. AKI, as defined by KDIGO criteria, occurred in 12% of patients within the first postoperative week. In multivariable analysis, increased BMI, acetylsalicylic acid use, and diuretic use were independently associated with AKI. AKI was also significantly associated with higher complication rates and a prolonged hospital stay, but not with increased mortality.

The incidence of AKI following thoracic surgery was found to be 15.1% in the study conducted by Naruka et al.[9] We identified a lower AKI rate; however, our results pertain to a one-week period. In a meta-analysis in which 34,826 patients were evaluated, the incidence of AKI following thoracic surgery was 8.8%.[10] We may have observed a different incidence due to multiple factors related to the patient, surgery, and anesthesia that contribute to the development of AKI, as well as to the methods used in our center.

Table 3. Intraoperative data of patients with and without AKI

		AKI		*P* value
		No	Yes	
Hypotension, n (%)	No	131 (90.3)	14 (9.7)	***P* = 0.024**[b]
	Yes	15 (71.4)	6 (28.6)	
Hypertension, n (%)	No	126 (88.7)	16 (11.3)	*P* = 0.50[b]
	Yes	20 (83.3)	4 (16.7)	
Red blood cell transfusion, n (%)	No	132 (88.6)	17 (11.4)	*P* = 0.50[b]
	Yes	14 (82.4)	3 (17.6)	
Colloid, n (%)	No	128 (90.8)	13 (9.2)	***P* = 0.015**[b]
	Yes	18 (72.0)	7 (28.0)	
Total fluid	Mean ± SD	1882.53±768.11	2525.00±1312.64	***P* = 0.006**[c]
	Median (min-max)	1500.00 (750–6000)	2000.00 (1000–7000)	
Urine output (kg/hour)	Mean ± SD	1.53±1.06	1.45±1.29	*P* = 0.42[c]
	Median (min-max)	1.20 (0.19–5.70)	1.05 (0.17–5.80)	
Anuria, n (%)	No	130 (89.7)	15 (10.3)	*P* = 0.14[b]
	Yes	16 (76.2)	5 (23.8)	
Oliguria, n (%)	No	125 (89.3)	15 (10.7)	*P* = 0.32[b]
	Yes	21 (80.8)	5 (19.2)	
Furosemide, n (%)	No	143 (87.7)	20 (12.3)	*P* > 0.99[b]
	Yes	3 (100.0)	0 (0.0)	
Duration of anesthesia	Mean ± SD	219.66±63.03	250.25±80.42	*P* = 0.22[c]
	Median (min-max)	210.00 (75.00–540.00)	210.00 (150.00–390.00)	
Duration of surgery	Mean ± SD	187.10±57.31	210.50±75.57	*P* = 0.52[c]
	Median (min-max)	180.00 (75.00–480.00)	172.50 (120.00–360.00)	

[b]: Fisher's exact test, [c]: Mann-Whitney U test
AKI: acute kidney injury, SD: standard deviation, min: minimum, max: maximum

The identification of risk factors for postoperative AKI following thoracic surgery remains a challenge. In the study conducted by Licker et al.[11], anesthesia duration, forced expiratory volume in 1 second, and ASA classification were significantly associated with postoperative AKI. Ishikawa et al.[5] identified HT, peripheral vascular disease, angiotensin II receptor blocker (ARB) use, and hydroxyethyl starch use as independent risk factors. Zhao et al.[12] found that older age, HT, DM, ARB use and angiotensin-converting enzyme inhibitor use, preoperative albumin and sCr levels, intraoperative blood loss, and hypotension were significantly associated with postoperative AKI. In the univariate analysis, increased BMI, higher CCI score, HT, DM, and presence of comorbidity were each significantly associated with AKI following pneumonectomy. Similarly, in a large cohort study of 161,185 patients undergoing major surgery, older age, male sex, higher BMI, DM, HT, malignancy, lung disease, a low glomerular filtration rate, use of ACEI/ARB, use of diuretics, and surgery later during hospitalization were associated with postoperative AKI following thoracic surgery.[2] The development of AKI during the perioperative period is influenced by a number of patient-, anesthetic-, and surgery-related factors. We identified increased BMI, acetylsalicylic acid use, and diuretic use as independent predictors of postoperative AKI. Our results highlight BMI as a strong risk factor for AKI, consistent with previous studies reporting increased susceptibility to renal complications among overweight and obese patients. The observed association between acetylsalicylic acid use and AKI is notable, as acetylsalicylic acid can impair renal autoregulation and may exacerbate perioperative renal hypoperfusion, especially in patients with multiple comorbidities or in those exposed to other nephrotoxic agents. However, the strong association between acetylsalicylic acid use and AKI in our analysis should be interpreted with caution due to the relatively small number of AKI events and potential residual confounding. Interestingly, diuretic use was associated with a markedly reduced risk of AKI. While this finding may suggest a protective effect, it is possible that patients receiving diuretics were more closely monitored and managed in the perioperative period, or that diuretic use reflects other aspects of patient care not fully captured in the present dataset. Traditional risk factors such as HT and DM were not independent predictors of AKI in multivariate analysis, likely due to overlap with BMI and medications and to the limited statistical power of our cohort. Our results are consistent with the findings reported in previous studies. However, the heterogeneity of thoracic surgery patient populations and varying treatment protocols across centers may affect the identification of AKI-related risk factors. Further

research is warranted to definitively determine the risk factors for AKI following pneumonectomy.

During the perioperative phase, patients are at increased risk of developing AKI because of surgical stress, systemic inflammation, renal hypoperfusion induced by hypovolemia, and the vasodilatory and cardiodepressant effects of anesthesia.[1] In this cohort, the principal intraoperative risk factors for postoperative AKI were hypotension and greater total intraoperative fluid administration. In thoracic surgery,

Table 4. Evaluation of postoperative data of patients with and without AKI

		AKI		*P* value
		No	Yes	
RBC-tx in 7 days, n (%)	No	127 (91.4)	12 (8.6)	***P* = 0.006**[b]
	Yes	19 (70.4)	8 (29.6)	
Total fluid in 24 hours	Mean ± SD	1949.47±325.34	2026.25±415.67	*P* = 0.66[c]
	Median (min-max)	1970.00 (1260–3600)	1975.00 (1350–3090)	
Patient controlled analgesia, n (%)	IV fentanyl	2 (66.7)	1 (33.3)	*P* = 0.47[b]
	IV contromal	135 (87.7)	19 (12.3)	
	Epidural bupivacaine	7 (100.0)	0 (0.0)	
Postoperative drug usage				
NSAID, n (%)	No	4 (33.3)	8 (66.7)	***P* < 0.001**[b]
	Yes	142 (92.2)	12 (7.8)	
Paracetamol, n (%)	No	113 (90.4)	12 (9.6)	*P* = 0.10[b]
	Yes	33 (80.5)	8 (19.5)	
Opioid, n (%)	No	104 (88.9)	13 (11.1)	*P* = 0.57[a]
	Yes	42 (85.7)	7 (14.3)	
ARB, n (%)	No	138 (87.9)	19 (12.1)	*P* > 0.99[b]
	Yes	8 (88.9)	1 (11.1)	
ACEI, n (%)	No	128 (86.5)	20 (13.5)	*P* = 0.13[b]
	Yes	18 (100.0)	0 (0.0)	
Furosemide, n (%)	No	128 (90.8)	13 (9.2)	***P* = 0.015**[b]
	Yes	18 (72.0)	7 (28.0)	
Cephalosporin, n (%)	No	5 (100.0)	0 (0.0)	*P* > 0.99[b]
	Yes	141 (87.6)	20 (12.4)	
Fluoroquinolones, n (%)	No	103 (92.0)	9 (8.0)	***P* = 0.022**[a]
	Yes	43 (79.6)	11 (20.4)	
LOS in hospital	Mean ± SD	13.23±8.22	23.90±20.21	***P* = 0.011**[c]
	Median (min-max)	12.00 (4.00–63.00)	19.50 (6.00–85.00)	
Complication, n (%)	No	110 (94.0)	7 (6.0)	***P* < 0.001**[a]
	Yes	36 (73.5)	13 (26.5)	
CDC, n (%)	<3	137 (90.1)	15 (9.9)	***P* = 0.015**[b]
	≥3	9 (64.3)	5 (35.7)	
Hospital mortality, n (%)	No	145 (88.4)	19 (11.6)	*P* = 0.23[b]
	Yes	1 (50.0)	1 (50.0)	
30-day mortality, n (%)	No	144 (87.8)	20 (12.2)	*P* > 0.99[b]
	Yes	2 (100.0)	0 (0.0)	
90-day mortality, n (%)	No	140 (88.1)	19 (11.9)	*P* > 0.99[b]
	Yes	6 (85.7)	1 (14.3)	
1-year mortality, n (%)	No	136 (88.9)	17 (11.1)	*P* = 0.16[b]
	Yes	9 (75.0)	3 (25.0)	

ACEI: angiotensin-converting enzyme inhibitors, AKI: acute kidney injury, ARB: angiotensin receptor blockers, CDC: Clavien-Dindo classification, RBC-tx: red blood cell transfusion, SD: standard deviation, LOS: length of stay, NSAID: non-steroidal anti-inflammatory drug

[a]: chi-square test, [b]: Fisher's exact test, [c]: Mann-Whitney U test

perioperative fluid and hemodynamic management is a complex area. We hypothesized that intravascular volume restriction might be linked to AKI following pneumonectomies. However, fluid restriction alone did not appear to be a risk factor for AKI in our study population. Kaufmann et al.[13] found that fluid restriction to prevent pulmonary complications did not cause AKI when regulated to prevent relative hypovolemia. In our study, the association between increased intraoperative fluid administration and AKI development is noteworthy. While adequate fluid resuscitation is essential, excessive fluid administration may contribute to renal congestion and impaired oxygenation, thus paradoxically increasing the risk of AKI. Our findings highlight the importance of individualized fluid management strategies and careful intraoperative monitoring.

In contrast to crystalloids, we established a connection between colloid administration (we prefer albumin) and AKI following lung resection. There is no definitive conclusion regarding the use of albumin as a colloid in thoracic surgery. However, our results indicated that 28% of patients receiving colloids developed postoperative AKI. Hydroxyethyl starch use has previously been shown to be associated with AKI.[5,14] There is no evidence linking albumin use to AKI. In contrast, lower preoperative albumin levels have been shown to be a risk factor for AKI.[14,15] Low albumin levels negatively impact patient outcomes, which is not surprising given that albumin levels are influenced by numerous factors, including malnutrition and stress levels.

We observed that the incidence of AKI was significantly higher in patients with intraoperative hypotension. It was shown that intraoperative vasopressor use by Licker et al.[11], intraoperative hypotension by Ren and Meng[16], and lower intraoperative mean arterial pressure values by Zhao et al.[12] were related to the development of AKI. Cardinale et al.[17] found that AKI occurred in patients who experienced greater intraoperative blood loss and received more postoperative transfusions. Hypotension during surgery was more frequently observed among patients who developed AKI, underscoring the critical role of hemodynamic stability in renal protection. These findings are consistent with previous literature demonstrating the vulnerability of renal function to intraoperative hemodynamic disturbances. Although no association was found between intraoperative blood component transfusions and AKI in our study, it is important to note that our exposure level was very low. However, transfusions during the first postoperative week were associated with AKI.

Other intraoperative factors, such as HT, blood transfusion, altered urine output (anuria or oliguria), anesthesia, and surgical duration, did not independently predict AKI in this cohort. Although surgical duration showed a borderline association, it did not reach statistical significance in the multivariate analysis. Zhao et al.[12] found a moderate association but poor predictive value of urine output for AKI. Prolonged anesthesia duration[11] and longer surgical time[18] have been shown to be associated with AKI. Matesanz et al.[18] found that patients who developed AKI following thoracic surgery had a greater postoperative systemic inflammatory response, which was associated with an increased risk of complications and mortality.

Our study demonstrates that postoperative AKI was independently associated with prolonged LOS and the absence of NSAID use. While use of RBC-tx, furosemide, and fluoroquinolones was more common in patients who developed AKI, these associations did not remain statistically significant after adjustment for confounders in the multivariate model.

This study comprehensively evaluated postoperative risk factors for AKI following major surgery, with particular focus on clinical interventions and postoperative complications. Our findings demonstrate that patients with AKI had higher rates of RBC transfusion, greater use of NSAIDs, furosemide, and fluoroquinolones, more complications, and longer hospital stays. In a multivariate model, postoperative AKI was independently associated with prolonged hospital stay and the absence of NSAID therapy. This finding is counterintuitive, as NSAIDs are generally considered nephrotoxic and have been widely reported as risk factors for AKI, particularly in vulnerable populations.[19] The unexpected association between postoperative NSAID use and a reduced risk of AKI is most likely attributable to selection bias, as clinicians may have avoided NSAIDs in patients considered to be at higher risk (e.g., those perceived to have increased susceptibility to renal impairment). However, due to the retrospective design and incomplete documentation regarding the reasons for withholding NSAIDs, we could not exclude such cases from the analysis. As such, this finding should be interpreted with great caution and cannot be considered evidence of a true protective effect. Additionally, the relatively small number of AKI events may have influenced the statistical outcome. Therefore, while our results suggest a potential inverse association between NSAID use and postoperative AKI, these findings should be interpreted with caution and require validation in larger prospective studies.

The LOS stay after thoracic surgery is influenced by the type of surgery, patient comorbidities, and postoperative complications. Patients without complications after wedge resection are discharged earlier than those with CDC grade ≥3 complications after pneumonectomy. The LOS has been shown to be twice as long in patients with CDC grade ≥3 complications than in those with CDC grade <3 complications.[20] AKI following thoracic surgery has also been consistently associated with prolonged hospital stays, as demonstrated in multiple studies.[3,5] In our study, the incidence of AKI was higher among patients with CDC grade ≥3 complications than among those with CDC grade <3. Similarly, Murphy et al.[21] reported that AKI after thoracic surgery was significantly associated with CDC grade 3, the comprehensive complications index, pneumonia, and respiratory failure.

In contrast to much of the existing literature, we were unable to demonstrate a positive relationship between mortality and postoperative AKI.[3,18] Differences in patient characteristics and treatment protocols at our center may have contributed to this finding. Additionally, the relatively small number of AKI cases in our cohort may have limited the statistical power to detect a significant difference in mortality rates. Because of the correlation between preoperative comorbidities and the risk of AKI, perioperative management must be conducted more carefully to prevent AKI and improve patient outcomes. The incidence of AKI has been reported to increase with the extent

of lung resection.[17] Because of the distinctive nature of their condition, pneumonectomy patients require increased levels of care. Moreover, mortality in pneumonectomy patients is multifactorial and may be influenced by a range of perioperative and non-renal factors, potentially diluting the impact of AKI on overall outcomes. The lack of an established clinical prediction rule for AKI following pneumonectomy may negatively affect clinicians' approach to this population. Recognition of risk factors could help establish clinical prediction rules and enhance patient outcomes. Further studies with larger sample sizes are warranted to clarify the relationship between AKI and mortality after pneumonectomy.

The strengths of this study include its relatively large, single-center cohort spanning ten years, comprehensive perioperative data collection, and detailed evaluation of short-term outcomes. These features provide valuable and reliable information for identifying high-risk patients and improving perioperative management in clinical practice.

Study Limitations

This study has several limitations. First, it was designed as a single-center retrospective cohort study, which may limit the generalizability of the findings to other populations. The retrospective nature of the study may also introduce bias due to incomplete or missing data, as well as potential inaccuracies in the recorded information. Second, AKI was defined exclusively by changes in sCr levels because postoperative urine output data were not routinely collected. As a result, episodes of AKI identified solely by oliguria may have been missed, potentially leading to an underestimation of the true AKI incidence. Third, several relevant perioperative factors and comorbidities could not be fully assessed due to data limitations, which may have resulted in residual confounding. The preoperative pulmonary function test results (e.g., FEV) and tumor stage were not consistently available in the retrospective dataset; therefore, these variables could not be incorporated into the statistical analyses, although they are clinically relevant factors known to be associated with postoperative outcomes. Fourth, the relatively limited sample size may have reduced the statistical power of our analyses and the ability to detect weaker associations. Fifth, the findings regarding the potential protective effects of NSAID and diuretic use should be interpreted with caution, as these results may be subject to confounding by indication and selection bias—NSAIDs and diuretics may have been avoided in higher risk patients or those with impaired renal function, resulting in an apparent protective association in the remaining cohort. Furthermore, the small number of AKI events may have limited the ability to robustly evaluate the association between these drug exposures and AKI risk. Finally, long-term outcomes beyond hospital discharge, including persistent renal dysfunction and overall survival, were not systematically assessed, thereby preventing evaluation of the extended impact of postoperative AKI.

CONCLUSION

Reducing the incidence of AKI could significantly improve patient outcomes, and we propose that AKI may serve as a valuable quality metric in thoracic surgery. To our knowledge, no clinical prediction rules have yet been established for forecasting AKI following thoracic surgery. Recognition of risk factors, as outlined in this systematic review, could lay the groundwork for the development of such predictive models. Nevertheless, AKI should be regarded as a crucial outcome measure in thoracic surgery, necessitating further studies focused on early prediction, risk minimization, and the long-term consequences of this complication.

Ethics

Ethics Committee Approval: The Ethics Committee of Ankara University Faculty of Medicine approval number is I8-568-21, date: 15/10/2021.

Informed Consent: The need for obtaining informed consent was waived because of the retrospective nature of the study.

Footnotes

Authorship Contributions

Concept: Ç.Y.G., S.K.E., Design: Ç.Y.G., S.K.E., Data Collection or Processing: B.Ş., Y.K., S.G.G., Analysis or Interpretation: B.Ş., Y.K., S.G.G., Literature Search: Ç.Y.G., S.K.E., B.Ş., B.C.M., Writing: Ç.Y.G., B.Ş., B.C.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. Gumbert SD, Kork F, Jackson ML, et al. Perioperative acute kidney injury. *Anesthesiology*. 2020;132(1):180-204. **[Crossref]**
2. Grams ME, Sang Y, Coresh J, et al. Acute kidney injury after major surgery: a retrospective analysis of veterans health administration data. *Am J Kidney Dis*. 2016;67(6):872-880. **[Crossref]**
3. Lei SH, Guo GF, Yan T, Zhao BC, Qiu SD, Liu KX. Acute kidney injury after general thoracic surgery: a systematic review and meta-analysis. *J Surg Res*. 2023;287:72-81. **[Crossref]**
4. Chudgar N, Yan S, Hsu M, et al. The American College of surgeons surgical risk calculator performs well for pulmonary resection: a validation study. *J Thorac Cardiovasc Surg*. 2021;163(4):1509-1516. **[Crossref]**
5. Ishikawa S, Griesdale DEG, Lohser J. Acute kidney injury after lung resection surgery: incidence and perioperative risk factors. *Anesth Analg*. 2012;114(6):1256-1262. **[Crossref]**
6. Kellum JA, Lameire N, Aspelin P, et al. Kidney Disease Improving Global Outcomes (KDIGO) acute kidney injury work group. KDIGO clinical practice guideline for acute kidney injury. *Kidney Int*. 2012;2(1):1-138. **[Crossref]**
7. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-383. **[Crossref]**
8. Dindo D, Demartines, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. *Ann Surg*. 2004;240(2):205-213. **[Crossref]**

9. Naruka V, Mckie MA, Khushiwal R, et al. Acute kidney injury after thoracic surgery: a proposal for a multicentre evaluation (MERITS). *Interact Cardiovasc Thorac Surg.* 2019;29(6):861-866. **[Crossref]**
10. Yu Y, Xu S, Yan B, et al. Incidence and associations of acute kidney injury after general thoracic surgery: a system review and meta-analysis. *J Clin Med.* 2022;12(1):37. **[Crossref]**
11. Licker M, Cartier V, Robert J, et al. Risk factors of acute kidney injury according to RIFLE criteria after lung cancer surgery. *Ann Thorac Surg.* 2011;91(3):844-850. **[Crossref]**
12. Zhao BC, Lei SH, Yang X, et al. Assessment of prognostic value of intraoperative oliguria for postoperative acute kidney injury: a retrospective cohort study. *Br J Anaesth.* 2021;126(4):799-807. **[Crossref]**
13. Kaufmann KB, Stein L, Bogatyreva L, et al. Oesophageal Doppler guided goal-directed haemodynamic therapy in thoracic surgery-a single centre randomized parallel-arm trial. *Br J Anaesth.* 2017;118(6):852-861. **[Crossref]**
14. Ahn HJ, Kim JA, Lee AR, Yang M, Jung HJ, Heo B. The risk of acute kidney injury from fluid restriction and hydroxyethyl starch in thoracic surgery. *Anesth Analg.* 2016;122(1):186-193. **[Crossref]**
15. Lee EH, Kim HR, Baek SH, et al. Risk factors of postoperative acute kidney injury in patients undergoing esophageal cancer surgery. *J Cardiothorac Vasc Anesth.* 2014;28(4):936-942. **[Crossref]**
16. Pang Z, Zou W. Advances of perioperative acute kidney injury in elderly patients undergoing non-cardiac surgery. *Zhong Nan Da Xue Xue Bao Yi Xue Ban.* 2023;48(5):760-770. English, Chinese. **[Crossref]**
17. Cardinale D, Cosentino N, Moltrasio M, et al. Acute kidney injury after lung cancer surgery: incidence and clinical relevance, predictors, and role of N-terminal pro B-type natriuretic peptide. *Lung Cancer.* 2018;123:155-159. **[Crossref]**
18. Matesanz CM, Gala FL, Rancan L, et al. Predictive value of plasma cytokines for acute kidney injury following lung resection surgery: prospective observational study. *Braz J Anesthesiol.* 2019;69(3):242-252. **[Crossref]**
19. Gumbert SD, Kork F, Jackson ML, et al. Perioperative acute kidney injury. *Anesthesiology.* 2020;132(1):180-204. **[Crossref]**
20. Finlev CJ, Begum HA, Pearce K, et al. The effect of majör and minör complications after lung surgery on length of stay and readmission. *J Patient Exp.* 2022;9:1-8. **[Crossref]**
21. Murphy CF, Dunne T, Elliott JA, et al. Acute kidney injury after esophageal cancer surgery: incidence, risk factors, and impact on oncologic outcomes. *Ann Surg.* 2022;275(5):683-689. **[Crossref]**

Thorac Res Pract. 2026;27(1):30-37 **DOI:** 10.4274/ThoracResPract.2025.2025-6-5

Original Article

Should the Mandatory Health Referral System be Implemented in Türkiye? A Perspective from a Tertiary Pulmonology Hospital

Hülya Abalı, Fatma Tokgöz Akyıl

Clinic of Pulmonology, University of Health Sciences Türkiye, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital, İstanbul, Türkiye

Cite this article as: Abalı H, Tokgöz Akyıl F. Should the mandatory health referral system be implemented in Türkiye? A perspective from a tertiary pulmonology hospital. *Thorac Res Pract.* 2026;27(1):30-37

ABSTRACT

OBJECTIVE: Providing care at the appropriate level is key to cost-effective healthcare. In Türkiye, a mandatory referral system has not yet been implemented. This study aimed to determine the proportion of cases presented to a tertiary outpatient clinic that genuinely required tertiary-level care and to analyze their characteristics compared with those better suited for management at lower levels of care.

MATERIAL AND METHODS: This observational study included 692 patients (mean age: 54±15 years; 51% male) who attended two pulmonologists' outpatient clinics. Data on demographics, reasons for application, and all institutions to which individuals applied (primary, secondary, tertiary) for chest conditions within the previous 12 months were collected. After clinical evaluations, the appropriate level of care for each case was assessed.

RESULTS: Overall, 70.2% (n = 486) of cases bypassed primary care. While 10.7% (n = 74) required tertiary-level care, 66.3% (n = 459) required secondary-level care. Emergency departments (30%, n = 208) were visited more frequently than family physicians (29.6%, n = 205). Among cases seen in primary care, 11% were referred to higher-level care. The most common diagnoses were asthma and chronic obstructive pulmonary disease (56.8%; n = 394). While 46% (n = 318) of all cases received prescriptions or follow-up, 18.6% (n = 129) required interventions. Among the cases, 70.1% avoided non-teaching public hospitals, while 15% (n = 104) made more than 10 return visits to the study hospital.

CONCLUSION: Approximately 70% of cases presenting to a tertiary pulmonology center could have been managed at lower levels of care. Healthcare policymakers should urgently implement sustainable solutions to address the factors that impede the referral chain.

KEYWORDS: Healthcare, referral chain, overcrowding, burnout, tertiary hospital

Received: 09.07.2025 ***Revision Requested:*** 07.08.2025 ***Last Revision Received:*** 18.11.2025

Accepted: 07.12.2025 ***Publication Date:*** 30.01.2026

INTRODUCTION

The health system seeks to offer patients more effective, efficient, and equitable access to organized levels of care, classified as primary, secondary, and tertiary. This goal can be achieved through a referral system that directs patients to more specialized care when necessary.[1]

In many developed countries, a "gatekeeper system" is used whereby residents first consult their family physicians for health issues and are referred to specialists for complex or unresolved cases, thereby avoiding unnecessary referrals to higher levels of care.[2,3] Since patients have most of their health needs met in primary care institutions, their satisfaction increases.[4] Furthermore, multiple referrals to health institutions are avoided; hospital and emergency visits decline; the use of limited resources and healthcare costs decrease; and high-quality care and treatment are provided.[5-7]

Corresponding author: Hülya Abalı, MD, e-mail: hulayab@gmail.com

The history of family medicine within the Turkish health system spans roughly four decades. In 1983, family medicine was officially recognized as a distinct field of medical education in the Medical Specialization Regulation. Specialization training in family medicine began in 1985, and by 1993, departments of family medicine were established within medical faculties. The organization for family medicine was founded in 1990 with the creation of the Turkish Family Physicians Specialist Association.[8] Family medicine model was implemented as part of the Health Transformation Program, published in 2003, and introduced across Türkiye in 2010, with its legal and institutional framework gradually established. Family medicine has become a key part of the referral chain by initiating referrals.[9]

According to the latest health services statistics (2023) 28,054 family medicine units operating in Türkiye met approximately 99% of applications to primary health-care institutions and approximately 43% of applications to all health institutions. Although the number of family physician applications was around 417 million, the referral rate for family physician units remained very low, at 0.2%.[10] Underutilization of the health referral system leads to severe patient overcrowding at secondary and tertiary care institutions. As a result, specialists are often overburdened by outpatient clinic services. Therefore, the extent to which the referral system is followed in tertiary care settings is frequently questioned. We aimed to discuss this issue based on the outcomes of a study in which we evaluated applications to the pulmonology outpatient clinics at a tertiary care hospital.

MATERIAL AND METHODS

This single-center, cross-sectional, descriptive study was conducted at a reference hospital for chest diseases located in a densely populated city and easily accessible from the city's most populous districts. The study protocol was approved by Yedikule Chest Diseases and Thoracic Surgery Education and Research Hospital's Local Ethics Committee (approval no: 21-1; date: 19.03.2025). This study was carried out in accordance with the ethical principles outlined in the Declaration of Helsinki. Participants were informed about the study's objectives, and verbal consent was obtained before the face-to-face interview.

Main Points

- At least seven of every ten cases presenting to the tertiary pulmonology hospital could have been managed at lower levels of care without compromising quality of care.
- Overall, 70.2% had not consulted a family physician in the past year, highlighting opportunities for improvement in the healthcare system.
- Visits to emergency departments were more frequent than those to family physicians (n = 214 vs. 205), which, in itself, suggests a potential overuse of emergency services.
- Only 11% of the 205 cases seen by family physicians were referred to other institutions, suggesting that family care has substantial potential to enhance overall system efficiency.

Data Collection

The pulmonologist researchers recorded data through face-to-face interviews during outpatient appointments scheduled through the Central Physician Appointment System between March and May 2020. Demographic characteristics, educational status, comorbidities, complaints and their duration, and reasons for presentation to our hospital were recorded. All visits related to chest issues within 12 months-including consultations with family physicians, visits to emergency departments, and admissions to public, private, and university hospitals-were retrieved from the "E-nabız" electronic database. These institutions were classified into primary, secondary, and tertiary categories.

Statistical Analysis

Quantitative data are expressed as mean ± standard deviation (SD), and qualitative data are expressed as frequencies. Variables were recorded and analyzed using the statistical software package SPSS for Windows, version 16.0 (SPSS Inc., Chicago, IL, USA). The quantitative data were presented as mean ± SD or median (minimum-maximum), and qualitative data were reported as numbers and frequencies. For comparative statistical analyses, a t-test was used to analyze quantitative data, and a chi-square test was used to analyze qualitative data. Logistic regression was conducted to identify independent predictors of family physician visits.

End Point

Based on examinations, final diagnoses, and treatments, the most appropriate level of care was determined retrospectively.

RESULTS

Of the 692 cases enrolled, 351 (51%) were male, with a mean age of 54±15 years (range, 16-92 years). 10% (n = 72) of the participants were illiterate, 46% (n = 316) had completed primary school, and 10% (n = 70) were university graduates or university students.

The four most common comorbidities were hypertension (21%), asthma (21%), diabetes (16.5%), and chronic obstructive pulmonary disease (COPD) (16%). The most common complaints were shortness of breath (n = 285, 41%) and cough (n = 217, 31%) (Table 1). Complaint durations (in months) were as follows: <1 (n = 94, 13.6%), 1-3 (n = 188, 27.2%), 3-12 (n = 87, 12.6%), and >12 (n = 127, 18.4%).

The reasons for cases' visits to our outpatient clinics, listed below in order of frequency, were: 48.3% (n = 334) presented with complaints; 18.8% (n = 130) attended for follow-up and control visits; 12.6% (n = 87) were referred for examination; 11.8% (n = 82) sought examination based on a recommendation without a preliminary diagnosis; 5.3% (n = 37) required a prescription or medical report; and 3.2% (n = 22) sought a second opinion from other hospitals (Table 1).

The distribution of cases based on the number of previous applications to our pulmonology outpatient clinics was as follows: no previous applications (n = 71, 10%), fewer than 3 (n = 286, 42%), 4-10 (n = 231, 33%), and more than 10

(n = 104, 15%). The etiological factors for follow-up among 321 cases included: asthma (n = 155, 22%); COPD (n = 110, 16%); pulmonary nodules (n = 22, 3%); radiological control, tuberculosis, and bronchiectasis (n = 18, 3%); pulmonary sarcoidosis (n = 6, 1%); pulmonary embolism (n = 5, 0.7%); interstitial lung diseases (n = 3, 0.4%); and obstructive sleep apnea syndrome (n = 2, 0.3%).

Among all cases, 70.2% (n = 486) had not visited a family physician in the past year, and 10% (n = 69) had visited once (Table 2). The main reason these patients did not visit a family physician was that they did not consider consulting one as their first point of contact for their complaints. Of the 205 patients who visited their family physicians, 92 (49%) were prescribed medication and received medical reports, 82 (40%) were prescribed medication for their complaints, and the remaining patients were advised to visit the hospital.

Over the last year, 70.1% of cases were not treated at public hospitals. Among the cases, 6.5% presented to private hospitals and 3.5% presented to other tertiary care hospitals. The proportion of cases presenting to emergency departments was 30% (n = 208) in the past year and 10.3% (n = 71) in the past month (Table 2).

When factors associated with family physician visits were analyzed, cases presenting to family physicians were older (56±14 vs. 53±15 years, $P = 0.005$), had lower follow-up rates at our hospital, and more frequently used emergency department services (Table 3). In logistic regression analysis, emergency

Table 1. Comorbidities, presenting complaints, and reasons for application among patients attending the pulmonology outpatient clinic in the tertiary care hospital

Total n = 692					
Additional diseases	**n (%)**	**Complaints**	**n (%)**	**Reasons for applying to the study center**	**n (%)**
HT	146 (**21**)	Shortness of breath	285 (**41**)	Direct application due to complaints	338 (**48.3**)
Asthma	147 (**21**)	Cough	217 (**31**)	Follow-up	130 (**18.8**)
DM	114 (**16.5**)	Sputum	64 (**9**)	Referral for examination	87 (**12.6**)
COPD	111 (**16**)	Chest pain	56 (**8**)	Prescription-medical reports	37 (**5.3**)
CAD	60 (**9**)	Back pain	39 (**5.6**)	Recommendation for pulmonology examination	82 (**11.8**)
Extrapulmonary malignancies	30 (**4**)	Wheeze	25 (**4**)	Consulting on examinations and diagnoses at external institutions	22 (**3.2**)
Thyroid diseases	27 (**4**)	Rash	22 (**3**)		
CHF	22 (**3**)	Flank pain	17 (**2.5**)		
CTD	22 (**3**)	Fatigue	14 (**2**)		
Arrhythmias	20 (**3**)	Hemoptysis	10 (**1.4**)		
Sequelae of tuberculosis	20 (**3**)	Sweating	8 (**1**)		
Bronchiectasis	9 (**1**)	Sneeze	8 (**1**)		
ILD	4 (**1**)	Snore	4 (**0.6**)		
OSAS	4 (**1**)	Weight loss	2 (**0.3**)		
Lung malignities	4 (**1**)				

N: number of patients, %: percentage of patients, HT: hypertension, DM: diabetes, COPD: chronic obstructive pulmonary disease, CAD, coronary artery disease, CHF: congestive heart failure, CTD: connective tissue disorders, ILD: interstitial lung diseases, OSAS: obstructive sleep apnea syndrome

Table 2. Distribution of patient visits to primary, secondary, and tertiary hospitals within one year of the index pulmonology outpatient clinic visit at our tertiary center

n = 692 **Family physician** **AN n (%)**	**n = 683** **Public hospital** **AN n (%)**	**n = 686** **Emergency department** **AN n (%)**	**n = 684** **Private hospital** **AN n (%)**	**n = 692** **University hospital** **AN n (%)**
0 486 (**70.2**)	**0** 485 (**70.1**)	**0** 478 (**69.1**)	**0** 639 (**92.3**)	**0** 668 (**96.5**)
1 69 (**10**)	**1** 151 (**21.8**)	**1** 134 (**19.4**)	**1** 35 (**5.1**)	**1** 24 (**3.5**)
2 31 (**4.5**)	**2** 21 (**3**)	**2** 39 (**5.6**)	**2** 4 (**0.6**)	
3 22 (**3.2**)	**3** 9 (**1.3**)	**3** 11 (**1.6**)	**3** 2 (**0.3**)	
4 71 (**10.3**)	**4** 10 (**1.4**)	**4** 10 (**1.4**)	**4** 3 (**0.4**)	
5 6 (**0.9**)	**5** 2 (**0.3**)	**5** 4 (**0.6**)	**7** 1 (**0.1**)	
6 3 (**0.4**)	**6** 1 (**0.1**)	**6** 3 (**0.4**)		
7 1 (**0.1**)	**7** 1 (**0.1**)	**7** 3 (**0.4**)		
8 1 (**0.1**)	**8** 1 (**0.1**)	**10** 4 (**0.6**)		
10 2 (**0.3**)	**10** 1 (**0.1**)			
	15 1 (**0.1**)			

AN: application number, n: number of patients, %: percentage of patients

department visits and lack of follow-up at our hospital were independently associated with a lower likelihood of visiting a family physician (Table 4).

Among all cases, the final diagnosis was asthma and COPD in 56.8%. When the final diagnoses and further examination requirements were evaluated, the application rate that could be managed in primary or secondary care institutions was 4.8% (n = 42), and the application rate that could be managed in a secondary care institution was 66.3% (n = 459). While 16.9% (n = 117) of the cases were considered to require an application to a secondary or tertiary care institution, the rate of cases that absolutely required to be examined in a tertiary care institution remained at 10.7% (n = 74) (Table 5).

Cases were divided into two groups: those manageable at secondary- or tertiary-level care and those requiring definitive tertiary-level care (n = 191). These groups were then compared (Table 6). Cases that did not require tertiary hospital admission were more likely to be male, younger, and to have visited family physicians.

DISCUSSION

The main finding of this study was that at least seven out of every ten cases presenting to the tertiary hospital were diagnosed with diseases that could have been managed at lower-level care facilities. Additionally, the referral rate from external institutions to the tertiary hospital for further examination and recommendations was only 24.4%. These findings alone demonstrate the potential for enhancements in the health referral system in Türkiye. The referral rate is a crucial parameter in the hierarchical diagnosis-and-treatment system, as its magnitude directly determines treatment accessibility and the efficiency of healthcare utilization.[11] To identify solutions for targeted healthcare services, the implications of congestion in the referral system are reviewed alongside the study outcomes.

Among the cases, 70.2% (n = 486) did not visit a family physician or use any primary care services last year. Of the 205 cases presenting to family physicians, only 11% were referred to other institutions, indicating that the "gatekeeper system" is not functioning as intended. The overcrowding in tertiary care hospitals results from the deliberate weakening of primary care under neoliberal transformations. Since 1960, the state has provided healthcare in Türkiye as a fundamental human right within the framework of a social state, in accordance with the Health Law 224 on socialization.[12] As was the case globally, healthcare in Türkiye came under the influence of market-driven policies in the 1990s with the launch of the Health Transformation Program.[13] Because the aim of this reform program is to rely heavily on technological and scientific advances cost-effectively, performance-based compensation for healthcare professionals has been the standard since 2004.[14] The current anomalies in the referral system may stem from policy reforms. Further studies may provide clearer insights into administrative measures aimed at liberalization and cost control.

Upon returning to family care as a key step in primary care, we found that family physicians who worked under performance

Table 3. Factors associated with family physicians' visits

	n = 486 No family physician visit n (%)	**n = 206 Any family physician visit n (%)**	***P***
Male gender	251 (52)	99 (48)	0.405
Age (mean ± SD years)	53±15	56±14	**0.005**
Education status			
Illiterate	47 (10)	25 (12)	
Primary school	212 (43)	104 (33)	
Middle/high school	174 (36)	60 (29)	0.154
University	53 (11)	17 (8)	
Follow-up in the study center	191 (39)	131 (64)	**<0.0001**
Any ED visits within 12 months	138 (28)	75 (36)	**0.039**
Any public hospital visits within 12 months	144 (30)	54 (26)	0.408
Any private hospital visits within 12 months	30 (6)	15 (7)	0.614

SD: standard deviation, ED: emergency department

Table 4. Results of logistic regression analysis for family physician visits

	Beta coefficient	**Standard error**	**Wald**	**Significance**	**Exponential beta**	**95% CI**
Age	0.011	0.006	3.450	0.063	1.011	0.999-1.024
Follow-up in the study center	-0.936	0.176	28.411	**<0.0001**	0.392	0.278-0.553
Any ED application within 12 months	-0.406	0.182	4.965	**0.026**	0.666	0.466-0.952
Constant	-0.751	0.384	3.831	0.05	0.472	

Wald: wald test in the logistic regression, CI: confidence interval, ED: emergency department

Table 5. Final diagnoses, examination results, and appropriate institutions for the management of patients when they applied to the pulmonology outpatient clinics at the tertiary study center

n = 679 **Final diagnosis n (%)**	**n = 563** **Examination results n (%)**	**n = 692** **Appropriate institutions for management of patients n (%)**
Asthma, COPD 394 (**56.8**) Pulmonary nodules 70 (**10.1**) Acute bronchitis 35 (**5.1**) Standard examination 33 (**4.8**) Pneumonia 30 (**4.3**) Extrapulmonary diseases 24 (**3.5**) Lung malignancy examination 23 (**3.3**) ILD, HP 13 (**1.9**) Pulmonary embolism 12 (**1.7**) Bronchiectasis, sequelae of tuberculosis 11 (**1.6**) Allergic rhinitis 10 (**1.4**) Pulmonary sarcoidosis 5 (**0.7**) Tuberculosis examination 5 (**0.7**) Pleural effusion examination 4 (**0.6**) Hemoptysis 4 (**0.6**) OSAS 4 (**0.6**) Pulmonary hypertension 1 (**0.1**) Pectus excavatum 1 (**0.1**)	Treatment continuation, prescription, and medical reports 318 (**46**) Routine assessment 124 (**17.9**) Non-invasive examination plan 73 (**13.5**) Invasive examination plan 35 (**5.1**) Referral to allergy clinic 6 (**0.9**) Referral to sleep clinic 5 (**0.7**) Evaluation of results 1 (**0.1**) Hospitalization 1 (**0.1**)	Primary care level 9 (**1.3**) Secondary care level 459 (**66.3**) Tertiary care level 74 (**10.7**) Secondary and tertiary care levels 117 (**16.9**) Primary and secondary care levels 33 (**4.8**)

n: number of patients, COPD: chronic obstructive pulmonary disease, ILD: interstitial lung disease, HP: hypersensitivity pneumonia, OSAS: obstructive sleep apnea syndrome

Table 6. Factors related to referral chain compliance

	n = 501 **Tertiary level not required** **n (%)**	**n = 191** **Tertiary level may be required** **n (%)**	***P***
Male gender	237 (48)	113 (60)	**0.006**
Age (mean ± SD years)	52±14	58±14	**<0.0001**
Education status			
Illiterate	51 (10%)	21 (11)	
Primary school	225 (45%)	91 (48)	
Middle/high school	174 (35%)	60 (31)	0.856
University	51 (10%)	19 (10)	
Follow-up in the study center	233 (47%)	89 (47)	0.998
Any family physician visits within 12 months	172 (34%)	34 (18)	**<0.0001**
Any ED visits within 12 months	157 (31%)	56 (29)	0.646
Any public hospital visits within 12 months	138 (27)	60 (31)	0.347
Any private hospital visits within 12 months	27 (5)	18 (9)	0.059

SD: standard deviation, ED: emergency department

pressure could only spend 3-5 minutes per patient. According to 2023 data, the number of patients per family physician in Türkiye is twice the OECD average.[15] This shortfall may suggest that healthcare professionals prefer not to work in primary care due to unfavorable factors such as excessive workload, low wages, and limited career prospects. An insufficient number of family physicians exacerbates inequities in access to tertiary care among patients who primarily require primary care. Additionally, family physicians are instructed to focus on specific populations rather than serving patients regionally.[16] Although we have not measured this, some low-income individuals may choose tertiary care services for convenience, for example, lower transportation costs and shorter commute times. Finally, prevailing discourse regarding service quality and the performance pressures faced by public professionals may contribute to the current suboptimal functioning of the referral system.

More specifically, the rate of applications to family physicians was lower than that to the emergency department (n = 205

vs. 214). The use of emergency services as primary-level care (30% of presentations) highlights a collapse of the system's triage mechanism. In this context, we also analyzed the factors associated with visits to family physicians. High emergency department visit rates and low follow-up rates at our hospital were found to be independent predictors of low family physician visit rates. These findings suggest that patients receiving follow-up care for chronic diseases at any institution are also more likely to visit family physicians and that these patients represent a more severe group requiring emergency care. Conversely, the finding that individuals without chronic diseases were less likely to seek care from family physicians may reflect greater reliance on tertiary care hospitals.

The most common final diagnoses among the cases were asthma and COPD (56.8%), which were manageable at primary- and secondary-level care institutions. This condition highlighted the weaknesses in the management of chronic pulmonary diseases, especially in primary-level care. Asthma and COPD are chronic pulmonary diseases that can be improved with the appropriate treatment protocol after risk factors and symptoms are controlled. Family physicians may be assigned responsibility for the diagnosis, control, and treatment of these diseases when appropriate conditions are met. A study found that family physicians were inadequately trained in these diseases and that the primary chest diseases they requested training in were COPD and asthma.[17] Nevertheless, family physicians can always enhance their skills and knowledge by participating in postgraduate training programs, notwithstanding performance pressure and funding limitations.[18,19] Additionally, investment in medical devices and equipment at primary care facilities can be reconsidered, especially for chronic pulmonary diseases, for which diagnostic and follow-up capacity is significantly lacking.[20]

In the examination results, the proportion of patients instructed to continue treatment, receive prescriptions, or submit medical reports was 2.5 times greater than the proportion of patients requiring invasive or non-invasive interventions (46% vs. 18.6%). This finding indicates that conditions that could have been managed at primary or secondary care levels unnecessarily increased the workload of the tertiary institution. Such a situation undermines the efficiency of tertiary-care delivery and may hinder critically ill patients' timely access to appropriate specialists-patients who strictly require tertiary-level interventions. This condition may cause serious adverse outcomes in patients with chest diseases. For instance, 2023 data from the Ministry of Health indicate that the incidence of advanced lung cancer is 52.7%.[21] We consider that this rate can be reduced among lung cancer patients who present to a tertiary hospital without delay, provided a well-functioning referral chain is in place.

The fact that 70.1% of patients did not seek care from a public hospital indicates a lack of trust in secondary-level care. This condition may be explained by the general perception of the public that better-quality healthcare services are provided at higher levels of care. Unfortunately, university and private hospitals, which are considered "reference hospitals" under neoliberalism, receive only 4% and 8% of cases, respectively. This result further demonstrates that public hospitals bear an enormous and disproportionate burden in providing basic public services. A functioning referral system would undoubtedly benefit all segments of the population, simplifying allocation of the required level of care across income groups.[22,23]

The factors associated with compliance in the referral chain were also analyzed. Male gender, younger age, and lower rates of applications to family physicians were associated with the need for care at non-tertiary facilities. This may indicate that cases perceive primary care assessments as inadequate, even when their conditions do not warrant tertiary-level care.

Specialist physicians working in tertiary-level care institutions are struggling to cope with excessive workloads caused by an inefficient referral system and performance-related pressures. A study conducted in training and research hospitals, including our center, reported that 92% of physicians experienced burnout syndrome.[23] In Türkiye, health policies, together with harsh working conditions, low wages, increasing rates of violence, and rising immigration in recent years, have been asserted to contribute to physician burnout.[24] The Turkish Medical Association reported that more than 1,400 physicians emigrated in 2021. In the last decade, the total number of physicians who left the country to work abroad has reached 4,000.[25] Additionally, a study conducted among medical students in Türkiye reported that a significant majority of them (94%) were considering practicing medicine overseas, with nearly half (46%) committed to that decision.[26] These data suggest that improving working conditions for healthcare professionals through more effective allocation of patients could reduce turnover and increase physicians' job satisfaction.

For comparison, in countries with a successful "gatekeeper system", such as Cuba, the healthcare system is entirely public and free, whereas in Türkiye it consists of public, semi-public, private, and non-governmental organizations.[27,28] In Cuba, family physicians are assigned on a geographical rather than a population basis, in contrast to Türkiye.[16,29] In Cuba, primary care plays a proactive role in preventive medicine and in the diagnosis, treatment, and follow-up of patients, with approximately 80% of all health problems resolved at the primary care level.[30] In Türkiye, primary care plays a reactive role in preventive medicine, performing procedures such as prescribing medication and reporting.[31] To ensure the feasibility of the mandatory system, we suggest: (1) providing additional funding and authority to family physicians; (2) applying penalties or providing incentives to encourage patients to follow the referral process; and (3) integrating the appointment system in tertiary hospitals with referral requirements. However, without free and accessible primary care services, the mandatory referral system would fail to fulfill its intended purpose.

Study Limitations

This study has some limitations. Because the study was single-center and included a homogeneous group of cases, the results cannot be generalized. The geographical distribution of individuals who did not visit a family physician could not be mapped because of missing data. The relatively small sample size (n = 692) and brief data collection period (two months) may compromise the statistical power. Additionally, no comparison was made with secondary-level institutions with respect to appropriate referrals. Either multi-center study or a

brief report, each with a larger sample size and an extended data collection period that includes secondary and tertiary healthcare institutions has been planned on this topic.

CONCLUSION

The current study demonstrates that shortcomings in the referral system have resulted in at least 70% of patients receiving tertiary-level care could have been appropriately managed at lower levels. This suggests that the challenge lies not only in the system's structure but also in patient behaviors and perceptions. Therefore, simply mandating referrals may not be sufficient.

The following policy measures could be considered in ongoing reform efforts: (1) Immediate reinvestment in community-based primary care, (2) Reevaluation of profit-driven physician quotas, and (3) Establishment of democratic health councils to help guide policy reforms.

Ethics

Ethics Committee Approval: The study protocol was approved by Yedikule Chest Diseases and Thoracic Surgery Education and Research Hospital's Local Ethics Committee (approval no: 21-1; date: 19.03.2025).

Informed Consent: Participants were informed about the study's objectives, and verbal consent was obtained before the face-to-face interview.

Footnotes

Authorship Contributions

Concept: H.A., F.T.A., Design: H.A., F.T.A., Data Collection or Processing: H.A., F.T.A., Analysis or Interpretation: H.A., F.T.A., Literature Search: H.A., Writing: H.A., F.T.A.

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. Kim Y, Chen AH, Keith E, Yee HF Jr, Kushel MB. Not perfect, but better: primary care providers' experiences with electronic referrals in a safety net health system. *J Gen Intern Med*. 2009;24(5):614-619. [Crossref]
2. Kroneman MW, Maarse H, van der Zee J. Direct access in primary care and patient satisfaction: a European study. *Health Policy*. 2006;76(1):72-79. [Crossref]
3. Öcek Z, Çiçeklioğlu M, Türk M. Birinci basamak sağlık hizmetlerinde kapı tutma nedir? Türkiye'de uygulanabilir mi? [What is gatekeeping in primary health care services? Can it be applied in Turkey?]. *Sürekli Tıp Eğitimi Dergisi*. 2006;15(8):144-148. [Crossref]
4. Vafaee-Najar A, Nejatzadegan Z, Pourtaleb A, et al. The quality assessment of family physician service in rural regions, Northeast of Iran in 2012. *Int J Health Policy Manag*. 2014;2(3):137-142. [Crossref]
5. Etter JF, Perneger TV. Introducing managed care in Switzerland: impact on use of health services. *Public Health*. 1997;111(6):417-422. [Crossref]
6. Escarce JJ, Kapur K, Joyce GF, Van Vorst KA. Medical care expenditures under gatekeeper and point-of-service arrangements. *Health Serv Res*. 2001;36(6 Pt 1):1037-1057. [Crossref]
7. Schwenkglenks M, Preiswerk G, Lehner R, Weber F, Szucs TD. Economic efficiency of gate-keeping compared with fee for service plans: a Swiss example. *J Epidemiol Community Health*. 2006;60(1):24-30. [Crossref]
8. Kringos DS, Boerma WGW, Hutchinson A, Saltman RB, editors. Building primary care in a changing Europe: case studies [Internet]. Copenhagen (Denmark): European Observatory on Health Systems and Policies; 2015. Accessed June 12, 2025. [Crossref]
9. Bulut S. Aile hekimlerinin zorunlu sevk zinciri uygulaması hakkındaki görüşlerinin değerlendirilmesi. Thesis. Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü; 2017. [Crossref]
10. T.C. Sağlık Bakanlığı. Sağlık istatistikleri yıllığı 2023 [Internet]. Ankara: T.C. Sağlık Bakanlığı; 2024. Accessed January 13, 2025. [Crossref]
11. Yang N, Shen L, Shu T, Liao S, Peng Y, Wang J. An integrative method for analyzing spatial accessibility in the hierarchical diagnosis and treatment system in China. *Soc Sci Med*. 2021;270:113656. [Crossref]
12. 224 sayılı Sağlık Hizmetlerinin Sosyalleştirilmesi Hakkında Kanun. Resmi Gazete. 12 Ocak 1961; Sayı: 10705. Accessed January 13, 2025. [Crossref]
13. Gönç T. Neoliberal politikaların küresel düzeyde sağlık üzerindeki etkileri. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*. 2017;17(1):161-162. [Crossref]
14. Aka A. Neoliberal transformation of the healty projections: basic healty actors preception of these outlines. *Toplum ve Sosyal Hizmet*. 2012;23(1):175. [Crossref]
15. OECD Health statistics (2023). Consultations. Accessed August 11, 2025. [Crossref]
16. Elbek O, Adaş EB. Sağlıkta dönüşüm: eleştirel bir değerlendirme. *Türkiye Psikiyatri Derneği Bülteni*. 2009;12(1):33-44. [Crossref]
17. Özdemir T, Aydın L, Türkkanı MH, Kılıç T. Perspective of family physicians on chest diseases. *Ankara Med J*. 2014;14(4)131-135. [Crossref]
18. Söğüt A, Çayır Y, Çayır A, et al. Skill of the managing asthma by family physicians and their need of education. *Euras J Fam Med*. 2014;3(1):21-26. [Crossref]
19. Günbatar H, Sertoğullarından B, Ekin S, Arisoy A, Sunnetcioglu A. Evaluation of information level of family practitioners about COPD. *Van Med J*. 2014;21(2):77-80. [Crossref]
20. Çalıştay Raporu. Aile Hekimliğinde Kronik Hastalık Yönetimi. TAHUD İstanbul Şubesi. 18.02.2017. Accessed August 15, 2025. [Crossref]
21. Akciğer Kanseri Farkındalık Ayı, 2023. T.C. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü. Accessed June 12, 2025. [Crossref]
22. Altay A. Sağlık hizmetlerinin sunumunda yeni açılımlar ve Türkiye açısından değerlendirilmesi. *Sayıştay Dergisi*. 2007;64:33-58. [Crossref]
23. Harman Yıldız G, Kubilay Ö. Neoliberal health policies and doctor migration. *Scientific Acad Res*. 2022;1:1-13. [Crossref]
24. Türkiye'de Doktorların Yaşam Kalitesi. Altıncı Bölüm: Doktorların Kariyer Tatmini ve Gelecek Tahayyülü. p.86; 2022. Accessed June 12, 2025. [Crossref]
25. Yıldız GH, Kubilay O. Neoliberal sağlık politikaları ve hekim göçü: neoliberal health policies and doctor migration. *Scientific Acad Res*. 2022;1:1-13. [Crossref]
26. Erdoğan Kaya A, Erdoğan Aktürk B, Aslan E. Factors predicting the motivation to study abroad in Turkish medical students: a causal

investigation into the problem of brain drain. *J Health Sci Med/ JHSM.* 2023;6(2):526-531. **[Crossref]**

27. Constitución de la República de Cuba (2019). Havana: Government of the Republic of Cuba. Accessed August 15, 2025. **[Crossref]**
28. Okursoy A. Health system of Turkey and evaluation of public hospitals performance. Dissertation. Adnan Menderes Üniversitesi; 2010. p. 70-71. **[Crossref]**
29. Dresang LT, Brebrick L, Murray D, Shallue A, Sullivan-Vedder L. Family medicine in Cuba: community-oriented primary care and complementary and alternative medicine. *J Am Board Fam Pract.* 2005;18(4):297-303. **[Crossref]**
30. Sáanchez L. Introducciáon a la Medicina General Integral. Selecciáon de temas literatura báasica. La Habana, Cuba: Editorial Ciencias Máedicas. 1999;105-110. **[Crossref]**
31. Öcek ZA, Çiçeklioğlu M, Yücel U, Özdemir R. Family medicine model in Turkey: a qualitative assessment from the perspectives of primary care workers. *BMC Fam Pract.* 2014;15:38. **[Crossref]**

Thorac Res Pract. 2026;27(1):38-46 **DOI:** 10.4274/ThoracResPract.2025.2025-6-3

Original Article

AI in Patient Care: Evaluating Large Language Model Performance Against Evidence-Based Guidelines for Pulmonary Embolism

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. AI in patient care: evaluating large language model performance against evidence-based guidelines for pulmonary embolism. *Thorac Res Pract.* 2026;27(1):38-46

ABSTRACT

OBJECTIVE: Artificial intelligence (AI)-driven large language models (LLMs) are increasingly used in patient education; however, their ability to interpret and apply clinical guidelines within real-world physician workflows remains uncertain. Pulmonary embolism (PE), with its well-established diagnostic and management protocols, provides a suitable model for evaluating these systems. This study assessed the performance of four widely used AI-driven LLMs—ChatGPT-4o, DeepSeek-V2, Gemini, and Grok—in applying the 2019 European Society of Cardiology guidelines for PE. The focus was on evaluating clinical accuracy, adherence to guidelines, and response consistency.

MATERIAL AND METHODS: Ten open-ended questions based on a simulated PE case were created, covering diagnosis, risk stratification, treatment, and follow-up. Guideline-based reference answers were used for scoring. LLMs were tested under identical conditions, and the responses were anonymized and scored by two emergency physicians using a 10-point scale. Inter-rater reliability was measured using the intraclass correlation coefficient (ICC), and group comparisons were made using Kruskal-Wallis tests.

RESULTS: ChatGPT-4o scored highest (76), followed by Gemini (73.75), Grok (71.25), and DeepSeek-V2 (65). No significant difference was found in total scores (P = 0.390). Performance varied by category; ChatGPT-4o excelled in follow-up, while DeepSeek-V2 performed best in diagnostics. Expert reviewers noted ChatGPT-4o's structured responses and Grok's practicality, but highlighted limitations such as insufficient personalization and guideline gaps. Inter-rater agreement was excellent (ICC: 0.986).

CONCLUSION: AI-driven LLMs show promise in supporting PE management, though none consistently excel in all domains. Further development is needed to enhance clinical integration and guideline compliance.

KEYWORDS: Pulmonary embolism, artificial intelligence, clinical decision support

Received: 23.06.2025 ***Revision Requested:*** 15.11.2025 ***Last Revision Received:*** 18.11.2025

Accepted: 22.12.2025 ***Epub:*** 20.01.2026 ***Publication Date:*** 30.01.2026

INTRODUCTION

Technological advancements have significantly transformed patient care by improving diagnostic accuracy, optimizing treatment planning, and enhancing overall healthcare efficiency. In recent years, artificial intelligence (AI)-driven large language models (LLMs) have emerged as potential clinical decision-support tools capable of assisting physicians in real time during routine patient management, particularly by synthesizing guideline-based recommendations for complex clinical scenarios. LLMs, designed to comprehend, process, and generate human language, are AI-driven systems trained on predefined datasets to respond efficiently to a wide range of queries and to retrieve accurate information from the internet, using advanced natural language processing (NLP) models. Current LLM technologies are capable of extracting evidence-

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

based information and presenting it in a natural conversational format.[1] Consequently, AI-driven LLMs have the potential to serve as valuable point-of-care reference tools by delivering clinical information in a clear, interactive, and context-specific manner. Given their capacity to process extensive medical literature and rapidly evolving clinical protocols, LLMs hold promise as point-of-care reference tools that may support physicians' diagnostic reasoning, risk stratification, and management planning.

The increasing body of literature on the applicability of AI in the medical field highlights its expanding role in healthcare. Studies evaluating the responses of AI-driven LLMs with NLP capabilities to patient inquiries have demonstrated that the accuracy and reliability of these responses often meet medically acceptable standards.[2-4] These findings suggest that AI-driven LLMs hold promise as potential medical advisors, offering reliable health-related information and assisting patients in understanding their conditions. While further validation and regulatory oversight are necessary, the growing evidence supports the potential integration of AI-driven systems into patient education and preliminary medical consultations.

Given the rapid advancements in AI and NLP technologies, LLMs have the potential to bridge this gap by providing physicians with quick and accurate access to updated clinical guidelines. By processing vast amounts of medical literature in real time, AI-driven systems can assist healthcare professionals in retrieving guideline-based recommendations relevant to specific clinical scenarios. Moreover, with the integration of LLMs that incorporate deep learning methods tailored for healthcare professionals, NLP-powered systems have significant potential to analyze and interpret clinical guidelines with greater precision.[5] Additionally, these systems can help synthesize complex information, highlight key updates, and present context-specific recommendations, thereby improving adherence to evidence-based practices. However, the reliability, accuracy, and clinical applicability of AI-generated recommendations remain critical areas that require further investigation and validation.

This study aims to assess the performance of four widely used AI-driven LLMs in managing simulated patient cases across various clinical scenarios. Specifically, we will examine their ability to apply the 2019 European Society of Cardiology (ESC) guidelines for the management of pulmonary embolism (PE) and will evaluate the accuracy of their responses and their adherence to established recommendations.[6] By comparing their outputs with evidence-based standards, we seek to determine their potential role as clinical decision-support tools. Assessing how LLMs handle guideline-driven decision steps within PE management provides meaningful insight into their potential integration into physician workflows.

Main Points

- This study evaluated the clinical accuracy and guideline adherence of four widely used artificial intelligence (AI) chatbots in managing a simulated case of acute pulmonary embolism based on the 2019 European Society of Cardiology guidelines.
- ChatGPT-4o demonstrated the highest overall score, while DeepSeek-V2 had the lowest performance, particularly in risk stratification and treatment planning.
- Despite individual strengths, none of the AI models consistently excelled across all phases of diagnosis, treatment, and follow-up, highlighting variability in clinical applicability.
- Expert reviewers found notable differences in the models' ability to provide structured, evidence-based, and patient-specific responses.

MATERIAL AND METHODS

Study Design and Setting

This study was designed as a comparative evaluation of AI-driven LLMs for patient management, specifically assessing their adherence to the 2019 ESC guidelines for the diagnosis and management of PE. Since the study does not involve human participants, patient data, or personally identifiable information, ethical approval was deemed unnecessary. All analyses were conducted using AI-generated responses to predefined clinical scenarios, ensuring a standardized and controlled assessment environment. This research adheres to ethical considerations relevant to AI-based studies in medicine and aligns with the principles outlined in the Declaration of Helsinki.

Development of Clinical Scenarios and Questions

A simulated clinical case was developed to reflect real-world patient presentations of PE. Based on this scenario, ten open-ended questions were formulated, each addressing a key aspect of PE diagnosis, risk stratification, and management as outlined in the 2019 ESC guidelines. The questions were designed to assess the ability of LLMs to generate evidence-based responses with an appropriate level of clinical accuracy.

To ensure standardization, guideline-based reference answers were pre-constructed for each question using a 10-point scoring system. These answers incorporated predefined key points, enabling an objective and structured evaluation of LLM-generated responses. The complete list of questions and their corresponding reference answers is provided in Table 1.

Selection of Artificial Intelligence-Driven Large Language Models

To ensure a broad and representative comparison, four widely used and advanced AI-driven LLMs were selected based on their global popularity, technological diversity, and NLP capabilities. The selected LLMs were:

ChatGPT-4o (OpenAI, USA): Chosen for its strong language understanding, widespread use, frequent updates, and prominence in research.

DeepSeek-V2 (DeepSeek AI, China) was selected to represent Eastern AI development owing to its rising popularity and notable performance gains.

Gemini (PaLM 2 Pro/Ultra) (Google, USA): Known for advanced search integration and robust data processing, backed by Google's AI expertise.

Table 1. Acute pulmonary embolism case-based assessment and scoring criteria

Questions	Predefined answer key
Case scenario A 51-year-old female patient was transferred to our emergency department from another healthcare facility with a preliminary diagnosis of syncope of unknown etiology. Her husband reported that, after a 20-hour flight from the United States, she initially experienced mild shortness of breath at the airport, which was followed by a complete loss of consciousness. Since he was holding her at that moment, no fall or trauma occurred. The loss of consciousness lasted 3–4 minutes, without excessive salivation or secretion, tongue biting, or uncontrolled convulsive movements of the limbs. The husband reported that his wife had no history of epilepsy or previous episodes of unconsciousness.	
Question 1: Based on the patient's presenting complaint, what are your differential diagnoses?	If the most probable diagnosis is pulmonary embolism (PE) → **5 points** If any of the following additional possible diagnoses are included, add **1 point** for each: Vasovagal syncope (21%) → **1 point** Cardiac syncope (10%) → **1 point** Orthostatic syncope (9%) → **1 point** Seizure (5%) → **1 point** Neurological causes (4.1%) → **1 point** **Maximum possible score: 10 points**
Case scenario continued The patient's family history is notable for venous thromboembolism in the mother. In terms of medical history, the patient has been using oral contraceptives for the past six years. On physical examination, the patient appeared to be in moderate condition. She is confused, partially oriented, and partially cooperative. Her vital signs indicate hemodynamic instability, with a blood pressure of 85/50 mmHg, a heart rate of 128 bpm, a respiratory rate of 30 breaths per minute, and an oxygen saturation of 88% on room air. Further examination reveals jugular venous distension and significant bilateral lower extremity edema. Additionally, there is notable tenderness in the right calf. Electrocardiographic evaluation demonstrates sinus tachycardia and an incomplete right bundle branch block.	
Question 2: Based on the patient's history, family history, and physical examination findings, what are the most likely preliminary diagnoses? Explain.	If the primary preliminary diagnosis is PE → **4 points** If the response includes relevant risk factors (e.g., prolonged immobilization, long-haul flight, oral contraceptive use) → **2 points** If the response includes a calculated Wells score → **4 points** **Maximum possible score: 10 points**
Question 3: Which laboratory tests would you order to confirm your diagnosis? Additionally, which tests would help assess the patient's prognosis and risk of mortality? Provide a rationale for your choices.	If the answer includes B-type natriuretic peptide (for right ventricular dysfunction and prognosis in PE) → **3 points** If the answer includes Troponin (for myocardial injury and right ventricular strain assessment) → **3 points** If the answer includes arterial blood gas (to assess hypoxemia, hypercapnia, respiratory alkalosis, and elevated lactate as a marker of tissue hypoxia and hemodynamic instability in PE) → **2 points** If the answer includes any of the following additional tests, add 0.5 points each: Hemogram (to assess anemia, leukocytosis, thrombocytopenia, or polycythemia) → **0.5 points** Biochemistry panel (to evaluate renal function, liver function, and metabolic status) → **0.5 points** Coagulation tests (PT, aPTT, INR) (to assess clotting status and risk of coagulopathy) → **0.5 points** D-dimer (for ruling out low-risk PE) → **0.5 points** **Maximum possible score: 10 points**
Question 4: Which imaging studies would you order to confirm your diagnosis? Explain your choices.	If the answer includes cardiac echocardiography (for assessing right ventricular strain, pulmonary hypertension, and indirect signs of PE) → **5 points** If the answer includes pulmonary computed tomography (CT) angiography (as the gold standard for confirming PE) → **3 points** If the answer includes pulmonary ventilation-perfusion scan (alternative in patients with contraindications to contrast-enhanced CT) → **1 point** If the answer includes chest X-ray (to rule out alternative diagnoses like pneumonia, pneumothorax, or pulmonary edema) → **1 point** **Maximum possible score: 10 points**
Case scenario continued Arterial blood gas analysis revealed respiratory alkalosis accompanied by hypoxemia. A bedside chest X-ray showed no signs of pulmonary edema or pneumonia, and neither Hampton's hump nor Westermark's sign was detected. The radiographic findings were evaluated as normal. Transthoracic echocardiography demonstrated right ventricular dysfunction and dilation, with a positive McConnell's sign. Additionally, tricuspid annular plane systolic excursion was found to be decreased, indicating right ventricular impairment. CT pulmonary angiography revealed massive bilateral thrombi in the right and left main pulmonary arteries, confirming the diagnosis of acute PE.	
Question 5: Which risk stratification tools would you use to assess this patient, and how do they contribute to PE management? Apply an appropriate tool to evaluate the severity of this patient's condition, calculate the corresponding score, and determine the risk category.	If the answer includes the pulmonary embolism severity index (PESI) as the risk stratification tool → **4 points** If the answer includes a correct PESI score calculation → **4 points** If the answer correctly classifies the patient into the appropriate risk category based on the PESI score → **2 points** (The patient's PESI score was calculated based on various clinical parameters. Because the patient was older than 65 years, 10 points were added to the score. Being female did not contribute any additional points. Systolic blood pressure below 100 mmHg contributed 30 points, while heart rate exceeding 110 bpm added 20 points. Additionally, a respiratory rate greater than 30 breaths per minute and oxygen saturation below 90% each contributed 20 points. The cumulative PESI score for this patient was calculated to be 140, categorizing her as high risk for PE). **Maximum possible score: 10 points**

Table 1. Continued

Questions	Predefined answer key
Question 6: Describe the stepwise clinical management of this patient in the emergency department. Outline the necessary interventions in order of priority, explaining the rationale for each step.	If the answer includes oxygen therapy for respiratory support → **2 points** If the answer includes IV fluid resuscitation (as appropriate for hemodynamic stabilization) → **2 points** If the answer includes vasopressors (for patients with shock or persistent hypotension) → **2 points** If the answer includes specific treatment targeting the diagnosis (anticoagulation therapy for PE) → **4 points** **Maximum possible score: 10 points**
Question 7: What is the appropriate dosing regimen and administration protocol for thrombolytic therapy in this patient?	If the answer includes alteplase (rtPA) at the standard dose of 100 mg over 2 hours → **6 points** If the answer includes the weight-based alternative alteplase regimen (0.6 mg/kg over 15 min, max 50 mg) → **4 points** If the answer includes the streptokinase regimen (250,000 IU loading dose over 30 min, followed by 100,000 IU/h for 12-24 hours) → **3 points** Alternative accelerated regimen: 1.5 million IU over 2 hours → **3 points** If the answer includes the urokinase regimen (4,400 IU/kg loading dose over 10 min, followed by 4,400 IU/kg/h for 12-24 hours) → **3 points** Alternative accelerated regimen: 3 million IU over 2 hours → **3 points** **Maximum possible score: 10 points**
Question 8: What clinical changes would you expect after thrombolytic therapy? How would you assess treatment success?	If the answer includes hemodynamic and clinical improvements (e.g., resolution of hypotension, improved oxygenation, symptom relief, reduced respiratory distress) → **4 points** If the answer includes echocardiographic changes (e.g., reduction in right ventricular strain, improved right ventricular function, decreased pulmonary artery pressure) → **4 points** If the answer includes ECG changes (e.g., resolution of S1Q3T3 pattern, T-wave normalization, reduction in right heart strain signs) or other supportive clinical markers → **2 points** **Maximum possible score: 10 points**
Case scenario continued Following thrombolytic therapy, the patient's hemodynamic status improved significantly. Her blood pressure stabilized at 110/70 mmHg, oxygen saturation increased to 95%, and heart rate decreased to 105 bpm. Given the initial high-risk presentation, the patient was admitted to the intensive care unit for close monitoring. Over the following days, her hemodynamic parameters remained stable, and repeat echocardiography showed improved right ventricular function. Once clinically stable, she was transferred to the inpatient ward for further management.	
Question 9: What tests should be performed before discharge, and what secondary prevention strategies and lifestyle modifications should be implemented?	If the answer includes lower extremity venous Doppler (to assess for residual DVT or ongoing thrombotic risk) → **2 points** If the answer includes a thrombophilia panel (to evaluate for inherited or acquired thrombophilia in selected patients) → **2 points** If the answer includes modifying or discontinuing oral contraceptive use (if applicable) → **2 points** If the answer includes prevention of prolonged immobilization (e.g., avoiding long sedentary periods, early ambulation, compression stockings if needed) → **2 points** If the answer includes regular exercise as part of secondary prevention → **1 point** If the answer includes weight loss (if the patient is overweight or obese) as a risk-reduction strategy → **1 point** **Maximum possible score: 10 points**
Question 10: How should this patient's discharge treatment plan be structured?	If the answer includes initiating low molecular weight heparin (e.g., dalteparin) for transition to oral anticoagulation (class 1, level A) → **3 points** If the answer includes long-term oral anticoagulation therapy (e.g., apixaban) (class 1, level A) → **3 points** If the answer includes the appropriate apixaban dosing regimen (first 7 days: 10 mg twice daily, then maintenance 5 mg twice daily) (class 1, level B) → **2 points** If the answer includes follow-up thoracic angiography to assess resolution of pulmonary artery thrombus (class 2a, level C) → **2 points** **Maximum possible score: 10 points**

This table evaluates the clinical reasoning of AI models across ten domains, including differential diagnosis, risk stratification, diagnostic testing, imaging, treatment, and follow-up. Responses adhering to guideline-based approaches are prioritized, emphasizing the importance of structured diagnosis, risk assessment, evidence-based treatment, and secondary prevention strategies

PT: prothrombin time, aPTT: activated partial thromboplastin time, INR: international normalized ratio, IV: intravenous, rtPA: recombinant tissue plasminogen activator, ECG: electrocardiography, DVT: deep vein thrombosis

Grok (X AI, USA): Included for its unique data sources, rapid growth via social media integration, and innovative approach under Elon Musk's X Corp.

Prompt Structure, Standardization and Data Collection

All models were accessed on March 10, 2025, using an anonymous user account to avoid any personalization or model adaptation based on prior interactions. Before initiating the formal assessment, a standardized role-conditioning prompt was applied to each model: "I am a physician. Please evaluate my questions as a physician and provide guideline-based clinical reasoning." This prompt was used to ensure a consistent baseline response style aligned with clinical decision-making.

Following this initial conditioning, the ten PE case–based questions were posed sequentially within the same conversation thread for each LLM, to maintain contextual continuity and to ensure that the models interpreted the scenario in a manner similar to real-world clinical reasoning.

To enhance reproducibility, the prompt engineering process was standardized across all LLMs. The same wording, order, and contextual flow were used for each model. No additional hints, sub-prompts, or clarifying questions were issued to the models beyond the predefined case scenario and the ten structured questions.

Each model received the same clinical vignette and follow-up questions without deviation. Importantly, no external information (e.g., computed tomography images) was uploaded or provided; instead, all radiological and laboratory findings were described textually to ensure consistent interpretation across platforms. All responses were saved immediately after generation and anonymized for evaluation.

Evaluation Process and Expert Review

The anonymized LLM responses and the predefined answer key were provided to two independent emergency medicine specialists, each with over 5 years of experience, who served as expert reviewers. Each expert was instructed to:

• Score each AI-generated response using the 10-point scoring system, based on adherence to evidence-based guidelines.

• In addition to quantitative scoring, reviewers were asked to comment qualitatively on clarity, structure, and clinical actionability. Although these qualitative dimensions were not included in the numerical scoring system, they were documented to enrich the interpretation of model performance.

Expert assessments were not limited to specific aspects of LLM-generated responses but rather aimed to capture broad observations and insights regarding their potential role in patient management.

Statistical Analysis

The collected expert ratings were compiled in Microsoft Excel and subsequently analyzed using IBM SPSS (latest version, IBM Corp., Armonk, NY, USA). Statistical analyses included descriptive (means, standard deviations, and frequency distributions) to summarize LLM performance. Normality was assessed using the Kolmogorov-Smirnov test. Inter-rater reliability was assessed using the intraclass correlation coefficient (ICC) to ensure consistency in scoring across expert reviewers. A P value < 0.05 was considered statistically significant. All statistical procedures followed standard methodological guidelines to ensure the reliability and reproducibility of the findings.

RESULTS

For analytical purposes, the evaluation questions were classified into four categories: Diagnosis and Initial Evaluation (Q1–Q4), Risk Stratification and Prognosis (Q5), Management and Treatment (Q6–Q7), and Post-Treatment Assessment and Follow-Up (Q8–Q10).

ChatGPT-4o achieved the highest overall score (76), followed by Gemini (73.75), Grok (71.25), and DeepSeek-V2 (65). In the Diagnosis and Initial Evaluation category, all models exhibited comparable performance, with Q3 receiving the highest score (9.5 for both ChatGPT-4o and DeepSeek-V2). In the Risk Stratification and Prognosis category, DeepSeek-V2 obtained the lowest score of 5, whereas the other models obtained scores of 7. In the Management and Treatment category, Gemini outperformed ChatGPT-4o in Q7 (8 vs. 5); in Q6, ChatGPT-4o, Gemini, and Grok attained the highest possible score (10), whereas DeepSeek-V2 scored lower (7). In the Post-Treatment Assessment and Follow-Up category, ChatGPT-4o demonstrated superior performance, particularly in Q9, where it obtained the highest score (10), while Gemini had the lowest (6.5). The most pronounced performance discrepancy was observed in Q10, where ChatGPT-4o (4.5) and DeepSeek-V2 (4) outperformed Gemini (3) and Grok (1.5). 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 categorization of questions, the average scores assigned to each AI model, and the total scores are presented in Table 2.

Overall, the models' performance trends are variable: some models consistently achieve higher scores, while others demonstrate noticeable fluctuations. ChatGPT-4o reaches peak performance on certain questions (e.g., Q6 and Q9) but shows a significant decline in performance on Q7 and Q10. DeepSeek-V2 generally maintains a more stable but lower score range, with its lowest performance on Q5 followed by a gradual recovery in subsequent questions. Gemini demonstrates relatively stable performance, peaking at Q6 but exhibiting a decline towards the final questions. Grok follows a trend similar

Table 2. Performance comparison of medical AI models in acute pulmonary embolism assessment

	Question	ChatGPT-4o	DeepSeek-V2	Gemini	Grok
Diagnosis and initial evaluation	**Q1**	8	6.5	8	8
	Q2	6	6	6	6
	Q3	9.5	9.5	9.25	8.75
	Q4	8	8	8	8
Risk stratification and prognosis	**Q5**	7	5	7	7
Management and treatment	**Q6**	10	7	10	10
	Q7	5	6	8	6
Post-treatment assessment and follow-up	**Q8**	8	6	8	8
	Q9	10	7	6.5	8
	Q10	4.5	4	3	1.5
	Score	**76**	**65**	**73.75**	**71.25**

The table highlights the performance scores for each model across these categories, demonstrating variability in their responses and overall effectiveness in pulmonary embolism assessment

This table presents the evaluation of four AI models—ChatGPT-4o, DeepSeek-V2, Gemini, and Grok—across ten questions, covering domains such as diagnosis, risk stratification, management, treatment, and post-treatment follow-up

Q: question, AI: artificial intelligence

to Gemini's but exhibits slightly greater variability, achieving higher scores on Q1 and Q6 while showing a sharp drop on Q10.

The interrater reliability between the two independent expert reviewers was assessed by calculating the ICC. The results indicated excellent agreement between the reviewers, with a single-measure ICC of 0.986 [95% confidence interval (CI): 0.975–0.992; $P < 0.001$] and an average-measure ICC of 0.993 (95% CI: 0.987–0.996; $P < 0.001$). The distribution of the average scores assigned by the expert reviewers to the AI-generated responses is visually represented in Figure 1.

Table 3 provides a qualitative summary of the expert reviewers' feedback, highlighting the strengths and weaknesses of each AI model. These findings underscore the variability in AI models' performance in clinical decision-making, demonstrating that while some models excel in structured and evidence-based responses, others offer advantages in risk stratification, accessibility, or practical applicability.

DISCUSSION

This study aimed to evaluate the performance of four popular AI LLMs in clinical decision-making based on the 2019 ESC guidelines for the management of PE. Investigating the potential of AI systems as clinical decision-support tools is important for accelerating and enhancing healthcare professionals' decision-making. The results demonstrated that while each AI model exhibited strong performance in specific domains, no model showed a clear overall superiority. ChatGPT-4o achieved the highest total score, whereas DeepSeek-V2 received the lowest total score. However, the varying performances of different AI models across clinical decision-making scenarios provide valuable insights into the strengths and limitations of each system.

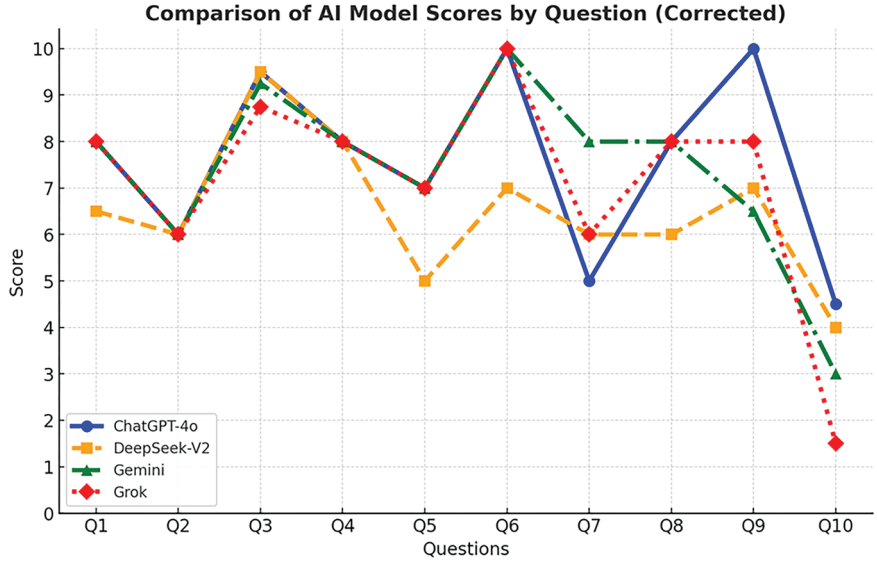

Figure 1. Comparison of AI models' scores on different questions

Comparison of AI models' scores on individual questions related to acute pulmonary embolism management. The line graph illustrates the performance of four AI models—ChatGPT-4o, DeepSeek-V2, Gemini, and Grok—across 10 structured questions (Q1–Q10). Each model's response was scored based on clinical accuracy, guideline adherence, clarity, and completeness

Q: question, AI: artificial intelligence

Table 3. Expert review of AI models in acute pulmonary embolism management

AI models	The expert reviewers' feedback
ChatGPT	ChatGPT was recognized for its comprehensive responses that adhered to guidelines, effectively addressing PE management in a structured, systematic manner. However, its emphasis on individualized treatment approaches was noted to be insufficient
DeepSeek	DeepSeek demonstrated notable proficiency in risk stratification and in clinical decision support, particularly in diagnostic test selection and stepwise management. Nevertheless, its explanations of thrombolytic therapy and post-discharge follow-up were deemed insufficient and required further elaboration
Gemini	Google Gemini provided fluent, accessible explanations consistent with general clinical practice, yet it lacked sufficient depth in risk stratification and in certain laboratory investigations
Grok	Grok was recognized for its practical and point-of-care management recommendations; however, its lack of a systematic approach and failure to provide direct references to established guidelines were identified as major limitations

Expert review of various AI models used in the management of acute pulmonary embolism. The table summarizes expert reviewers' evaluations regarding each model's strengths and limitations in areas such as diagnostic support, adherence to guidelines, risk stratification, and individualized treatment recommendations
AI: artificial intelligence, PE: pulmonary embolism

The variability in model performance likely reflects a combination of factors, including differences in model design, training objectives, and the general composition of training data; however, the exact sources and weighting of medical versus non-medical content are not publicly disclosed for these systems. Therefore, any explanation at the level of specific data sources or internal optimization strategies remains speculative. In this context, our findings are best interpreted as empirical evidence that each LLM exhibits domain-specific strengths and weaknesses, rather than as direct reflections of their proprietary training pipelines.

The literature encompasses a wide range of studies on the use of AI-based systems in the healthcare sector. It emphasizes that the use of AI in the medical field is rapidly increasing and may significantly enhance the efficiency of clinical decision support systems.[3,7-9] Our findings are consistent with previous studies showing that LLMs can generate clinically relevant, structured answers. However, the domain-specific strengths and weaknesses observed in each model underscore that LLMs should not be assumed to perform uniformly across all phases of patient management. From a clinical perspective, our results suggest that different LLMs may be better suited to particular components of PE management. For instance, DeepSeek-V2 demonstrated relative strength in diagnostics, Gemini performed strongly in treatment planning, and ChatGPT-4o excelled in post-treatment follow-up and in interpreting structured guidelines. These findings indicate that physicians should be aware of model-specific performance patterns rather than relying on a single system for all components of clinical care.

AI technology is evolving rapidly and becoming increasingly integrated into healthcare.[8] Continuous updates enhance AI performance in medical evaluation, as demonstrated in a gastroenterology study in which AI effectively managed real-world, guideline-based patient scenarios.[10] Studies investigating the medical reliability of publicly available LLMs suggest that these systems generate clinically relevant and guideline-compliant outputs. Our results support this conclusion by demonstrating that LLMs may provide reliable assistance in certain tasks, yet they also reveal that LLM-generated responses remain vulnerable to omissions, insufficient detail, or incomplete integration of risk stratification criteria in other tasks. Importantly, none of the evaluated models provided fully comprehensive or consistently guideline-aligned recommendations, reinforcing that LLMs should function as supportive tools rather than autonomous decision-makers. In a study evaluating the accuracy and clarity of patient education provided by three AI models on atrial fibrillation and cardiac implantable electronic devices, the responses generated by the AI systems were, on average, over 90% accurate and understandable.[3] In another AI study focusing on patient education regarding hypertension, the responses provided by ChatGPT to 100 questions were found to be appropriate and accurate in 92.5% of the cases.[7] In another study comparing two versions of ChatGPT in answering questions related to heart failure, GPT-3.5 achieved over 94% accuracy, while GPT-4 scored 100% across all 107 questions, demonstrating the high accuracy of these AI models in clinical decision-making for heart failure.[9] Our study supports these findings and provides valuable insights into how AI models can be applied to the management of critical medical conditions, such as PE.

Applying AI to PE assessment within a guideline-based framework offers valuable insight into its potential role in patient management. Additionally, studies examining healthcare professionals' perspectives on AI adoption indicate that these models could function as reliable clinical decision-support tools.[11,12] Beyond publicly available AI models, machine learning algorithms have been developed to improve acute PE diagnosis through electrocardiogram (ECG)-based analysis, further expanding AI's role in clinical medicine. In a study evaluating 1,014 ECGs obtained from patients admitted to the emergency department who underwent pulmonary computed tomography angiography for suspected PE, the AI model demonstrated greater specificity for detecting PE than commonly used prediction rules. The AI model achieved 100% specificity and 50% sensitivity.[13] Our study is consistent with these findings and provides important data on how AI models can be used to manage critical medical conditions such as PE. In this context, ChatGPT-4o was observed to produce highly accurate results, demonstrating a structured approach aligned with medical guidelines. However, some AI models, especially DeepSeek-V2, were found to produce less accurate responses or to provide insufficient coverage of specific treatment steps.

Given the life-threatening nature of PE, AI appears promising for real-world applications in emergency medicine. Integrating AI into medical decision-support systems may help mitigate healthcare workforce shortages, particularly in resource-limited settings.[14] While this study presents findings similar to those of those reported in the literature, it also highlights certain differences. For example, earlier studies have emphasized that AI generally contributes to diagnostic and therapeutic processes, though in some cases, clinical decision-making still requires significant human oversight.[8-10] In our study, the shortcomings of DeepSeek-V2, particularly in risk stratification and treatment processes, reveal notable limitations in the extent to which AI can be integrated into medical decision-making.

In this study, each AI model showed distinct strengths and weaknesses across different aspects of PE management. Below is a summary of their performance based on expert evaluations and guideline adherence:

ChatGPT-4o achieved the highest overall score, excelling in diagnosis and initial assessment. However, its limited ability to provide personalized treatment plans highlights the need for improved patient-specific adaptability—especially important in emergency settings.

DeepSeek-V2 showed stable but generally low performance, particularly in risk stratification and lab interpretation. While it performed well on some individual questions, its limited coverage of treatment and follow-up reduces its clinical utility.

Gemini stood out in treatment management, offering well-structured therapeutic recommendations. However, it had difficulty with risk assessment and interpretation of lab results. These gaps may be addressed in future updates, given Google's resources and AI capabilities.

Grok provided practical, point-of-care suggestions but lacked a clear structure and direct references to guidelines, thereby limiting its scientific reliability. Its wide reach via social media is a strength, but better integration with evidence-based content is needed to enhance clinical applicability.

In our study, AI models demonstrated varying strengths and limitations in the management of PE, as reflected in expert evaluations. ChatGPT was praised for providing comprehensive management due to its structured and guideline-compliant approach; however, its lack of focus on individualized treatment was noted as a drawback. DeepSeek performed well in risk stratification and diagnostic test selection, enhancing clinical decision-making, but its limited coverage of thrombolytic therapy and post-discharge follow-up represented a limitation. Gemini provided clear, accessible explanations and aligned with general clinical practice; nevertheless, it lacked depth in risk assessment and in certain laboratory evaluations critical to PE management. Grok was found to be useful in offering practical, point-of-care recommendations; however, the absence of a structured approach and of direct references to guidelines were cited as major limitations. These findings highlight that although AI models can provide valuable clinical insights, their performance varies across different aspects of PE management, underscoring the need for model refinement, guideline integration, and a more personalized approach.

For future research, more robust evaluation strategies are recommended, including multi-scenario studies, comparative assessments across different sets of clinical guidelines, and investigations of combined or ensemble approaches that leverage the strengths of multiple LLMs. Additionally, prospective studies examining how LLM support influences physician decision-making accuracy, workflow efficiency, and patient outcomes would provide meaningful insights into real-world applicability. Frameworks conceptualized as "physician–LLM collaboration models" may also help define safer and more effective integration pathways.

Study Limitations

Several considerations should guide the interpretation of this study. The analysis was based on a single clinical scenario, which enabled standardized comparison but naturally limits the breadth of clinical contexts to which the findings can be generalized. Different presentations may challenge LLMs in distinct ways.

Although all models were evaluated on the same day and under uniform conditions, LLMs evolve rapidly, and their performance reflects a moment in time rather than a stable characteristic. Future updates may alter their reasoning patterns.

As with all generative models, the risk of factual distortion or overconfident statements remains inherent. Our structured scoring system reduced this risk but could not completely eliminate it.

Finally, while quantitative scoring allowed for reproducible evaluation, the absence of dedicated qualitative metrics—such as omission tracking or actionability—represents a conceptual limitation that future studies may address.

CONCLUSION

In our study, each model demonstrated distinct strengths across the diagnostic, treatment, and follow-up processes; however, performance fluctuations were particularly notable in areas such as personalized patient management and risk assessment. Although AI models show promising potential as clinical decision support tools, they should be further trained with real-world patient data to enhance adherence to clinical guidelines and better align with the principles of personalized medicine. In this context, they should be developed not to replace physicians but to support clinical decision-making.

Ethics

Ethics Committee Approval: Since our research did not involve any human participants and did not include any patient records, no application to an ethics committee was submitted. Nevertheless, throughout all stages of the study, the principles of scientific research and publication ethics were strictly observed. Should any additional information or clarification be required, we are prepared to provide detailed information.

Informed Consent: This study is a simulation-based investigation and does not involve human participants or the use of real patient data.

Artificial Intelligence Usage Statement

Artificial intelligence (AI) models, including ChatGPT-4o, Gemini, Grok, and DeepSeek-V2, were used solely as subjects in the comparative evaluation. No generative AI tool was used to create or edit the manuscript content. All analyses, interpretations, and manuscript writing were performed by the authors.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ö.F.K., H.E.K., M.E.Ö., Y.G., Concept: Ö.F.K., H.E.K., M.E.Ö., B.Y., Design: Ö.F.K., H.E.K., M.E.Ö., B.Y., Data Collection or Processing: Ö.F.K., H.E.K., Ö.H.S., M.E.Ö., Analysis or Interpretation: Ö.F.K., H.E.K., Ö.H.S., M.E.Ö., Y.G., Literature Search: Ö.F.K., H.E.K., Ö.H.S., Y.G., Writing: Ö.F.K., H.E.K., B.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. Ghorashi N, Ismail A, Ghosh P, Sidawy A, Javan R. AI-powered Chatbots in medical education: potential applications and implications. *Cureus*. 2023;15(8):e43271. **[Crossref]**
2. Niko MM, Karbasi Z, Kazemi M, Zahmatkeshan M. Comparing ChatGPT and Bing, in response to the Home Blood Pressure

Monitoring (HBPM) knowledge checklist. *Hypertens Res*. 2024;47(5):1401-1409. **[Crossref]**

3. Hillmann HAK, Angelini E, Karfoul N, Feickert S, Mueller-Leisse J, Duncker D. Accuracy and comprehensibility of chat-based artificial intelligence for patient information on atrial fibrillation and cardiac implantable electronic devices. *Europace*. 2023;26(1):euad369. **[Crossref]**
4. Hernandez CA, Vazquez Gonzalez AE, Polianovskaia A, et al. The future of patient education: AI-driven guide for type 2 diabetes. *Cureus*. 2023;15(11):e48919. **[Crossref]**
5. Cascella M, Semeraro F, Montomoli J, Bellini V, Piazza O, Bignami E. The breakthrough of large language models release for medical applications: 1-year timeline and perspectives. *J Med Syst*. 2024;48(1):22. **[Crossref]**
6. Konstantinides SV, Meyer G, Becattini C, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). *Eur Heart J*. 2020;41(4):543-603. **[Crossref]**
7. Almagazzachi A, Mustafa A, Eighaei Sedeh A, et al. Generative artificial intelligence in patient education: ChatGPT takes on hypertension questions. *Cureus*. 2024;16(2):e53441. **[Crossref]**
8. Holzinger A, Keiblinger K, Holub P, Zatloukal K, Müller H. AI for life: trends in artificial intelligence for biotechnology. *New Biotechnol*. 2023;74:16-24. **[Crossref]**
9. King RC, Samaan JS, Yeo YH, Mody B, Lombardo DM, Ghashghaei R. Appropriateness of ChatGPT in answering heart failure related questions. *Heart Lung Circ*. 2024;33(9):1314-1318. **[Crossref]**
10. Sciberras M, Farrugia Y, Gordon H, et al. Accuracy of information given by ChatGPT for patients with inflammatory bowel disease in relation to ECCO guidelines. *J Crohns Colitis*. 2024;18(8):1215-1221. **[Crossref]**
11. Vearrier L, Derse AR, Basford JB, Larkin GL, Moskop JC. Artificial intelligence in emergency medicine: benefits, risks, and recommendations. *J Emerg Med*. 2022;62(4):492-499. **[Crossref]**
12. Sauerbrei A, Kerasidou A, Lucivero F, Hallowell N. The impact of artificial intelligence on the person-centred, doctor-patient relationship: some problems and solutions. *BMC Med Inform Decis Mak*. 2023;23(1):73. **[Crossref]**
13. Valente Silva B, Marques J, Nobre Menezes M, Oliveira AL, Pinto FJ. Artificial intelligence-based diagnosis of acute pulmonary embolism: development of a machine learning model using 12-lead electrocardiogram. *Rev Port Cardiol*. 2023;42(7):643-651. English, Portuguese. **[Crossref]**
14. Long P, Lu L, Chen Q, Chen Y, Li C, Luo X. Intelligent selection of healthcare supply chain mode - an applied research based on artificial intelligence. *Front Public Health*. 2023;11:1310016. **[Crossref]**

Thorac Res Pract. 2026;27(1):47-56 **DOI:** 10.4274/ThoracResPract.2025.2025-2-6

Review

Breaking: The New 9th Version TNM Classification for Lung Cancer is Now in Use

Yusuf Kahya[1], Feride Marım[2], Gökçen Ömeroğlu Şimşek[3], Haydar Soydaner Karakuş[4], Sevin Başer Öncel[5], Müge Meltem Tor[6], Tuncay Göksel[4]

[1]Department of Thoracic Surgery, Ankara University Faculty of Medicine, İbni Sina Hospital, Ankara, Türkiye
[2]Department of Pulmonology, Kütahya Health Sciences University Faculty of Medicine, Kütahya, Türkiye
[3]Department of Pulmonology, Dokuz Eylül University Faculty of Medicine, İzmir, Türkiye
[4]Department of Pulmonology, Ege University Faculty of Medicine, İzmir, Türkiye
[5]Department of Pulmonary Medicine, Pamukkale University Faculty of Medicine, Denizli, Türkiye
[6]Department of Pulmonology, Zonguldak Bülent Ecevit University Faculty of Medicine, Zonguldak, Türkiye

Cite this article as: Kahya Y, Marım F, Ömeroğlu Şimşek G, et al. Breaking: the new 9th version TNM classification for lung cancer is now in use. *Thorac Res Pract.* 2026;27(1):47-562026;27(1):47-56

ABSTRACT

Lung cancer is the leading cause of cancer-related mortality worldwide. Among the parameters determining prognosis in lung cancer, the stage of the disease holds primary importance. Staging provides a universally accepted terminology for describing the anatomical characteristics of cancer, facilitating reliable communication in clinical research, evaluation of treatment outcomes, and prognosis. The tumor, node, metastasis (TNM) staging system, developed by the American Joint Committee on Cancer (AJCC) and the International Union for Cancer Control (UICC), serves as a simple, practical, and globally recognized staging framework. Over the past two decades, the International Association for the Study of Lung Cancer has been conducting a global three-phase project aimed at revising the TNM classification. The first two phases of this project were focused on revising the 7th and 8th lung cancer TNM staging revisions under the guidance of AJCC and the UICC. The third and final phase, the 9th staging project, has been completed and has been implemented as of January 1, 2025. This review aims to examine the 9th version of the TNM staging system compared to previous versions and evaluate the structural modifications, statistical foundations, and clinical implications of the new system. In the study, current data regarding the 9th version of the TNM staging system have been analysed; the revisions made to the T, N, and M components are detailed; the fundamental changes between the 8th and 9th versions are compared using tables. Furthermore, the impacts of the staging system on daily clinical practice are discussed.

KEYWORDS: Lung and pleural malignancies, lung cancer, TNM classification

Received: 27.02.2025 ***Revision Requested:*** 14.04.2025 ***Last Revision Received:*** 14.06.2025

Accepted: 29.07.202 ***Epub:*** 18.09.2025 ***Publication Date:*** 30.01.2026

INTRODUCTION

Lung cancer is the leading cause of cancer-related mortality worldwide for both women and men.[1] Among the parameters determining prognosis in lung cancer, the stage of the disease holds primary importance.[2] Staging provides a straightforward and universally accepted terminology for describing the anatomical characteristics of cancer, facilitating reliable communication in clinical research, evaluation of treatment outcomes, and prognosis.[3] The tumor, node, metastasis (TNM) staging system, developed by the American Joint Committee on Cancer (AJCC) and the International Union for Cancer Control (UICC), serves as a simple, practical, and globally recognized staging framework.[3]

The origins of the TNM staging system currently in use can be traced back to the 1940s. This system, originally developed by the French surgeon Pierre Denoix during that decade, was adopted by the UICC in 1953, with its first official version published in 1958.[4]

Corresponding author: Gökçen Ömeroğlu Şimşek, Asst. Prof. MD, e-mail: gokcen.simsek@deu.edu.tr

The TNM system stages cancers based on three primary anatomical parameters: the size and extent of the primary tumor (T), the involvement of regional lymph nodes (N), and the presence of distant metastasis (M). Each T, N, and M component is further classified into multiple components (e.g., T1, T2, N1, N2) and subcomponents (e.g., T1a, T1b, T1c). Specific combinations of these T, N, and M components are grouped into stage classifications that share similar prognostic outcomes, culminating in the final staging designation.[3] However, it is crucial to recognise that due to the heterogeneity in the biological behaviour of tumor, cases with identical T, N, and M classifications may exhibit differing prognoses and treatment responses.

To specify the context of TNM staging, certain prefixes are utilised. Clinical staging (cTNM) is determined based on all available information prior to surgical resection (e.g., symptoms, physical examination findings, imaging studies, biopsies). Pathological staging (pTNM) incorporates additional data obtained through pathological evaluation following surgical resection, supplementing the information from clinical staging. The accuracy of staging is directly related to the concordance between clinical and pathological staging. Restaging (yTNM) is performed following partial or complete treatment and includes stages such as post-treatment clinical staging (ycTNM) and post-treatment pathological staging (ypTNM). The stage determined at the time of recurrence is referred to as rTNM, while staging identified during autopsy is termed aTNM.[5]

Recent advancements in immunotherapy and targeted therapies have significantly altered diagnostic and therapeutic algorithms, thereby underscoring the increasing importance of staging, particularly restaging.

The aim of this review is to provide a comprehensive evaluation of the 9th version of the TNM staging system for lung cancer, which came into effect in 2025. In the study, the development process of the new version, the statistical methods employed and the decision criteria are discussed and a comparative analysis with the 8th edition is presented. Updates to the T, N, and M components are assessed along with their scientific foundations based on survival analyses. Furthermore, the effects of these changes on routine clinical practice, surgical decision-making, and stage-based treatment planning are examined. In addition, the role of artificial intelligence and machine learning methods in the development of the staging system is discussed; the potential of non-anatomic prognostic factors that could be integrated into the TNM staging system in the future is also addressed. Thus, this review aims to present both the structural aspects of the system and its clinical significance and practical applications from multiple perspectives.

Methodology

This review has been prepared to concisely present the most current updates regarding the 9th version of the TNM classification in lung cancer, the rationales underlying these changes, and their potential clinical implications. International guidelines, expert opinions, and original studies that contributed to the development of the 9th version have been brought together. A thorough literature search was conducted across PubMed, Scopus, Web of Science, and Google Scholar databases, covering publications from January 2002 to April 2025. The keywords used in the search are as follows: "TNM staging system," "lung cancer," "IASLC," "TNM 9th version," "stage groupings," "N descriptors," "M descriptors," "T descriptors," "prognosis," and "molecular data in NSCLC."

The studies included in this review were selected based on specific eligibility criteria. First, each study was required to address the staging of non-small cell lung cancer (NSCLC) using the TNM classification system. In addition, these studies were expected to include original data, recommendations, or consensus reports provided by recognized authorities such as the International Association for the Study of Lung Cancer (IASLC), the AJCC, or the UICC. Only publications published in peer-reviewed scientific journals with full-text access in English were included in the evaluation. Editorials, case reports, and conference abstracts presented in summary form only were excluded from the scope of this review.

Primarily, priority was given to the official publications of the IASLC Lung Cancer Staging Project, which contributed to the development of the 9th version. In addition, foundational resources conveying the historical evolution of the staging system, along with earlier editions, were also taken into consideration. To assess the current state, narrative-style reviews and position papers were reviewed. The most recent staging manual published by the AJCC specific to the 9th version, as well as the IASLC's molecular database project, were specifically evaluated regarding the integration of anatomical and molecular prognostic factors.

The data were narratively summarized by being classified according to the TNM components (T, N, and M); then, recommendations and validation studies in stage groups were examined. Emerging trends in the literature; defining oligometastatic disease, subclassification of N2 disease, and integration of molecular biomarkers into prognostic groupings, were also discussed.

The International Association for the Study of Lung Cancer Staging Project

The TNM system in lung cancer undergoes periodic evidence-based revisions. Under the guidance of the AJCC, the TNM system has been revised eight times to date, with a revision cycle of 6–8 years. Over the past two decades, the IASLC has been conducting a global three-phase project aimed at revising the TNM classification. The first two phases of this project were focused on revising the 7th and 8th lung cancer TNM staging editions under the guidance of AJCC and the UICC. The third phase involved collecting data on cases diagnosed with lung cancer between January 2011 and December 2019 to establish the database for the 9th version of staging.[6,7]

To initiate the first phase of the international staging project, the IASLC formed the International Staging Committee [currently known as the Staging and Prognostic Factors Committee (SPFC)] in 1997. Between 1990 and 2000, data on 100,869 lung cancer cases from 45 centres in 20 countries across Europe, North America, Asia, and Australia were submitted to the Cancer Research and Biostatistics (CRAB) database. Subcommittees under SPFC analysed this database, which was managed by

CRAB, to propose revisions to the TNM system. The resulting recommendations were submitted by IASLC to AJCC and UICC, which led to the creation of the 7th edition of the lung cancer TNM staging system, implemented in 2010. This phase was led by British thoracic surgeon Goldstraw et al.[8]

The second phase of the IASLC's international staging project began in 2009 under the leadership of Spanish thoracic surgeon Ramon Rami Porta. Between 1999 and 2010, data on 94,708 lung cancer cases from 35 centres in 16 countries across Europe, Asia, North America, Australia, and South America were submitted to the CRAB database. After analysing this data, the IASLC conveyed its recommendations to AJCC and UICC, resulting in the 8th edition of the lung cancer TNM staging system, which was implemented in 2017.[9]

Main Points

- The 9th tumor, node, metastasis (TNM) classification introduces key structural updates to N and M descriptors for better prognostic accuracy.
- The N2 component is now subdivided into N2a (single-station) and N2b (multi-station). At the same time, the M1c category is split into M1c1 (multiple lesions in one organ) and M1c2 (multiple organs involved), reflecting differences in survival outcomes and guiding personalized treatment planning.
- No major changes were made to the T descriptors themselves; however, additional clarifications were introduced regarding specific anatomical invasion sites. While the overall T staging framework remains consistent with the 8th edition, the current revision provides clearer definitions for invasion of structures such as the azygos vein, vagus nerve, brachial plexus, and thoracic nerve roots, all of which continue to be classified within the T3 or T4 categories.
- Although T staging criteria remain structurally similar to the 8th edition, clarifications were added regarding invasion into structures such as the azygos vein, vagus nerve, brachial plexus, and thoracic nerve roots, with all these maintaining their classifications under T3 or T4.
- The 9th version is based on a global, high-volume, multi-institutional dataset of over 124,000 lung cancer cases.
- For the first time, the International Association for the Study of Lung Cancer's staging project included data from all continents, including underrepresented regions like Africa and the Middle East, thus increasing the global representativeness and robustness of survival models.
- Updated stage groupings reflect new prognostic data and survival distinctions.
- For example, T1N1 was downstaged from IIB to IIA, and new combinations such as T1N2a were included in stage IIB. These refinements were validated through hazard ratio-based survival analyses, demonstrating improved stage discrimination over the 8th edition.
- While still anatomically based, the 9th TNM system incorporates AI modeling and molecular data collection to prepare for future personalized staging.
- Recursive partitioning (a machine learning method) was used in model development, and molecular biomarkers (EGFR, ALK, KRAS, PD-L1) were recorded, laying the groundwork for integrating biological features and AI-driven decision tools in upcoming editions.

The 7th and 8th editions of the TNM staging system differed significantly from previous editions. For instance, the 6th edition, led by U.S. thoracic surgeon Clifton Mountain, relied on data from a single centre that only included cases treated surgically between 1975 and 2002.[10] In contrast, the 7th edition incorporated data from a multinational project that also included cases treated with non-surgical modalities (e.g., chemotherapy, radiotherapy, or chemoradiotherapy).[8]

The third and most recent phase of the IASLC international staging project was conducted under the leadership of Japanese thoracic surgeon Asamura et al.[6] This phase was the most comprehensive in terms of the number of participating centres and the volume of data collected. Notably, for the first time, data from Africa and the Middle East were included, making it the only project representing all continents.[6] Details of the 9th version of the lung cancer TNM staging system will be discussed extensively.

The Ninth Version TNM Classification of Lung Cancer

The third IASLC international staging project was initiated with the aim of improving the anatomical staging system and enhancing its clinical applicability. This project began shortly after the 8th edition of the TNM staging system was implemented. The IASLC emerged as the sole global organisation undertaking this project. Between January 1, 2011, and December 31, 2019, data from 124,581 cases were entered into the database from 78 centres across 25 countries spanning five continents. The data distribution was as follows: Asia and Australia (69,749 cases, 56%), Europe (30,827 cases, 24.7%), North America (19,608 cases, 15.7%), South and Central America (4,225 cases, 3.4%), and Africa and the Middle East (172 cases, 0.1%).[6,11]

Data entry was conducted through batch data submission for 81.1% (101,033 cases) and electronic data capture (EDC) for 18.9% (23,548 cases), resulting in an impressive and comprehensive database. Batch data submission refers to the process of uploading data to a database in bulk at defined intervals (e.g., weekly or monthly), typically via file transfer, either manually or semi-automatically. EDC, on the other hand, involves entering and storing data in real time directly into a digital system, typically through a web-based software or interface.

After excluding cases with missing or erroneous data, 87,043 cases were included in the statistical analysis. Notable data contributors included the Japanese Joint Lung Cancer Registry (Japan, 23,663 cases), Heidelberg University Hospital (Germany, 8,840 cases), West China Hospital at Sichuan University (China, 7,345 cases), the Korean Association for Lung Cancer (South Korea, 4,022 cases), and Samsung Medical Centre (South Korea, 3,645 cases). Türkiye was represented with 1,395 cases (1.1%).[6,11]

Surgical treatment was performed in 67% of the cases in the database. Among 47,933 surgical cases with available surgical margin information, 42,623 cases (88.9%) achieved

R0 resection.[12] Histopathological analysis of the 87,043 cases showed the following distribution: 84% (73,197 cases) non-small cell carcinoma, 6% (5,530 cases) small cell carcinoma, and 10% (8,316 cases) others. When analysed in detail, the distribution is presented as follows: 59.8% (52,069 cases) invasive adenocarcinoma, 18.2% (15,872 cases) squamous cell carcinoma, 1.3% (1,142 cases) adenocarcinoma *in situ*, 1.3% (1,100 cases) adenosquamous carcinoma, 1.2% (1,057 cases) large cell carcinoma, 6.4% (5,530 cases) small cell lung cancer, and 0.8% (689 cases) large cell neuroendocrine carcinoma.[6,11]

Both clinical and pathological data were recorded for 51.5% (44,831 cases), only clinical data for 38.2%, and only pathological data for 10.3% of the cases. Among 77,811 cases with clinical staging data, the most frequent clinical stage was IA2 (10,402 cases, 13.4%) according to the 8th edition of the TNM system. Among 54,248 cases with pathological staging data, the most common pathological stage was IA (22,206 cases, 40.9%).[6]

Future versions of the TNM staging system are expected to incorporate non-anatomical data, particularly molecular biomarkers. Unlike previous projects, this project allowed for the recording of molecular data (e.g., genetic mutations such as EGFR, ALK, ROS1, KRAS, and biomarkers like PD-L1 expression), marking the first step toward developing a molecular database. The cohort for this project included 9,931 (13.6%) patients with available molecular data.[13]

Collaborative efforts among 14 subcommittees within the lung cancer working group (domains) of CRAB and SPFC, including those for T, N, M, neuroendocrine tumors, staging, lepidic/adenocarcinoma *in situ*, lymph node charts, validation methodology, multiple nodules, prognostic factors, R factors, molecular data, imaging, and database management, enabled comprehensive statistical analyses of the data. Additionally, statistical analysis of subgroups has also been conducted. For example, EDC vs. batch datasets, squamous vs. non-squamous carcinoma, region (Asia, Europe, North America, Rest of World), Zubrod performance status (PS: 0 vs. PS≥1), year of diagnosis (2017 or earlier vs. 2018 or later), treatment modality (surgical, non-surgical, neoadjuvant). The resulting recommendations for the TNM staging system were submitted to AJCC and UICC for approval, and the 9th version was officially implemented on January 1, 2025.[6] Details of the changes to the T, N, and M components and stage groups in the 9th version are discussed below.

T component

There have been several updates to T descriptors, differing from the 8th edition of staging. These changes can be summarised as follows: 1. A tumors, with direct invasion of an adjacent lobe, across the fissure or by direct extension at a point where the fissure is deficient, should be classified as T2a unless other criteria assign a higher T component. 2. Invasion of the azygos vein is classified as T3. 3. Invasion of thoracic nerve roots (e.g., T1, T2) or stellate ganglion is classified as T3. 4. Invasion of the thymus is classified as T4. 5. Invasion of subclavian vessels, vertebral body, lamina, spinal canal, cervical nerve roots, or brachial plexus (e.g., trunks, divisions, cords, or terminal nerves) is classified as T4. 6. Invasion of the supra-aortic arteries or brachiocephalic veins is classified as T4. 7. Invasion of the vagus nerve is classified as T4.

When compared to the 8th edition, no changes have been made to the T component. An analysis was conducted to test the hypothesis that the presence of chest wall invasion within the T3 component group could be treated as a distinct descriptor or even reclassified as a T4 descriptor. However, no reproducible survival differences were observed between chest wall invasion and other T3 descriptors in either the clinical or pathological stage.[14,15]

Tx Primary tumor cannot be assessed[a]

T0 No evidence of primary tumor

Tis Carcinoma *in situ*[b]

T1 Tumor surrounded by lung or visceral pleura, or in a lobar or more peripheral bronchus[c]

T1mi Minimally invasive adenocarcinoma[d]

T1a Tumor ≤1 cm in greatest dimension

T1b Tumor >1 cm but ≤2 cm in greatest dimension

T1c Tumor >2 cm but ≤3 cm in greatest dimension

T2

Tumor with any of the following features:

T2a

- Tumor >3 cm but ≤4 cm in greatest dimension;
- Invades visceral pleura;
- Invades an adjacent lobe;
- Involves main bronchus (up to but not including the carina) or is associated with atelectasis or obstructive pneumonitis extending to the hilar region, involving either part of or the entire lung.

T2b Tumor >4 cm but ≤5 cm in greatest dimension

T3

Tumor with any of the following features:

- Tumor >5 cm but ≤7 cm in greatest dimension;
- Invades parietal pleura or chest wall;
- Invades pericardium, phrenic nerve, or azygos vein;[e]
- Invades thoracic nerve roots (ie T1, T2) or stellate ganglion;
- Separate tumor nodule(s) in the same lobe as the primary

T4

Tumor with any of the following features:

- Tumor >7 cm in greatest dimension;

• Invades mediastinum, thymus, trachea, carina, recurrent laryngeal nerve, vagus nerve, esophagus or diaphragm;

• Invades heart, great vessels (aorta, superior/inferior vena cava, intrapericardial pulmonary arteries/veins), supra-aortic arteries, or brachiocephalic veins;

• Invades subclavian vessels, vertebral body, lamina, spinal canal, cervical nerve roots, or brachial plexus (ie trunks, divisions, cords, or terminal nerves);

• Separate tumor nodule(s) in a different ipsilateral lobe than that of the primary

[a]This includes tumors proven by the presence of malignant cells in sputum or bronchial washings but not visualized by imaging or bronchoscopy

[b]This includes adenocarcinoma *in situ* – Tis – and squamous cell carcinoma *in situ* – Tis

[c]The uncommon superficial spreading tumor of any size with its invasive component limited to the bronchial wall, which may extend proximal to the main bronchus, is also classified as T1a

[d]Solitary adenocarcinoma (not more than 3 cm in greatest dimension), with a predominantly lepidic pattern and not more than 5 mm invasion in greatest dimension

[e]Although these structures lie within the mediastinum, the degree of mediastinal penetration by the tumor needed to invade these structures is not counted as T4

N component

The lymph node map developed by IASLC in 2009 was first implemented with the 7th edition of the TNM staging system (Figure 1).[16] This map provides clear anatomical definitions and numbering of lymph node stations, but some areas remain contentious (e.g., the distinction between the right paratracheal and right hilar lymph nodes, the left lower paratracheal and left hilar lymph nodes, and the subcarinal and hilar lymph nodes).

In the 9th version, no changes were made to the lymph node map. However, to aid in the understanding of anatomical markers and minimise misinterpretations of the N component, realistic illustrations and intraoperative photographs have been added to the map.[17,18] These adjustments are significant in reducing stage migration.

Compared to the 8th edition, the N2 component group has been subdivided into two components: N2a and N2b.

• N2a: Refers to metastasis confined to a single ipsilateral mediastinal or subcarinal lymph node station.

• N2b: Indicates metastasis involving multiple mediastinal or subcarinal lymph node stations.

Survival analyses demonstrated a clear and consistent prognostic difference between single and multiple N2 station involvement in both clinical and pathological stages (Table 1).[17,18] This distinction is important as it helps address the heterogeneity within N2 disease (e.g., skip vs. non-skip N2, micrometastatic vs. bulky N2, single vs. multiple zone N2 involvement). However, further subcategorisation of the N2 component group remains necessary.

At the N1 level, no consistent and significant differences were observed between single and multiple station involvement in both clinical and pathological stages. Therefore, subdividing the N1 component group into additional components was not recommended for the 9th version.[17,18]

NX Regional lymph nodes cannot be assessed

N0 No regional lymph node metastasis

N1 Metastasis in ipsilateral peribronchial and/or ipsilateral hilar and/or intrapulmonary lymph nodes, including involvement by direct extension

N2 Metastasis in ipsilateral mediastinal and/or subcarinal lymph node(s)

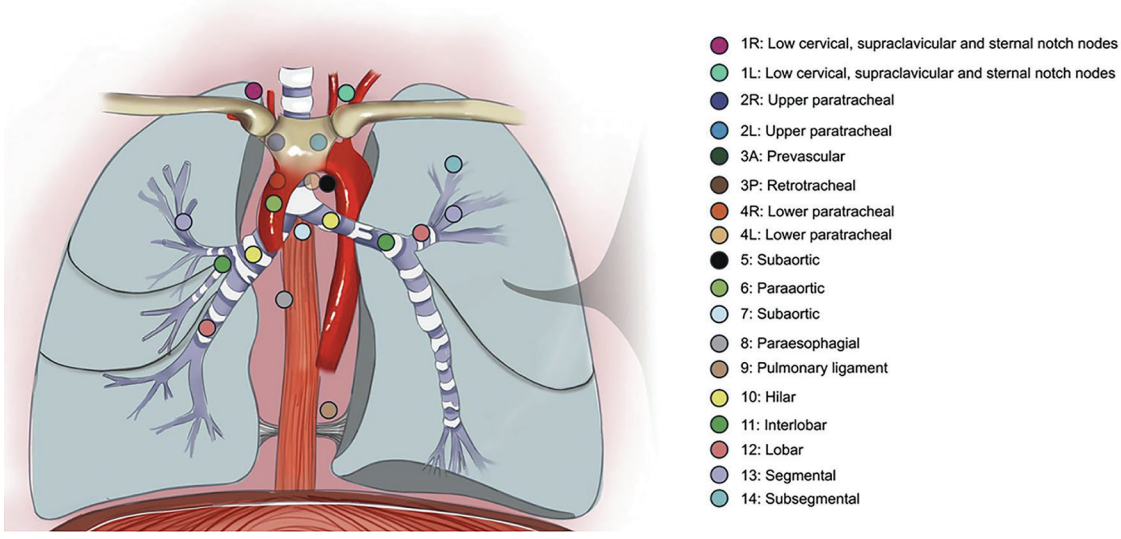

Figure 1. IASLC nodal chart with stations (redrawn by the author inspired by the original figure)

IASLC: International Association for the Study of Lung Cancer

N2a – Single N2 station involvement

N2b – Multiple N2 station involvement

N3 Metastasis in contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene or supraclavicular lymph node(s)

M component

Compared to the 8th edition, the M1c component has been subdivided into two components in the 9th version: **M1c1** and **M1c2** (Table 2).[19]

M0 No distant metastasis

M1 Distant metastasis

M1a Tumor with pleural or pericardial nodules or malignant pleural or pericardial effusions,[a] separate tumor nodule(s) in a contralateral lobe

M1b Single extrathoracic metastasis in a single organ system[b]

M1c Multiple extrathoracic metastases

M1c1 Multiple extrathoracic metastases in a single organ system[c]

M1c2 Multiple extrathoracic metastases in multiple organ systems

[a]Most pleural (or pericardial) effusions in patients with lung cancer are due to the tumor. In a few patients, however, multiple microscopic examinations of pleural (or pericardial) fluid are negative for tumor, and the fluid is non-bloody and is not an exudate. When these elements and clinical judgment dictate that the effusion is not related to the tumor, the effusion should be excluded as a stage descriptor. An effusion thought to be malignant is thus counted as M1a , whether it is microscopically proven or not.

[b]This includes involvement of a single non-regional node.

[c]For example, the skeleton is considered one organ system. Multiple metastases in several bones are classified as M1c1. Multiple metastases in the liver are classified as M1c1. Metastasis involving liver and bone would be considered M1c2.

Table 1. Adjusted hazard ratios for overall survival of patients between 9th version N components[17]

	cN (44,309 patients)		pN (34,342 patients)	
	HR (95% CI)	***P* value**	**HR (95% CI)**	***P* value**
N1 vs. N0	1.96 (1.84-2.08)	<0.0001	2.40 (2.26-2.55)	<0.0001
N2a vs. N1	1.42 (1.28-1.56)	<0.0001	1.45 (1.31-1.60)	<0.0001
N2b vs. N2a	1.27 (1.13-1.43)	<0.0001	1.46 (1.32-1.62)	<0.0001
N3 vs. N2b	1.51(1.35-1.70)	<0.0001	1.62 (1.29-2.03)	<0.0001

Overall survival was compared between 9th version N components based on a Cox proportional hazards model with covariates of 9th N component, sex, age, histologic type, history of prior malignancy, geographical region, and completeness of resection
HR, Hazard ratio; 95% CI, 95% confidence interval, *P* value from chi-square test score in Cox regression model

Table 2. Cox regression for overall survival by number of lesions and sites, stratified by datasource; analysis of M components[19]

Component	Variable	n/N (%)	HR (95% CI)	*P* value
M1 components: M1a, M1b, M1c1 (single organ system), and M1c2 (multiple organ systems)				
M1a	M1a	5406/14926 (36%)	Reference level	N/A
M1b	M1b; single organ system, single lesions (vs. M1a)	1927/14926 (13%)	1.18 (1.10-1.27)	<0.001
M1c single organ system	M1c1; single organ system, multiple lesions (vs. M1b)	2207/14926 (15%)	1.17 (1.08-1.27)	<0.001
M1c2 multiple organ systems	M1c2; multiple organ systems, multiple lesions (vs. M1c1 single organ systems)	5386/14926 (36%)	1.33 (1.25-1.41)	<0.001
Adjustment factors				
Age ≥65		8577/14926 (57%)	1.35 (1.30-1.41)	<0.001
Male		8838/14926 (59%)	1.32 (1.27-1.38)	<0.001
Squamous		2529/14926 (17%)	1.34 (1.27-1.41)	<0.001
Region: Asia (vs. other)		6872/14926 (46%)	0.93 (0.89-0.97)	<0.001

HR, Hazard ratio; 95% CI, 95% confidence interval, N/A: not applicable

Stage Groups

The changes in stage group classifications can be summarised as follows:

1. T1N2a has been included in stage 2B.

2. T2N2b has been included in stage 3B.

3. T3N2a has been included in stage 3A.

4. T1N1 has been downstaged from stage 2B to stage 2A.

The 9th TNM stage groups are summarised in the tables below (Tables 3, 4).[20]

DISCUSSION

The staging system in lung cancer plays a critical role in diagnosis, treatment, and prognosis. The 9th TNM staging system is a modernised framework designed to support clinical decision-making by more precisely classifying the anatomical extent of the disease. Implemented worldwide as of January 2025, the most significant changes in the 9th version pertain to the N and M components.

In comparison to the 8th edition:

• The **N2** component has been subdivided into **N2a** and **N2b**.

• The **M1c** component has been subdivided into **M1c1** and **M1c2**.

These detailed subdivisions are expected to contribute to the personalisation of treatment strategies. No changes have been made to the T component.[19]

In lung cancer, N2 disease exhibits a highly heterogeneous profile in terms of its clinical, anatomical, and biological characteristics. This heterogeneity significantly influences both patient prognosis and response to treatment. N2 disease can be subclassified based on several factors, including the number of metastatic lymph node stations involved (single-station vs. multi-station), the morphological features of nodal metastases (bulky vs. non-bulky), the presence or absence of concurrent hilar or intrapulmonary nodal involvement (skip vs. non-skip), and the timing of diagnosis (preoperative, intraoperative, or incidental). These subgroups differ substantially with respect to overall survival, treatment responsiveness, and suitability for surgical intervention.[21] Therefore, acknowledging this heterogeneity is critical in managing N2 disease and in guiding personalized treatment strategies. Given this pronounced heterogeneity exhibited by N2 disease, the division of the N2 component into two subgroups in the 9th version of the TNM staging system is an important step towards at least partially reducing this heterogeneity.

According to the National Comprehensive Cancer Network (NCCN) guidelines, the treatment of N2 disease is based on two primary approaches. The first approach involves concurrent chemoradiotherapy followed by maintenance therapy with immune checkpoint inhibitors or targeted agents. The second approach entails evaluating the feasibility of surgical resection after neoadjuvant systemic therapy using immune checkpoint inhibitors or targeted therapies. In this strategy, surgery should only be considered for patients who do not exhibit significant tumor progression following systemic therapy. However, for cN2 patients with single-station involvement, upfront surgery followed by adjuvant chemotherapy is also listed

Table 3. Stage groups of the 9th version of the tumor, node, metastasis (TNM) classification of lung cancer[20]

T/M	Components and descriptors	N0	N1	N2		N3
				N2a single station	N2b multiple station	
T1	**T1a:** ≤1 cm	IA1	IIA	IIB	IIIA	IIIB
	T1b: >1 to ≤2 cm	IA2	IIA	IIB	IIIA	IIIB
	T1c: >2 to ≤3 cm	IA3	IIA	IIB	IIIA	IIIB
T2	**T2a:** Visceral pleura/central invasion	IB	IIB	IIIA	IIIB	IIIB
	T2a: >3 to ≤4 cm	IB	IIB	IIIA	IIIB	IIIB
	T2b: >4 to ≤5 cm	IIA	IIB	IIIA	IIIB	IIIB
T3	**T3:** >5 to ≤7 cm	IIB	IIIA	IIIA	IIIB	IIIC
	T3: Invasion	IIB	IIIA	IIIA	IIIB	IIIC
	T3: Same lobe separate tumor nodules	IIB	IIIA	IIIA	IIIB	IIIC
T4	**T4:** >7 cm	IIIA	IIIA	IIIB	IIIB	IIIC
	T4: Invasion	IIIA	IIIA	IIIB	IIIB	IIIC
	T4: Ipsilateral separate tumor nodules	IIIA	IIIA	IIIB	IIIB	IIIC
M1	**M1a:** Contralateral tumor nodules	IVA	IVA	IVA	IVA	IVA
	M1a: Pleural/pericardial effusion, nodules	IVA	IVA	IVA	IVA	IVA
	M1b: Single extrathoracic metastasis	IVA	IVA	IVA	IVA	IVA
	M1c1: Multiple metastases in 1 organ system	IVB	IVB	IVB	IVB	IVB
	M1c2: Multiple metastases in >1 organ systems	IVB	IVB	IVB	IVB	IVB

as a conditional treatment option in the NCCN guidelines. Especially in patients with non-bulky, single-station N2 disease and adequate cardiopulmonary reserve, if a complete resection with lobectomy is anticipated, an upfront surgical approach can be a suitable option in selected cases.[22] In this context, the subclassification of N2 disease into N2a and N2b can be considered an important approach in treatment planning.

In lung cancer, histopathological confirmation plays a critical role in differentiating between benign and malignant mediastinal lymph nodes. For the histopathologic evaluation of mediastinal lymph nodes, both minimally invasive techniques such as endobronchial ultrasound-guided transbronchial needle aspiration (EBUS/TBNA) and endoscopic ultrasound-guided fine needle aspiration and invasive procedures such as standard cervical mediastinoscopy, video-assisted mediastinoscopy, extended mediastinoscopy, video-assisted mediastinoscopic lymphadenectomy, transcervical extended mediastinal lymphadenectomy, anterior mediastinotomy, and video thoracoscopy are utilized.[23,24] Current guidelines recommend that all patients diagnosed with NSCLC who are surgically resectable should undergo invasive mediastinal staging. The exception to this are tumors that are ≤3 cm in diameter, peripherally located, and have no clinical lymph node involvement.[22] However, in patients with negative mediastinal lymph nodes on positron emission tomography computed tomography (PET-CT), invasive staging is still recommended if additional risk factors are present such as non-squamous histology, centrally located tumors, or concomitant cN1 disease. Invasive staging is also recommended in cases where suspicion of cN2 persists based on thorax CT and PET-CT findings, despite a negative EBUS/TBNA result.[25] Restaging of mediastinal lymph nodes following neoadjuvant therapy holds substantial clinical significance in the surgical planning of lung cancer patients. Various studies have demonstrated that patients with residual metastatic lymph nodes after neoadjuvant therapy have significantly lower survival rates compared to those who achieve nodal downstaging to ypN0–1 status.[26,27] Therefore, selecting surgical candidates after neoadjuvant therapy can only be reliably performed using restaging techniques with high sensitivity and specificity.

Oligometastatic NSCLC, similar to N2 disease, represents a clinically and biologically heterogeneous group. This new subclassification facilitates a clearer definition of the concept of oligometastatic disease and the guidance of curative treatment options for this patient group. In the M1c1 subgroup, the combination of local therapies (such as surgery or stereotactic radiotherapy) with immune checkpoint inhibitors has shown promising clinical outcomes.[28]

External validation studies on the 9th version of the TNM staging system have demonstrated its superiority over previous editions in terms of prognostic accuracy, survival discrimination power, and clinical applicability. In a retrospective analysis by Son et al.[29],

Table 4. Comparison of stage groups in the 8th and 9th version staging systems for lung cancer[20]

8th TNM components						**9th TNM components**						
T/M	**Components and descriptors**	**N0**	**N1**	**N2**	**N3**	**T/M**	**Components and descriptors**	**N0**	**N1**	**N2**		**N3**
										N2a	**N2b**	
T1	T1a ≤1 cm	IA1	IIB	IIIA	IIIB	T1	T1a ≤1 cm	IA1	IIA	**IIB**	IIIA	IIIB
	T1b >1 to ≤2 cm	IA2	IIB	IIIA	IIIB		T1b >1 to ≤2 cm	IA2	IIA	**IIB**	IIIA	IIIB
	T1c >2 to ≤3 cm	IA3	IIB	IIIA	IIIB		T1c >2 to ≤3 cm	IA3	IIA	**IIB**	IIIA	IIIB
T2	T2a	IB	IIB	IIIA	IIIB	T2	T2a	IB	IIB	IIIA	**IIIB**	IIIB
	T2a >3 to ≤4 cm	IB	IIB	IIIA	IIIB		T2a >3 to ≤4 cm	IB	IIB	IIIA	**IIIB**	IIIB
	T2b >4 to ≤5 cm	IIA	IIB	IIIA	IIIB		T2b >4 to ≤5 cm	IIA	IIB	IIIA	**IIIB**	IIIB
T3	T3 >5 to ≤7 cm	IIB	IIIA	IIIB	IIIC	T3	T3 >5 to ≤7 cm	IIB	IIIA	**IIIA**	IIIB	IIIC
	T3 invasion	IIB	IIIA	IIIB	IIIC		T3 invasion	IIB	IIIA	**IIIA**	IIIB	IIIC
	T3 satellite nodules	IIB	IIIA	IIIB	IIIC		T3 satellite nodules	IIB	IIIA	**IIIA**	IIIB	IIIC
T4	T4 >7 cm	IIIA	IIIA	IIIB	IIIC	T4	T4 >7 cm	IIIA	IIIA	IIIB	IIIB	IIIC
	T4 Invasion	IIIA	IIIA	IIIB	IIIC		T4 invasion	IIIA	IIIA	IIIB	IIIB	IIIC
	T4 ipsilateral nodules	IIIA	IIIA	IIIB	IIIC		T4 ipsilateral nodules	IIIA	IIIA	IIIB	IIIB	IIIC
M1	M1a contralateral nodules	IVA	IVA	IVA	IVA	M1	M1a contralateral nodules	IVA	IVA	IVA	IVA	IVA
	M1a pleural, pericardial effusion	IVA	IVA	IVA	IVA		M1a pleural, pericardial effusion	IVA	IVA	IVA	IVA	IVA
	M1b single extrathoracic lesion	IVA	IVA	IVA	IVA		M1b single extrathoracic lesion	IVA	IVA	IVA	IVA	IVA
	M1c multiple lesions	IVB	IVB	IVB	IVB		M1c1 multiple lesions, single organ system	IVB	IVB	IVB	IVB	IVB

TNM: tumor, node, metastasis

which included 4,029 patients diagnosed with stage I–III NSCLC, the 9th version exhibited greater discriminatory power for overall survival and freedom from recurrence compared to the 8th edition. Furthermore, this study also showed that subclassifying N2 disease into N2a and N2b significantly improved survival prediction and enabled a more refined prognostic stratification.[29]

Similarly, in a study where Kim et al.[30] analyzed the data of 7,429 patients, they emphasized that the 9th version of the TNM staging system exhibited higher prognostic accuracy and discriminatory ability for both clinical and pathological stages compared to the 8th edition. In another study conducted by Wang et al.[31] using the Surveillance, Epidemiology, and End Results database, a total of 19,193 patients with stage IA–IIIA NSCLC who underwent lobectomy were retrospectively evaluated. In this analysis, the 9th version of the TNM staging system was able to distinguish the survival difference between stage IB and IIA more distinctly than the previous edition, and this difference was found to be statistically significant. The findings reveal that the TNM 9th version is valid and applicable for NSCLC patients; furthermore, it provided a prognostic accuracy in survival prediction that was almost equivalent to the 8th version.[31]

The 9th TNM staging system focuses solely on anatomical spread and does not evaluate tumor biological features, molecular profiles, or immunological status. Non-anatomical variables were excluded, and the static structure of staging has been maintained. Numerous prognostic factors have been identified in lung cancer, and their number continues to grow in parallel with scientific and clinical advancements. However, integrating these factors into the staging system can transform it into a complex and difficult-to-apply structure. Therefore, while the TNM stage should remain the cornerstone for prognostication, other prognostic variables not included in the staging system should also be considered in clinical decision-making processes.

This limitation may lead to significant variability in clinical outcomes among patients with the same TNM stage. In an era where personalised medicine is becoming increasingly important, this restricts the system's applicability. Integrating molecular biology, dynamic staging approaches, and artificial intelligence-driven analyses could further enhance the system's clinical utility. In the development of the 9th version of the TNM staging system, recursive partitioning modeling—a machine learning method based on decision trees—was utilized. This method is an artificial intelligence approach with a limited scope, used to evaluate the effects of the T, N, and M components on survival. Today, the use of artificial intelligence methods, such as deep learning, which can automatically learn meaningful features in data without human intervention, is increasing. Through deep learning, it is possible to analyze complex, high-dimensional data such as staging parameters and to uncover subtle, non-obvious relationships within them. Consequently, the gradual integration of artificial intelligence into the TNM staging system holds the potential to render cancer staging more dynamic and personalized. Future research should focus on advancing the TNM staging system in these directions. The fourth phase of the staging project has already begun under the leadership of U.S. thoracic surgeon Valerie Rusch, with the 10th version of the lung cancer TNM staging system expected to be implemented by 2031.[32]

CONCLUSION

In conclusion, the 9th TNM staging system has provided a significant advancement in staging and treatment planning by improving survival predictions. This review prepares the ground for a more precise and personalized staging system in the next TNM revision.

Ethics

Footnotes

Authorship Contributions

Surgical and Medical Practices: Y.K., F.M., G.Ö.Ş., H.S.K., S.B.Ö., M.M.T., T.G., Concept: Y.K., F.M., G.Ö.Ş., H.S.K., S.B.Ö., M.M.T., T.G., Design: Y.K., F.M., G.Ö.Ş., H.S.K., S.B.Ö., M.M.T., T.G., Data Collection or Processing: Y.K., G.Ö.Ş., T.G., Analysis or Interpretation: Y.K., G.Ö.Ş., Literature Search: Y.K., Writing: Y.K., G.Ö.Ş.

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 cancer observatory (2024). Last Accessed Date: 17.10.2024. **[Crossref]**
2. Garinet S, Wang P, Mansuet-Lupo A, Fournel L, Wislez M, Blons H. Updated prognostic factors in localized NSCLC. *Cancers (Basel).* 2022;14(6):1400. **[Crossref]**
3. Kahya Y, Kayı Cangır A. The philosophy of staging in lung cancer: prognosis or treatment planning? *Eur J Ther.* 2018;24(Suppl 1);S1-S3. **[Crossref]**
4. Rami-Porta R. The TNM classification of lung cancer-a historic perspective. *J Thorac Dis.* 2024;16(11):8053-8067. **[Crossref]**
5. Amin MB, Edge SB, Greene FL, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin.* 2017;67(2):93-99. **[Crossref]**
6. Asamura H, Nishimura KK, Giroux DJ, et al. IASLC lung cancer staging project: the new database to inform revisions in the ninth edition of the TNM classification of lung cancer. *J Thorac Oncol.* 2023;18(5):564-575. **[Crossref]**
7. Asamura H, Farjah F, Gill R, et al. *AJCC Cancer Staging System*: lung (version 9). Kindle. Chicago: Springer; 2024. **[Crossref]**
8. Goldstraw P, Crowley J, Chansky K, et al. International Association for the Study of Lung Cancer International Staging Committee; participating institutions. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. *J Thorac Oncol.* 2007;2(8):706-714. **[Crossref]**
9. Goldstraw P, Chansky K, Crowley J, et al. The IASLC Lung Cancer Staging Project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM Classification for lung cancer. *J Thorac Oncol.* 2016;11(1):39-51. **[Crossref]**

10. Sobin L, Wittekind C. UICC TNM Classification of Malignant Tumours. 6th ed. New York: Wiley-Liss; 2002. **[Crossref]**
11. Detterbeck FC, Nishimura KK, Cilento VJ, et al. The International Association for the Study of Lung Cancer staging project: methods and guiding principles for the development of the ninth edition TNM classification. *J Thorac Oncol.* 2022;17(6):806-815. **[Crossref]**
12. Detterbeck FC, Ostrowski M, Hoffmann H, et al. The International Association for the Study of Lung Cancer lung cancer staging project: proposals for revision of the classification of residual tumor after resection for the forthcoming ninth edition of the TNM classification of lung cancer. *J Thorac Oncol.* 2024;19(7):1052-1072. **[Crossref]**
13. Osarogiagbon RU, Rami-Porta R, Tsao MS, et al. International Association for the Study of Lung Cancer molecular database project: objectives, challenges and opportunities. *J Thorac Oncol.* 2021;16(6):897-901. **[Crossref]**
14. Van Schil PE, Asamura H, Nishimura KK, et al. The IASLC lung cancer staging project: proposals for the revisions of the T descriptors in the forthcoming ninth edition of the TNM classification for lung cancer. *J Thorac Oncol.* 2024;19(5):749-765. **[Crossref]**
15. Ugalde Figueroa PA, Marques E, Cilento VJ, et al. Completeness of resection and long-term survival of patients undergoing resection for T3 non-small-cell lung cancer: an International Association for the Study of Lung Cancer analysis. *J Thorac Oncol.* 2024;19(1):141-152. **[Crossref]**
16. Rusch VW, Asamura H, Watanabe H, et al. The IASLC lung cancer staging project: a proposal for a new international lymph node map in the forthcoming seventh edition of the TNM classification for lung cancer. *J Thorac Oncol.* 2009;4:568-577. **[Crossref]**
17. Huang J, Osarogiagbon RU, Giroux DJ, et al. The International Association for the Study of Lung Cancer staging project for lung cancer: proposals for the revision of the N descriptors in the forthcoming ninth edition of the TNM classification for lung cancer. *J Thorac Oncol.* 2024;19(5):766-785. **[Crossref]**
18. Osarogiagbon RU, Van Schil P, Giroux DJ, et al. The International Association for the Study of Lung Cancer lung cancer staging project: overview of challenges and opportunities in revising the nodal classification of lung cancer. *J Thorac Oncol.* 2023;18:410-418. **[Crossref]**
19. Fong KM, Rosenthal A, Giroux DJ, et al. The International Association for the Study of Lung Cancer staging project for lung cancer: proposals for the revision of the M descriptors in the forthcoming ninth edition of the TNM classification of lung cancer. *J Thorac Oncol.* 2024;19(5):786-802. **[Crossref]**
20. Rami-Porta R, Nishimura KK, Giroux DJ, et al. The International Association for the Study of Lung Cancer lung cancer staging project: proposals for revision of the TNM stage groups in the forthcoming ninth edition of the TNM classification for lung cancer. *J Thorac Oncol.* 2024;19(7):1007-1027. **[Crossref]**
21. Crinò L, Weder W, van Meerbeeck J, Felip E; ESMO Guidelines Working Group. Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2010;21:103-115. **[Crossref]**
22. National Comprehensive Cancer Network. (2025). *NCCN clinical practice guidelines in oncology: non-small cell lung cancer (version 4.2025)*. **[Crossref]**
23. Gürün Kaya A, Çiledağ A, Erol S, et al. Evaluation of lung cancer biomarkers profile for the decision of targeted therapy in EBUS-TBNA cytological samples. *Scott Med J.* 2022;67(1):18-27. **[Crossref]**
24. Leiro-Fernández V, Fernández-Villar A. Mediastinal staging for non-small cell lung cancer. *Transl Lung Cancer Res.* 2021;10(1):496-505. **[Crossref]**
25. Boada M, Sánchez-Lorente D, Libreros A, et al. Is invasive mediastinal staging necessary in intermediate risk patients with negative PET/CT? *J Thorac Dis.* 2020;12(8):3976-3986. **[Crossref]**
26. Steger V, Walker T, Mustafi M, et al. Surgery on unfavourable persistent N2/N3 non-small-cell lung cancer after trimodal therapy: do the results justify the risk? *Interact Cardiovasc Thorac Surg.* 2012;15(6):948-953. **[Crossref]**
27. Stamatis G, Müller S, Weinreich G, et al. Significantly favourable outcome for patients with non-small-cell lung cancer stage IIIA/IIIB and single-station persistent N2 (skip or additionally N1) disease after multimodality treatment. *Eur J Cardiothorac Surg.* 2022;61(2):269-276. **[Crossref]**
28. Antonoff MB, Deboever N, Werner R, Altan M, Gomez D, Opitz I. Surgery for oligometastatic non-small cell lung cancer. *J Thorac Cardiovasc Surg.* 2024;167(2):508-516. **[Crossref]**
29. Son JW, Lee J, Jeon JH, et al. Validation of IASLC 9th edition TNM classification for lung cancer: focus on N descriptor. *BMC Cancer.* 2024;24(1):1460. **[Crossref]**
30. Kim S, Ahn Y, Lee GD, et al. Validation of the 9th tumor, node, and metastasis staging system for patients with surgically resected non-small cell lung cancer. *Eur J Cancer.* 2025;222:115436. **[Crossref]**
31. Wang RR, Li MJ, Peng Q, Huang ZY, Wu LL, Xie D. Validation of the 9th edition of the TNM staging system for non-small cell lung cancer with lobectomy in stage IA-IIIA. *Eur J Cardiothorac Surg.* 2024;65(3):071. **[Crossref]**
32. International Association for the Study of Lung Cancer (IASLC). IASLC staging project: lung cancer, thymic tumors, and mesothelioma [Internet]. Aurora (CO): IASLC; Cited: 25.08.2025. **[Crossref]**

Thorac Res Pract. 2026;27(1):57-67 **DOI:** 10.4274/ThoracResPract.2025.2025-6-2

Review

Artificial Intelligence in Pleural Diseases: Current Applications and Next Steps

Ferhan Karataş[1], Öner Dikensoy[1,2]

[1]Department of Pulmonary Diseases, Koç University Hospital, İstanbul, Türkiye
[2]Department of Pulmonary Diseases, Koç University Faculty of Medicine, İstanbul, Türkiye

Cite this article as: Karataş F, Dikensoy Ö. Artificial intelligence in pleural diseases: current applications and next steps. *Thorac Res Pract.* 2026;27(1):57-67

ABSTRACT

Pleural diseases pose a significant burden on healthcare systems due to diagnostic challenges and high costs. Artificial intelligence (AI) has the potential to provide faster, more accurate, and more reliable results in the diagnosis of these diseases. This review evaluates the current status of AI technologies in the diagnosis of pleural effusion (PE), malignant PE, tuberculosis pleurisy (TP), pneumothorax, and malignant pleural mesothelioma (MPM). Deep learning algorithms developed for radiological diagnosis provide high sensitivity and specificity in determining the presence and severity of PE. AI models that integrate clinical parameters such as chest computed tomography (CT), positron emission tomography (PET)-CT, and tumour markers in distinguishing between benign and malignant effusions have significantly improved diagnostic accuracy (area under the curve: >0.90). In cytological diagnosis, computer-assisted systems such as Aitrox have demonstrated performance comparable to that of expert cytopathologists in diagnosing malignant effusions. In the diagnosis of TP, AI models outperform conventional diagnostic methods, particularly when combined with laboratory parameters such as adenosine deaminase. Food and Drug Administration-approved AI models are effectively used for the rapid diagnosis of pneumothorax and for emergency interventions. In MPM diagnosis, AI models using PET-CT images and three-dimensional segmentation offer significant advantages in prognostic evaluation and treatment response monitoring. However, large-scale, multi-centre studies are needed to standardise and generalise AI models. In light of these developments, AI may fundamentally change the diagnostic management of pleural diseases.

KEYWORDS: Artificial intelligence, diagnostic methods, mesothelioma, pleural diseases, pneumothorax, tuberculosis

Received: 18.06.2025 ***Revision Requested:*** 03.09.2025 ***Last Revision Received:*** 09.09.2025

Accepted: 24.09.2025 ***Epub:*** 15.01.2026 ***Publication Date:*** 30.01.2026

INTRODUCTION

Pleural diseases are heterogeneous, with diverse etiologies posing complex diagnostic and therapeutic challenges.[1,2] Pleural diseases constitute a serious burden on healthcare systems and patients. They represent a significant burden, with an estimated annual incidence of 1.5 million cases in the USA and hospitalization costs exceeding 10 billion dollars.[1-4] Despite established diagnostic principles, no single method adequately meets clinical needs, underscoring the demand for faster and more accurate tools.[2,3]

Artificial intelligence (AI), defined as the simulation of human cognition through machine learning (ML), deep learning (DL), and data analysis, has recently become an essential component of healthcare.[5-7] Advances in big data and computational power have enabled AI to provide rapid, precise, and reliable analysis of radiological and clinical data, improving diagnostic accuracy and reducing clinician workload.[3,5-7] While initially developed for imaging, AI in respiratory medicine is now applied not only to lung cancer screening but also to pleural effusion (PE) and pneumothorax, and has demonstrated high diagnostic performance in these applications. Well-trained AI systems integrating clinical, radiological, and laboratory data support physicians with greater accuracy and reduced need for immediate specialist access. Consequently, AI is expected to play an increasingly important role in chest medicine.[3,5,6] A systematic review confirmed that AI-assisted observers outperformed human readers in detecting thoracic pathologies

Corresponding author: Ferhan Karataş, MD, e-mail: fkaratas19@hotmail.com, fkaratas@kuh.ku.edu.tr

on chest X-rays (CXR) and computed tomography (CT) scans, achieving higher sensitivity, specificity, accuracy, and area under the curve (AUC).[8]

This review discusses the current status, potential advantages, and limitations of AI applications in the diagnosis of pleural diseases, and comprehensively evaluates the potential contributions of these technologies to future clinical applications inlight of the existing literature.

METHODS

A comprehensive search of the PubMed database was performed, without restriction on publication dates, using combinations of the following search terms: "pleural disease", "pleural effusion", "pneumothorax", "pleural mesothelioma", "tuberculous pleurisy", "artificial intelligence", "machine learning", "deep learning", "ultrasound", "cytology", and "PET-CT". We included English-language human studies that reported the development, validation, or clinical impact of AI systems in pleural diseases across imaging [CXR, chest CT, positron emission tomography (PET)-CT, lung ultrasound], cytology/whole-slide imaging (WSI), and laboratory/biomarker data. We excluded non-pleural AI studies, editorials, letters without primary data, and purely technical papers lacking clinical outcomes. Reference lists and recent reviews were screened to identify additional relevant publications.

Technique and Modality Overview

Various imaging and laboratory modalities have been used in AI applications for pleural diseases. CXR is mainly used for triage of conditions such as pneumothorax and gross effusion. Chest CT allows detailed segmentation and complexity scoring of effusions. PET-CT provides additional differentiation between malignant and benign processes, as well as prognostic information in malignant pleural mesothelioma (MPM). Cytology with WSI supports benign-malignant triage in effusion samples. Lung ultrasound is increasingly being studied for real-time guidance and monitoring, though data remain preliminary. Figure 1 illustrates the overall workflow of AI applications in pleural diseases across different imaging and laboratory modalities, while Table 1 summarizes their main tasks, reported performance metrics, and current readiness levels.

Main Points

- Artificial intelligence (AI)-based imaging algorithms achieve high diagnostic performance in pleural diseases, with deep learning models on chest computed tomography (CT) and positron emission tomography-CT consistently reporting area under the curve values >0.90 for detecting pleural effusion (PE) and differentiating malignant from benign effusions.
- Automated segmentation and quantification tools improve reproducibility and workflow efficiency, particularly for PE volume assessment and malignant pleural mesothelioma follow-up, reaching near–expert-level agreement while reducing clinician workload.
- Multimodal AI models that integrate imaging with clinical and laboratory data outperform single-modality approaches, enhancing diagnostic accuracy in malignant PE and tuberculous pleurisy, often surpassing traditional biomarkers alone.
- AI-assisted cytology systems demonstrate diagnostic accuracy comparable to experienced cytopathologists, offering objective and rapid triage of malignant PEs and helping address inter-observer variability and workforce limitations.
- Despite promising results, most AI applications lack large-scale prospective validation, highlighting the need for multicenter studies, standardized protocols, and careful clinical integration before routine adoption in pleural disease management.

Artificial Intelligence in Radiological Diagnosis of Pleural Effusion

In a study, the AI model developed using 600 posterior-anterior chest radiographs had a sensitivity of 95%, a specificity of 97%, and an AUC of 97% for the detection of PE.[9] The same study also showed that abnormal radiological findings, such as air-fluid level, atelectasis, bleb, cardiomegaly, bone fracture, infiltration, mass, nodule, obstructive airway disease, pneumonia, and scoliosis, which may introduce bias, did not affect the algorithm's performance.

The AI model developed to assess PE severity may aid in treatment effectiveness and management. In a study evaluating the presence and severity of PE by AI in chest radiographs of chronic obstructive pulmonary disease patients, the model showed 85.4% accuracy (AUC: 0.95) for non-PE images; 12.5% of the 14.5% errors were mild PE.[10] Prediction accuracy rates for non-PE, small, moderate, and large effusions were 83.95%, 74.19%, 62.16%, and 50%, respectively. Zhou et al.[11] developed DL models to detect and segment cardiomegaly, pneumothorax, and PE on chest radiographs. A high-quality, labeled radiograph dataset was created, with lesion regions annotated by radiologists. The model used AP75 (overlap ≥75% between predicted and actual lesion regions) as the performance metric. High accuracy was achieved in detecting cardiomegaly (AP75: 98.0%), pneumothorax (AP75: 71.2%), and PE (AP75: 78.2%). Segmentation performance, evaluated using the dice similarity coefficient (DSC), was also strong (e.g., lung-field dice: 0.960). Detection and semi-quantitative analysis times with DL were significantly shorter than those of radiologists ($P < 0.001$), potentially expediting clinical workflows. These models show promise in automating lesion detection and quantitative analysis, supporting radiologists' diagnostic decision-making.

In a retrospective study, a series of DL-based sequential models was developed to automatically detect, segment, and classify PE as simple or complex using CT images from 2,659 patients.[12] A detection-segmentation network based on the self-configuring nnU-Net architecture achieved 99% sensitivity, 98% specificity, and a DSC of 0.89, matching human-level volumetric consistency.

Five classification models using the random forest (RF) algorithm, based on radiological features (hyperdense fluid, pleural thickening, gas, loculation) and radiomic features, reached an AUC of 0.77 for classifying simple vs. complex PE. Notably, pleural thickening yielded an AUC of 0.91, and hyperdense fluid achieved a negative predictive value (NPV) of

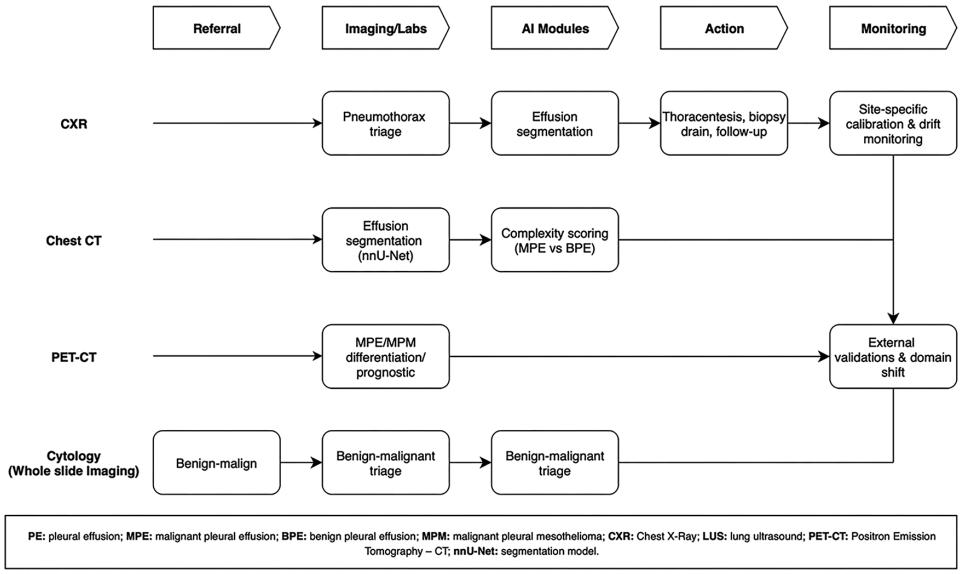

Figure 1. Workflow of artificial intelligence in pleural diseases by modality

Referral → imaging/labs (CXR, chest CT, PET-CT, LUS; cytology if present) → AI modules (triage, detection/segmentation, differential diagnosis, risk flags) → clinician review with thresholds → action (thoracentesis, biopsy, drain, follow-up) → monitoring (volume change, cytology triage)

Note: Performance pitfalls (e.g., small pneumothoraces on AP CXR and skin folds mimicking pneumothorax) highlight the need for site-specific calibration and continuous monitoring. In addition, all AI outputs are intended to function as decision-support tools and must be reviewed and validated by a physician before being used to guide clinical action

CXR: chest X-ray, CT: computed tomography, PET: positron emission tomography, LUS: lung ultrasound, AI: artificial intelligence, MPE: malignant pleural effusion, BPE: benign pleural effusion, MPM: malignant pleural mesothelioma, LUS: lung ultrasound, nnU-Net: segmentation model

Table 1. Overview of imaging and laboratory modalities used in artificial intelligence applications for pleural diseases

Modality	Main AI task(s)	Reported performance	Notes
CXR	Pneumothorax triage, effusion detection	Sensitivity 84-95%, specificity 88-97%, AUC ≈ 0.95	FDA-cleared tools available
Chest CT	Effusion segmentation, simple vs. complex PE classification	DSC 0.89; ICC ≥0.995	Highly reproducible, underestimates absolute volume
PET-CT	MPE vs. BPE differentiation, MPM prognosis	AUC up to 0.97	Small cohorts; potential in multimodal fusion
Cytology (WSI)	Malign vs. benign triage	AUC ≈ 0.95; accuracy >90%	Comparable to expert cytopathologists
LUS	Effusion detection/segmentation	Accuracy >90%, DSC 0.70	Operator dependence reduced with AI

Note: Reported performance values vary depending on dataset size, study design, and validation method. "Readiness" refers to current status: Research (tested in retrospective or single-center studies) vs. FDA-cleared/deployed (approved tools already used in clinical workflows)

AI: artificial intelligence, CXR: chest X-ray, CT: computed tomography, PET: positron emission tomography, WSI: whole-slide imaging, LUS: lung ultrasound, PE: pleural effusion, MPE: malignant pleural effusion, BPE: benign pleural effusion, MPM: malignant pleural mesothelioma, DSC: Dice similarity coefficient, ICC: intraclass correlation coefficient, AUC: area under the curve, FDA: Food and Drug Administration

0.94. Segmentation accuracy remained unaffected by contrast use or effusion complexity, confirming the model's robustness across diverse clinical scenarios.

In a prospective study of 79 patients, the accuracy of an AI algorithm that quantifies PE volume change via automatic segmentation of pre- and post-thoracentesis CT images was compared with actual drainage volumes.[13] The fully automated method underestimated drainage by 13.1%, while the semi-automated method, corrected by a thoracic radiologist, underestimated by 10.9% drainage, and the error increased linearly with fluid volume. Despite this, both agreement between methods [fully automated vs. semi-automated; intraclass correlation coefficient (ICC): 0.99] and test-retest reliability ICC: ≥0.995 were excellent. These results indicate that, although AI-based CT measurements are highly reproducible for clinical use, they consistently underestimate actual volumes, underscoring the need for calibration or correction, and highlighting the importance of accounting for algorithmic bias in quantitative monitoring of PE treatment response.

Artificial Intelligence in the Diagnosis of Malignant Pleural Effusion

The gold standard for differentiating malignant from benign PE is cytopathological examination of samples obtained by thoracentesis or pleural biopsy. Although they have high

specificity, their disadvantages include low positivity rates in pathological diagnosis, invasiveness, and high cost.[14,15] Therefore, non-invasive diagnostic methods with high sensitivity and improved diagnostic performance are needed.

Chest CT is one of the initial tools for the differential diagnosis of PE, and AI models may help reduce radiologists' workload. Wang et al.[16] developed an AI model that used chest CT features. The study analysed 918 PE cases—607 internal and 311 external—and created a training cohort of 362 cases from another centre. The model follows a two-stage structure: first, PE areas were segmented from CT images, producing masks with high overlap with expert-defined boundaries (mean DSC: 87.6%). Second, three dimensional (3D) PE masks and full CT volumes were used to classify the effusion as malignant or benign. The internal test cohort showed an AUC of 0.883, a sensitivity of 78.4%, and a specificity of 86.2%. In the external cohort, the AUC was 0.842, with 89.4% sensitivity and 65.1% specificity. Incorporating clinical data—such as sex, age, laterality, and PE volume—further improved AUCs in all three cohorts. The model detects suspicious areas, performs fine segmentation, and holistically analyzes chest CT features to differentiate benign pleural effusion (BPE) from MPE. The authors concluded that this AI model may assist radiologists and clinicians in PE case management.

Ozcelik et al.[17] compared the quantitative features of PE on CT scans with cytological results. The positive predictive value (PPV), NPV, sensitivity, specificity, and accuracy of the DL model for the diagnosis of MPE were reported as 93.3%, 86.67%, 87.5%, 92.86%, and 90%, respectively, for differentiating benign from malignant PE. In another study, the applicability of one DL model and five ML models in differentiating MPE from BPE was investigated.[18] In this study, a total of 898 patients were included; data from 726 patients were used for training and testing the models, and data from 172 patients were used for the prospective validation. The diagnostic performance, as measured by the AUC, was 90.9% in the training set, 88.3% in the test set, and 86.6% in the validation set. When laboratory findings were included, the addition of the carcinoembryonic antigen (CEA) level in PE showed the best diagnostic performance, with an AUC of 90.9%, a sensitivity of 82.09%, and a specificity of 91.37% at a cut-off value of 3.6 ng/mL. Clinically, this threshold can help non-invasively distinguish malignant from benign effusions and guide decisions about whether to proceed with or defer invasive procedures. The results of this study show that using AI to guide physicians in PE management is a highly effective, non-invasive diagnostic aid.

In another study, an ML model based on clinical, blood, and pleural-fluid examination features was developed to classify the etiology of PE into five groups, including transudate, malignant, parapneumonic, tuberculous pleurisy (TP), and others.[19] In this retrospective study, 18 basic features were identified by feature selection (FS), including history of malignancy; blood test results such as C-reactive protein and albumin; and pleural fluid characteristics such as lactate dehydrogenase (LDH), protein, adenosine deaminase (ADA), and CEA. The model's AUC for detecting MPE was 0.930 in the validation set and 0.916 in the extra validation set. The light gradient boosting machine (GBM) model showed the best performance, with accuracies of 81.8% and 78.7% in the validation and extra validation datasets, respectively. This ML algorithm provided high accuracy in the differential diagnosis of PE and may be useful as a clinical decision-support system by guiding clinicians regarding the necessity of invasive procedures.

Zhang et al.[20] developed five ML models—XGBoost, logistic regression (LR), Bayesian additive regression trees, RF, and support vector machine (SVM)—to assess the diagnostic performance of CEA, CA19-9, CA125, and CA15-3 in PE. Using these models, 319 PE cases were analyzed using both individual and combined tumor markers. Among single-marker models, the XGBoost model with CEA showed the highest AUC (0.895) and sensitivity (80%), while that with CA153 had the highest specificity (98%). Among marker combinations, the XGBoost model with CEA + CA153 achieved the highest AUC (0.921) and sensitivity (85%), whereas both the XGBoost model with CA125 + CA153 and the LR model with CEA + CA153 + CA19-9 achieved the highest specificity (97%). In another study, five AI models were developed using laboratory data from 2,352 patients.[21] The XGBoost model outperformed other models, with AUCs of 0.903, 0.918, and 0.886 in the training, validation, and test cohorts, respectively; it also achieved specificities of 89.2%, 93.4%, and 91.8%, and sensitivities of 86.1%, 84.4%, and 80.4%. PE CEA was the most important predictor, followed by serum CYFRA21-1, PE CA125, haematocrit, creatinine, calcium, and neutrophil percentage. The model also outperformed the standalone PE CEA in differentiating MPE from BPE.

The reported sensitivity of pleural fluid cytology is between 40% and 90%, depending on the tumor cell type.[22] However, inconsistencies are observed among pathologists in manual screening results, which suggest that these examinations are subjective. Although pleural metastasis indicates advanced-stage disease with a survival of 3-12 months, patients' life expectancy can be prolonged by prompt diagnosis. Early, accurate, rapid, and objective diagnoses can be achieved using automatic image analysis. Therefore, the development of computer-aided diagnosis (CAD) systems is essential. Win et al.[23] developed a CAD system for the diagnosis of MPE. In this system, 201 cellular features were analysed, and high diagnostic performance was achieved, with 87.9% sensitivity, 99.4% specificity, and 98.7% accuracy for the diagnosis of MPE. In another study, an AI model called "Aitrox" was applied to classify benign and malignant lung cancer cells in PE cytology.[24] The diagnostic performance of the model was compared between junior and senior cytopathologists. The "Aitrox" AI model featured in this study is a weakly supervised DL method based on a deep convolutional neural network (DCNN) designed to classify benign and malignant cases in lung cytological images at the WSI level. Aitrox AI achieved 91.67% accuracy, 87.5% sensitivity, and 94.4% specificity, with an AUC of 0.9526. These rates were higher than those of young cytopathologists and similar to those of senior cytopathologists. It has been demonstrated that this AI model provides more objective and consistent results than those of cytopathologists and may help alleviate the shortage of cytopathologists. Current research shows that AI can be used as an effective tool in the rapid cytological diagnosis of MPE. In addition to classifying PE as malignant or benign, AI

models could be developed to detect atypical cases or cases suspected of being malignant.

In a study using an AI algorithm to screen PET-CT scans for the diagnosis of PE in patients with malignancy, a sensitivity of 95.5%, a specificity of 92.6%, and an AUC of 97.7% were obtained for the diagnosis of MPE.[25] In the study, various clinical characteristics of the patients were not addressed, and the sample size was small; previous reports have also indicated that the results may differ if these limitations are mitigated.

Ultrasonography (USG) is widely used for diagnosing and monitoring lung pathologies due to its safety, bedside applicability, and low cost, though its main limitation is operator dependence. In a study aiming to develop an automated PE diagnosis system, a dataset of 623 videos comprising 99,209 two dimensional (2D) USG images from 70 patients was created.[26] Two models were trained using frame- and video-based labels. Based on expert interpretations, the models achieved accuracies of 92.4% and 91.1%, respectively, with no significant difference between them. The study demonstrated that PE can be reliably and efficiently diagnosed from USG images using AI. In another study, a DL model was developed for accurate segmentation of lung USG images of PE, and its performance was compared with that of experts.[27] A total of 3,041 USG images from 24 PE patients were segmented by two USG experts and used as the ground truth. The model's performance, assessed by DSC, was 0.70 for comparisons between the AI and experts, while the expert-to-expert DSC was 0.61. These results indicate that the algorithm's accuracy is comparable to that of human experts, supporting its potential as a reliable tool in PE diagnosis and management.

Artificial Intelligence in the Diagnosis of Tuberculous Pleurisy

TP is one of the most common forms of extrapulmonary tuberculosis (TB).[28,29] The incidence of TP varies according to the prevalence of TB. While it constitutes 4% of all TB cases in the USA, TP constitutes 20% of cases in South Africa.[29] The gold standard for the diagnosis of TP is the demonstration of TB bacilli in PE or biopsy samples. Elevated ADA and interferon-gamma (IFN-γ) levels in PE are important supportive diagnostic markers.[30,31] Comprehensive PE analysis, including ADA and IFN-γ levels, yields higher diagnostic accuracy; however, different cut-off values are reported.[30-32] On the other hand, the efficiency of these markers varies with disease prevalence and age group.[30,31] The microbiological culture positivity of PE for the diagnosis of TP ranges from 12% to 70%, and the highest reported diagnostic yield is 30%.[30] When solid media are used, it can take up to 8 weeks to obtain the results.[30] Pleural biopsy is another invasive, costly, and complication-prone diagnostic procedure. TP usually heals spontaneously, but 43-65% of these patients develop TB within a few years.[30,31] Therefore, early diagnosis and treatment of TP are important to prevent both fibrothorax and recurrent TB disease. It is necessary to develop a less-invasive, low-cost method with a high diagnostic yield for the diagnosis of TP.[28,30-32]

Several studies have explored the potential of AI models to enhance clinicians' diagnostic performance in TP. Ren et al.[32] compared the performance of four AI models—LR, k-nearest neighbour (KNN), SVM, and RF—against pleural fluid adenosine deaminase (pfADA) levels using a dataset of 443 PE cases that included demographic, clinical, and fluid analysis data. Combining AI models with pfADA yielded the highest diagnostic performance: 85.4% sensitivity, 84.1% specificity, and 84.7% accuracy at a pfADA cut-off of 17.5 U/L. When the cut-off is set at 17.5 U/L, the threshold can support clinical decision-making by favoring early initiation of anti-TB treatment in patients with a high pretest probability and concordant imaging or clinical findings, while deferring invasive procedures such as pleural biopsy. LR estimates outcome probabilities; KNN classifies based on the majority class of nearby samples; SVM identifies a hyperplane maximizing class separation; and RF uses multiple decision trees, combining their outputs to reduce overfitting and improve stability. Among all models, RF achieved the best performance (AUC: 0.971; sensitivity: 89.1%; specificity: 93.6%; PPV: 91.3%; NPV: 91.5%), outperforming LR (AUC: 0.876), KNN (AUC: 0.895), and SVM (AUC: 0.918), and pfADA alone.

A prospective multicentre study assessed, using ML, the diagnostic value of ADA levels, routine pleural fluid parameters (pH, glucose, protein, LDH, cell counts), and age for TP diagnosis in low-prevalence areas.[33] The SVM model showed the best performance with 97% accuracy, AUC: 0.98, 91% sensitivity, and 98% specificity. When ADA >40 U/L and lymphocyte ratio >50% were used, sensitivity and specificity were 97% and 93%, respectively, with a 100% NPV. In low-prevalence settings, this high NPV can help avoid unnecessary invasive diagnostic procedures by reliably ruling out TP. Including age and routine clinical parameters increased ADA specificity to 98% and PPD positivity rate to 64%, supporting its non-invasive diagnostic value. Another study used a decision tree and a weighted sparse representation-based classification (WSRC) model to differentiate TP from MPE based on pleural biomarkers.[34] Among 236 patients, ADA had the best individual performance (sensitivity: 91.9; specificity: 74%). Accuracy improved when age, polynuclear leukocytes, and lymphocytes were added. WSRC achieved an AUC of 0.877, sensitivity 93.3%, specificity 82.0%, PPV 87.5%, and NPV 90.1%. A decision flowchart based on these markers achieved an accuracy of 88.8% and provided a cost-effective and reliable tool. Liu et al.[35] developed several ML models for early TP diagnosis using data from 1,435 patients (plus 153 for external validation). The SVM model performed best, with 87.7% accuracy, 85.3% precision, an AUC of 0.914, 94.7% sensitivity, and 80.7% specificity. PE-ADA, PE-CEA, and serum CYFRA21-1 were the top predictors. In another study, five AI models—including LR, RF, gradient boosting, deep neural networks, and a contrastive-loss model—were compared to assess the etiology of PE, with ADA levels incorporated.[36] The contrastive-loss model had the best performance (sensitivity: 84.1%; specificity: 94.1%), surpassing those of traditional ADA-based criteria (sensitivity: 80.2%; specificity: 90.3%).

Li et al.[37] aimed to develop a novel AI model to diagnose TP. The study used 77 features, including clinical symptoms, routine blood tests, biochemical markers, pleural fluid cell counts, and fluid biochemistry. To optimize performance and identify key diagnostic factors, a FS model was used. The top five features were PE-ADA level, percentage of lymphocytes in pleural effusion (PELP), age, body temperature, and pleural fluid

color. An SVM model classified patients as TP or non-TP. The moth flame optimization (MFO) algorithm optimized the SVM parameters to achieve optimal performance. The resulting FS-MFO-SVM model achieved 95% accuracy, 93.35% sensitivity, 97.57% specificity, and an AUC of 95.6. This model predicts TP using patient data, such as clinical signs, blood tests, and pleural fluid analysis, with decisions primarily based on PE-ADA, PELP, age, temperature, and fluid color. The authors suggested that this non-invasive method could be an alternative to pleural biopsy and could hold promise for use in resource-limited settings due to its low cost and portability.

Artificial Intelligence in the Diagnosis of Pneumothorax

Pneumothorax can cause significant morbidity and mortality depending on the severity of the air leak and on the patient's cardiopulmonary reserve, as it impairs oxygenation and ventilation. Rapid detection and intervention are therefore essential. Chest radiography is the first-line and most commonly used imaging modality for diagnosis. However, pneumothorax may go undetected in 20% of cases, requiring thoracic CT. Researchers have evaluated the performance of radiologists and AI using large CXR datasets, with promising results. AI can be particularly helpful in settings that lack experienced clinicians or radiologists, both for detection and for severity grading. However, digital image quality and the presence of chest tubes have been shown to limit AI performance. On CXRs, AI models may misidentify a chest tube as a pneumothorax or focus on the tube instead of the actual air leak, reducing diagnostic accuracy.[6,38]

Various AI models have been developed for pneumothorax detection, and many of these models have received Food and Drug Administration (FDA) approval. The sensitivity of FDA-approved AI tools for diagnosing pneumothorax ranges from 84.3% to 94.6%, and specificity ranges from 87.9% to 95.1%. AI can be used primarily for triage and to provide a second opinion. AI requires less time to read radiographs than physicians do. Pneumothorax among patients presenting to the emergency department is very rare, when considering all radiographs obtained. An AI-enabled pneumothorax detection tool can rapidly identify true-positive cases and reduce waiting times for these cases in the emergency department.[39,40]

Studies comparing the performance of AI with that of radiologists are crucial for understanding AI's effectiveness and limitations in medical imaging. In a meta-analysis of 63 studies, the overall AUC was found to be 0.97 for both AI and physicians.[39] In this meta-analysis, the average sensitivity was 84% for AI and 85% for physicians; the average specificity was 96% for AI and 98% for physicians. In a multicenter retrospective study of 2,040 patients, Plesner et al.[41] reported that AI models detected pneumothorax with sensitivities of 63-90% and specificities of 98-100%. False-negative rates were comparable to those in radiology reports, but AI showed higher false-positive rates. Sensitivity decreased for small lesions, and specificity was lower for A-P chest radiographs and cases with multiple simultaneous pathologies. Kim et al.[42] developed an AI-based model to calculate pneumothorax area and compared its performance with that of an experienced radiologist and a gold-standard CT-based method. The measured pneumothorax rates were 5.41% for the radiologist and 8.45% for AI, representing a 3.04 percentage-point gap favoring the radiologist. However, no significant difference was observed between AI predictions and the gold standard ($P = 0.11$). The AI model successfully predicted the need for thoracostomy when the proportion of the chest radiograph occupied by pneumothorax exceeded 21.6%. Despite limitations such as a retrospective design and small sample size, the study highlights AI's potential utility in clinical practice. A study of 500 chest radiographs that investigated radiologists' performance with and without AI assistance found that unaided radiologists achieved the highest sensitivity (75.7%) and specificity (99-99.7%) in the diagnosis of pneumothorax.[43] AI alone achieved 99% specificity for pneumothorax and improved sensitivity for all findings compared with unaided readers. Additionally, AI assistance reduced average reading times by 25 seconds (from 81 to 56 seconds, $P < 0.001$) without affecting specificity. However, AI-assisted radiologists did not outperform AI alone, possibly because mistrust of AI led to accurate detections being dismissed. The study highlights AI's potential to enhance diagnostic sensitivity and efficiency across expertise levels while reducing reading times for chest radiographs.

Hillis et al.[40] evaluated the performance of an AI model designed to detect pneumothorax and tension pneumothorax. This AI model identified pneumothorax with 94.3% sensitivity, 92.0% specificity, and an AUC of 0.979, and detected tension pneumothorax with 94.5% sensitivity, 95.3% specificity, and an AUC of 0.987. When subgroup analyses were conducted for gender, age, patient's position, and projections, all subgroups achieved at least a sensitivity of 80%, and a specificity of 80% with an AUC of 0.95. Even in the presence or absence of potentially biasing findings, detection of both pneumothorax and tension pneumothorax maintained 80% sensitivity and 80% specificity, with an AUC of 0.95.

Thian et al.[44] investigated the performance and generalizability of an AI model on 2,931 chest radiographs. The mean AUC was 0.94, with a sensitivity of 88%, specificity of 88%, PPV of 71%, and NPV of 95%. The model demonstrated lower performance for small pneumothoraces compared to larger ones (AUC 0.88 vs. 0.96; $P = 0.005$). The model's performance did not differ between radiographs with and without chest tubes (AUC 0.95 vs. AUC 0.94; $P > 0.99$) or between radiographic projections (A-P vs. P-A; AUC 0.92 vs. AUC 0.96; $P = 0.05$). Thian et al.[44] showed that the AI model trained on a large database became generalizable after limiting factors unrelated to the training data were eliminated. In the study by Lee et al.[45], the overall PPV for AI-assisted pneumothorax diagnosis was reported to be 41.1%. The PPV was higher for P-A views (88.2%), but dropped to 20.1% for A-P views. Younger age, P-A projection, and larger pneumothorax were associated with increased PPV for AI-assisted pneumothorax diagnosis. In 31.3% of false-positive cases flagged by AI, the underlying cause was identified: 20.5% were due to skin folds, 5.6% to chest wall abnormalities, 3.1% to bullae, and 2.1% to rib opacities.

Hong et al.[46] evaluated 1,319 patients who underwent transthoracic lung biopsy, using an AI model to assess post-procedural chest radiographs for iatrogenic pneumothorax.

They found that the AI-assessed group had higher sensitivity (85.4% vs. 67.1%), NPV (96.8% vs. 91.3%), and accuracy (96.8% vs. 92.3%) than radiologists' interpretations ($P < 0.001$).[46] There were no significant differences between the groups in specificity and PPV ($P = 0.46$ and $P = 0.45$). In analyses based on pneumothorax volume, the AI model showed higher sensitivity in patients with pneumothorax <10% (74.5% vs. 51.4%, $P = 0.009$) and in those with pneumothorax 10-15% (92.7% vs. 70.2%, $P = 0.008$). Among patients with pneumothorax, the AI-assessed group required fewer catheters than the other group (2.4% vs. 5%, $P = 0.009$). At the conclusion of the study, the researchers suggested that the AI model could be used to diagnose iatrogenic pneumothorax on CXRs, allowing patients with small pneumothoraces to be diagnosed earlier and treated conservatively before air leakage progresses.

Artificial Intelligence in the Diagnosis of Malignant Pleural Mesothelioma

MPM is associated with asbestos exposure in 80% of cases. Although asbestos is banned in many countries, MPM incidence continues to rise. Given a poor prognosis and average survival of 9-14 months after diagnosis, rapid and accurate detection and assessment of treatment response are essential. Prognosis and treatment monitoring often rely on tumor segmentation in serial thoracic CT scans and interpretation of PET-CT images-processes prone to human error and a high workload for radiologists. 3D CT image analysis offers advantages for more accurate prognosis and treatment evaluation. Accordingly, various AI models have been developed and investigated for MPM applications.[47-50]

Segmentation of MPM is highly challenging, as it relies solely on density differences to distinguish tumor tissue from surrounding benign soft tissue. In Sensakovic et al.[47] study, a system was developed to automate segmentation and volumetric measurement. This method yielded 3D positional and volumetric information that was validated against 2D manual detections, significantly reducing human error and saving time. This allowed pleural abnormalities to be objectively monitored through serial imaging and for changes in size to be detected. In a study by Karapınar Şentürk and Çekiç[48], the performance of five AI models for diagnosing MPM was evaluated in terms of accuracy and sensitivity. The models' accuracies ranged from 80% to 100%, and sensitivities ranged from 50% to 100%. SVM and Artificial Neural Networks (ANN) demonstrated the highest accuracy and sensitivity (100%). However, because the majority of samples in this study belonged to non-diseased groups, these rates may differ in real-world settings. Another study investigated the performance of a 3D DCNN-based AI model that used PET-CT to differentiate MPM from benign pleural diseases.[49] Results from four datasets were analyzed: PET-CT-AI; radiologist interpretation; maximum standardized uptake value (SUV_{max}) quantitative method; and PET-CT combined with SUV_{max}, gender, and age, with AI. The AUC values were 82.5%, 85.4%, 88.1%, and 89.6. In the last protocol, sensitivity, specificity, and accuracy were 88.5%, 73.6%, and 82.4%, respectively. AI integrated solely with PET-CT demonstrated inferior performance compared with expert interpretation and SUV_{max}-based quantitative methods. However, AI integrating PET-CT, SUV_{max}, gender, and age demonstrated superior diagnostic performance compared with human interpretation and SUV_{max}-based methods. In a study by Er and Tanrikulu[50] involving 324 MPM cases, a newly developed AI system achieved an accuracy rate of 97.7%, outperforming the ANN algorithm. This algorithm has been reported to be an excellent auxiliary tool for diagnosing MPM.

No single PE biomarker provides sufficient diagnostic accuracy for MPM. Therefore, using multiple biomarkers could be a suitable approach to improve diagnostic efficiency. The integration of multiple biomarkers using AI models can enhance diagnostic yield. In a study involving 188 patients with undiagnosed PE, six ML-based AI models were tested using the biomarkers SMRP, CEA, and CYFRA21-1: LR, linear discriminant analysis, multivariate adaptive regression splines, KNN, GBM, and RF.[51] Among these algorithms, the LR model significantly improved the diagnostic accuracy for MPM, achieving an AUC of 0.97 and an accuracy of 91%. Similarly, another study examined four AI models for the differential diagnosis of MPM based on cytological analysis and tumor marker concentrations.[52] The logic learning machine (LLM) showed the highest performance, with an accuracy of 77.5%, while the accuracies of the other three methods —KNN, ANN, and decision tree models— ranged from 54.4% to 72.8%. Furthermore, the LLM achieved diagnostic accuracies of 79% for MPM, 66% for pleural metastasis, and 89% for benign pleural diseases.

DISCUSSION

Most of the current AI studies in pleural diseases are based on relatively small, single-center cohorts, which inherently limit their external validity and generalizability. While reported accuracy and AUC values are often high, these findings should be interpreted with caution, as the real-world performance of AI models may differ substantially in heterogeneous patient populations and across diverse clinical settings. Therefore, large-scale, multicenter validation studies are urgently needed to confirm their clinical applicability.

Strengths and Limitations of Artificial Intelligence in Pleural Diseases

Despite promising results, AI models in pleural diseases face several important limitations (Table 2). Data imbalance and selection bias may lead to overfitting, reducing the reliability of predictions when applied to broader populations. Another critical issue is explainability: the so-called "black-box" nature of DL algorithms limits clinicians' ability to understand how specific outputs are generated, thereby reducing clinicians' trust in the technology. Furthermore, the lack of transparency complicates clinical decision-making, especially in high-stakes scenarios such as invasive interventions. Finally, the majority of existing models are validated retrospectively, which underscores the need for prospective, multicenter studies with rigorous external validation before routine clinical adoption.

Implementation and Clinical Integration

For the rational integration of AI into pleural disease

Table 2. Strengths and limitations of artificial intelligence models in pleural diseases

Modality/AI application	Strengths	Limitations	References
Chest X-ray	Widely available, fast triage (PE, pneumothorax); FDA-cleared tools exist; high sensitivity/specificity in large datasets	Lower accuracy in small effusions/pneumothorax; susceptible to artifacts (skin folds, tubes); generalizability depends on dataset diversity	9-11,38-46
Chest CT	Excellent anatomical detail; robust segmentation with nnU-Net; reproducible volume analysis; useful for simple vs. complex PE classification	Consistently underestimates absolute fluid volume; requires calibration; high radiation exposure; most studies single-center with limited validation	12,13,17-19
PET-CT	Adds functional data for MPE vs. BPE and MPM prognosis; multimodal fusion improves performance; high reported AUC	Small sample sizes; expensive and limited availability; performance may drop in real-world heterogeneous populations	16,25,49
Cytology/WSI	AI systems (e.g., Aitrox) achieve performance comparable to expert cytopathologists; potential to reduce inter-observer variability; supports rapid triage	Dependent on quality of slide preparation; limited datasets; still requires pathologist oversight; "black box" decisions reduce trust	23,24
Lung ultrasound	Safe, bedside, radiation-free; AI reduces operator dependency; early studies show >90% accuracy	Limited training data: performance varies with probe settings and acquisition quality; segmentation less robust than CT	26,27
Tuberculous pleurisy models	AI + ADA and lab parameters improve diagnostic yield; potential non-invasive alternative to biopsy in resource-limited settings	Cut-off values differ by geography/epidemiology; small sample sizes; external validation rare	32-37
Malignant pleural mesothelioma	AI-based segmentation improves reproducibility in tumor volume monitoring; multimodal models integrate PET-CT + biomarkers with promising accuracy	Challenging tumor-benign tissue differentiation; many studies use retrospective datasets; real-world validation lacking	47-52

AI: artificial intelligence, CT: computed tomography, PET: positron emission tomography, WSI: whole-slide imaging, PE: pleural effusion, FDA: Food and Drug Administration, nnU-Net: no-new-U-Net segmentation framework, MPE: malignant pleural effusion, BPE: benign pleural effusion, MPM: malignant pleural mesothelioma, AUC: area under the curve, ADA: adenosine deaminase

management, stepwise frameworks should be considered. At the initial stage, AI can serve as a triage and screening tool, for example in chest radiography for effusion or pneumothorax detection. At the intermediate level, AI models may provide decision support by predicting the likelihood of malignancy and guiding the necessity of invasive procedures such as thoracentesis or biopsy. At the advanced level, AI may assist in treatment monitoring, such as evaluating changes in effusion volume or tumor burden in MPM. Importantly, AI should be viewed as an adjunctive decision-support system rather than a replacement for physician judgment, with outputs always interpreted within the broader clinical context. Our findings are consistent with the recent narrative review by Marchi et al.,[7] which also emphasized the importance of critically evaluating AI models, particularly regarding their generalizability and the necessity of developing frameworks for rational integration into clinical workflows.

AI applications for pleural diseases require careful planning for translation into clinical workflows. Intended clinical roles include triage and second-read tools for pneumothorax on chest radiographs, automated segmentation and volume tracking of PEs on CT, and cytology triage systems to prioritize suspicious samples for pathologist review.

• Integration: Seamless interoperability with Picture Archiving and Communication Systems and Electronic Health Records is essential. AI outputs should be provided in DICOM-compatible formats, incorporated into structured reports, and logged with audit trails to maintain accountability.

• Regulatory examples: The approval and deployment of FDA-cleared pneumothorax detection algorithms provide precedents for the integration of AI in pleural disease workflows. These experiences highlight the importance of prospective validation, calibration, and clinician oversight in safe implementation.

• Training: Adoption requires short, role-specific training sessions for radiologists, pulmonologists, and pathologists. Training should cover interpretation of confidence scores, handling of false positives, and recognition of model limitations.

• Cost and time-saving considerations: Evidence from real-world radiology practice suggests that AI triage can reduce report turnaround times and improve sensitivity. Automated CT segmentation and cytology triage also have the potential to reduce workload and shorten diagnostic delays. Formal cost-effectiveness analyses remain necessary to confirm the economic value of these tools in pleural disease management.

Future Directions

To accelerate translation of AI in pleural diseases into clinical practice, the following priorities are needed:

• Multicentre prospective trials with predefined thresholds and patient-relevant outcomes to validate diagnostic performance in real-world settings.

• Domain-shift and out-of-distribution testing across different scanners, institutions, and patient populations to ensure generalizability.

• Multimodal data fusion combining imaging (CT, PET-CT, USG), cytology, and clinical data to enhance predictive accuracy and robustness.

• Real-time USG guidance tools to support novice operators and improve procedure safety and efficiency.

• Economic evaluations and cost-effectiveness analyses to clarify resource savings and sustainability within healthcare systems.

• Bias and calibration checks at deployment sites, with continuous monitoring dashboards to maintain safety and fairness.

• Creation of open datasets and benchmarks for pleural imaging, cytology, and biomarker AI tasks to foster reproducibility and global collaboration.

• These steps will provide the field with a clear roadmap, bridging experimental success toward responsible clinical implementation.

CONCLUSION

Based on current evidence, the most promising AI applications in pleural diseases are listed in Table 3. These include radiological segmentation models for PE quantification, DL algorithms to differentiate malignant from benign effusions using CT and PET-CT, and ML-based classification systems utilizing pleural fluid biomarkers. Among these, ensemble models like XGBoost, RF, and SVM that combine clinical, radiological, and laboratory data have demonstrated high diagnostic accuracy—particularly in distinguishing MPE from TP, often with AUCs above 0.90. Automated segmentation tools such as nnU-Net and contrastive learning models have shown near-expert accuracy and consistency in volume and lesion analysis. Additionally, computer-aided cytological tools like Aitrox and CAD systems for MPE detection have matched the diagnostic performance of experienced cytopathologists, suggesting potential for workflow standardization. Altogether, these AI innovations offer non-invasive, rapid, and objective clinical support and may redefine diagnostic approaches in pleural disease management.

Although AI applications hold great promise for the future, their integration into routine clinical practice still faces several challenges. First, the algorithms require further optimization and performance improvement. Most current studies are based on single-center, small-scale datasets, lacking standardized performance metrics and limiting generalizability. To enable routine use, large-scale, multicenter studies with standardized protocols are essential, along with strong collaboration among patients, clinicians, institutions, and medical technology companies.

Table 3. Top performing artificial intelligence applications in pleural diseases

Application area	Model/algorithm	Key performance metrics	Readiness	References
PE diagnosis (CXR)	Deep learning	AUC: 97% Sensitivity: 95% Specificity: 97%	Research	[9]
PE volume measurement (CT)	nnU-Net segmentation model	DSC: 0.89 ICC ≥0.995	Research	[12,13]
Pneumothorax detection (CXR)	FDA-approved AI systems	Sensitivity: 84-94%, specificity: 88-95%	FDA-cleared	[39,40]
MPE-BPE differentiation (CT)	XGBoost + clinical variables	AUC >0.90 Sensitivity >85%	Research	[16]
Cytological MPE diagnosis	Aitrox (DCNN, WSI)	AUC: 0.95 Accuracy: 91.7%, specificity: 94.4%	Research	[24]
MPM diagnosis (PET-CT + markers)	LLM, LR, 3D DCNN	AUC: 0.89-0.97 Accuracy: 82-91%	Research	[49,51]

PE: pleural effusion, CXR: chest X-ray, CT: computed tomography, MPE: malignant pleural effusion, BPE: benign pleural effusion, MPM: malignant pleural mesothelioma, PET: positron emission tomography, nnU-Net: no-new-U-Net segmentation framework, FDA: Food and Drug Administration, DCNN: deep convolutional neural network, WSI: whole-slide imaging, LLM: logic learning machine, LR: logistic regression, AUC: area under the curve, DSC: dice similarity coefficient, ICC: intraclass correlation coefficient, 3D: three dimensional

Footnotes

Authorship Contributions

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

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. Mummadi SR, Stoller JK, Lopez R, Kailasam K, Gillespie CT, Hahn PY. Epidemiology of adult pleural disease in the United States. *Chest: Elsevier Inc.* 2021:1534-1551. **[Crossref]**
2. Sundaralingam A, Bedawi EO, Rahman NM. Diagnostics in pleural disease. *Diagnostics (Basel).* 2020;10(12):1046. **[Crossref]**
3. Addala DN, Rahman NM. Man versus machine in pleural diagnostics: does artificial intelligence provide the solution? *Ann Am Thorac Soc.* 2024;21(2):202-203. **[Crossref]**
4. Feller-Kopman D, Light R. Pleural disease. *N Engl J Med.* 2018;378(8):740-751. **[Crossref]**
5. Zhang C, Wu W, Yang J, Sun J. Application of artificial intelligence in respiratory medicine. *J Digit Health.* 2022;1(1):30-39. **[Crossref]**
6. Kalaiyarasan K, Sridhar R. Artificial intelligence in respiratory medicine: the journey so far - a review. *J Assoc Pulmonologist Tamil Nadu.* 2023;6(2):53-68. **[Crossref]**
7. Marchi G, Mercier M, Cefalo J, et al. Advanced imaging techniques and artificial intelligence in pleural diseases: a narrative review. *Eur Respir Rev.* 2025;34(176):240263. **[Crossref]**
8. Li D, Pehrson LM, Lauridsen CA, et al. The added effect of artificial intelligence on physicians' performance in detecting thoracic pathologies on CT and chest X-ray: a systematic review. *Diagnostics (Basel).* 2021;11(12):2206. **[Crossref]**
9. Chang J, Lin BR, Wang TH, Chen CM. Deep learning model for pleural effusion detection via active learning and pseudo-labeling: a multisite study. *BMC Med Imaging.* 2024;24(1):92. **[Crossref]**
10. Huang T, Yang R, Shen L, et al. Deep transfer learning to quantify pleural effusion severity in chest X-rays. *BMC Med Imaging.* 2022;22(1):100. **[Crossref]**
11. Zhou L, Yin X, Zhang T, et al. Detection and semiquantitative analysis of cardiomegaly, pneumothorax, and pleural effusion on chest radiographs. *Radiol Artif Intell.* 2021;3(4):e200172. **[Crossref]**
12. Sexauer R, Yang S, Weikert T, et al. Automated detection, segmentation, and classification of pleural effusion from computed tomography scans using machine learning. *Invest Radiol.* 2022;57(8):552-559. **[Crossref]**
13. Hwang EJ, Hong H, Ko S, et al. Accuracy of fully automated and human-assisted artificial intelligence-based CT quantification of pleural effusion changes after thoracentesis. *Radiol Artif Intell.* 2025;7(1):e240215. **[Crossref]**
14. Feller-Kopman DJ, Reddy CB, Gould MK, et al. Management of malignant pleural effusions: an official ATS/STS/STR clinical practice guideline. *Am J Respir Crit Care Med.* 2018;198(7):839-849. **[Crossref]**
15. Reddy CB, DeCamp MM, Diekemper RL, et al. Summary for clinicians: clinical practice guideline for management of malignant pleural effusions. *Ann Am Thorac Soc.* 2019;16(1):17-21. **[Crossref]**
16. Wang S, Tan X, Li P, et al. Differentiation of malignant from benign pleural effusions based on artificial intelligence. *Thorax.* 2023;78(4):376-382. **[Crossref]**
17. Ozcelik N, Ozcelik AE, Guner Zirih NM, Selimoglu I, Gumus A. Deep learning for diagnosis of malign pleural effusion on computed tomography images. *Clinics (Sao Paulo).* 2023;78:100210. **[Crossref]**
18. Li Y, Tian S, Huang Y, Dong W. Driverless artificial intelligence framework for the identification of malignant pleural effusion. *Transl Oncol.* 2021;14(1):100896. **[Crossref]**
19. Kim NY, Jang B, Gu KM, Park YS, Kim YG, Cho J. Differential diagnosis of pleural effusion using machine learning. *Ann Am Thorac Soc.* 2024;21(2):211-217. **[Crossref]**
20. Zhang Y, Wang J, Liang B, Wu H, Chen Y. Diagnosis of malignant pleural effusion with combinations of multiple tumor markers: a comparison study of five machine learning models. *Int J Biol Markers.* 2023;38(2):139-146. **[Crossref]**
21. Wei TT, Zhang JF, Cheng Z, Jiang L, Li JY, Zhou L. Development and validation of a machine learning model for differential diagnosis of malignant pleural effusion using routine laboratory data. *Ther Adv Respir Dis.* 2023;17:17534666231208632. **[Crossref]**
22. Loveland P, Christie M, Hammerschlag G, Irving L, Steinfort D. Diagnostic yield of pleural fluid cytology in malignant effusions: an Australian tertiary centre experience. *Intern Med J.* 2018;48(11):1318-1324. **[Crossref]**
23. Win KY, Choomchuay S, Hamamoto K, Raveesunthornkiat M, Rangsirattanakul L, Pongsawat S. Computer aided diagnosis system for detection of cancer cells on cytological pleural effusion images. *Biomed Res Int.* 2018;2018:6456724. **[Crossref]**
24. Xie X, Fu CC, Lv L, et al. Deep convolutional neural network-based classification of cancer cells on cytological pleural effusion images. *Mod Pathol.* 2022;35(5):609-614. **[Crossref]**
25. Wei J, Li P, Zhang H, Zhu R. Clone selection artificial intelligence algorithm-based positron emission tomography-computed tomography image information data analysis for the qualitative diagnosis of serous cavity effusion in patients with malignant tumors. *J Healthc Eng.* 2021;2021:4272411. **[Crossref]**
26. Tsai CH, van der Burgt J, Vukovic D, et al. Automatic deep learning-based pleural effusion classification in lung ultrasound images for respiratory pathology diagnosis. *Phys Med.* 2021;83:38-45. **[Crossref]**
27. Vukovic D, Wang A, Antico M, et al. Automatic deep learning-based pleural effusion segmentation in lung ultrasound images. *BMC Med Inform Decis Mak.* 2023;23(1):274. **[Crossref]**
28. Chan KKP, Lee YCG. Tuberculous pleuritis: clinical presentations and diagnostic challenges. *Curr Opin Pulm Med.* 2024;30(3):210-216. **[Crossref]**
29. Global Tuberculosis Report 2024. Accessed September 7, 2025. **[Crossref]**
30. Jeon D. Tuberculous pleurisy: an update. *Tuberc Respir Dis (Seoul).* 2014;76(4):153-159. **[Crossref]**
31. Shaw JA, Diacon AH, Koegelenberg CFN. Tuberculous pleural effusion. *Respirology.* 2019;24(10):962-971. **[Crossref]**
32. Ren Z, Hu Y, Xu L. Identifying tuberculous pleural effusion using artificial intelligence machine learning algorithms. *Respir Res.* 2019;20(1):220. **[Crossref]**
33. Garcia-Zamalloa A, Vicente D, Arnay R, et al. Diagnostic accuracy of adenosine deaminase for pleural tuberculosis in a low prevalence setting: a machine learning approach within a 7-year prospective multi-center study. *PLoS One.* 2021;16(11):0259203. **[Crossref]**

34. Darooei R, Sanadgol G, Gh-Nataj A, et al. Discriminating tuberculous pleural effusion from malignant pleural effusion based on routine pleural fluid biomarkers, using mathematical methods. *Tanaffos*. 2017;16(2):157-165. **[Crossref]**
35. Liu Y, Liang Z, Yang J, et al. Diagnostic and comparative performance for the prediction of tuberculous pleural effusion using machine learning algorithms. *Int J Med Inform*. 2024;182:105320. **[Crossref]**
36. Lee JH, Choi CM, Park N, Park HJ. Classification of pleural effusions using deep learning visual models: contrastive-loss. *Sci Rep*. 2022;12(1):5532. **[Crossref]**
37. Li C, Hou L, Sharma BY, et al. Developing a new intelligent system for the diagnosis of tuberculous pleural effusion. *Comput Methods Programs Biomed*. 2018;153:211-225. **[Crossref]**
38. Irmici G, Cè M, Caloro E, et al. Chest X-ray in emergency radiology: what artificial intelligence applications are available? *Diagnostics (Basel)*. 2023;13(2):216. **[Crossref]**
39. Sugibayashi T, Walston SL, Matsumoto T, Mitsuyama Y, Miki Y, Ueda D. Deep learning for pneumothorax diagnosis: a systematic review and meta-analysis. *Eur Respir Rev*. 2023;32(168):220259. **[Crossref]**
40. Hillis JM, Bizzo BC, Mercaldo S, et al. Evaluation of an artificial intelligence model for detection of pneumothorax and tension pneumothorax in chest radiographs. *JAMA Netw Open*. 2022;5(12):e2247172. **[Crossref]**
41. Plesner LL, Müller FC, Brejnebøl MW, et al. Commercially available chest radiograph AI tools for detecting airspace disease, pneumothorax, and pleural effusion. *Radiology*. 2023;308(3):e231236. **[Crossref]**
42. Kim D, Lee JH, Kim SW, et al. Quantitative measurement of pneumothorax using artificial intelligence management model and clinical application. *Diagnostics (Basel)*. 2022;12(8):1823. **[Crossref]**
43. Bennani S, Regnard NE, Ventre J, et al. Using AI to improve radiologist performance in detection of abnormalities on chest radiographs. *Radiology*. 2023;309(3):e230860. **[Crossref]**
44. Thian YL, Ng D, Hallinan JTPD, et al. Deep learning systems for pneumothorax detection on chest radiographs: a multicenter external validation study. *Radiol Artif Intell*. 2021;3(4):e200190. **[Crossref]**
45. Lee S, Kim EK, Han K, Ryu L, Lee EH, Shin HJ. Factors for increasing positive predictive value of pneumothorax detection on chest radiographs using artificial intelligence. *Sci Rep*. 2024;14(1):19624. **[Crossref]**
46. Hong W, Hwang EJ, Lee JH, Park J, Goo JM, Park CM. Deep learning for detecting pneumothorax on chest radiographs after needle biopsy: clinical implementation. *Radiology*. 2022;303(2):433-441. **[Crossref]**
47. Sensakovic WF, Armato SG 3rd, Straus C, et al. Computerized segmentation and measurement of malignant pleural mesothelioma. *Med Phys*. 2011;38(1):238-244. **[Crossref]**
48. Karapınar Şentürk Z, Çekiç N. A machine learning based early diagnosis system for mesothelioma disease. *Düzce University Journal of Science Technology*. 2020;8(2):1604-1611. **[Crossref]**
49. Kitajima K, Matsuo H, Kono A, et al. Deep learning with deep convolutional neural network using FDG-PET/CT for malignant pleural mesothelioma diagnosis. *Oncotarget*. 2021;12(12):1187-1196. **[Crossref]**
50. Er O, Tanrikulu AÇ. Use of artificial intelligence techniques for diagnosis of malignant pleural mesothelioma. *Dicle Med J*. 2015;42(1):5-11. **[Crossref]**
51. Niu Y, De Hu Z. Diagnostic accuracy of pleural effusion biomarkers for malignant pleural mesothelioma: a machine learning analysis. *J Lab Precis Med*. 2021;6:4. **[Crossref]**
52. Parodi S, Filiberti R, Marroni P, et al. Differential diagnosis of pleural mesothelioma using logic learning machine. *BMC Bioinformatics*. 2015;16(Suppl 9):3. **[Crossref]**

Thorac Res Pract. 2026;27(1):68-74 **DOI:** 10.4274/ThoracResPract.2026.2026-1-6

Position Statement

Turkish Thoracic Society Declaration on Peace and Health the Importance of Breathing in a World Without War

Pelin Duru Çetinkaya[1], Zehra Nur Töreyin[2], Hande Yüce Özdemir[3], Yeşim Yasin[4], Oğuz Kılınç[5], Arzu Yorgancıoğlu[6], Haluk C. Çalışır[7], Çağlar Çuhadaroğlu[8], Nurdan Köktürk[9], Ezgi Demirdöğen[10], Nesrin Öcal[11], Züleyha Bingöl[12], Aslı Gürün Kaya[13], Dorina Esendağlı[14], Ethem Yıldız[15], Mehmet Baran Balcan[16], Barış Çil[17], Hacıali Kılıçgün[18], Erencan Begiç[19]

[1]TTD Advocacy Committee Member, Department of Chest Diseases, Çukurova University Faculty of Medicine, Adana, Türkiye
[2]TTD Advocacy Committee Member, Department of Chest Diseases, Division of Occupational Diseases, Ege University Faculty of Medicine, İzmir, Türkiye
[3]TTD Advocacy Committee Secretary, Independent Researcher, Chest Diseases, Ankara, Türkiye
[4]TTD Advocacy Committee Member, Department of Public Health, Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, İstanbul, Türkiye
[5]TTD 2nd Vice President, TTD Advocacy Committee Member, Department of Chest Diseases, İzmir University of Economics Faculty of Medicine, İzmir, Türkiye
[6]TTD Advocacy Committee Member, Department of Chest Diseases, Manisa Celal Bayar University Faculty of Medicine, Manisa, Türkiye
[7]TTD Advocacy Committee Chairman, Department of Chest Diseases, Bezmialem Vakıf University Faculty of Medicine, İstanbul, Türkiye
[8]TTD President, Department of Chest Diseases, Acıbadem Mehmet Ali Aydınlar University Faculty of Medicine, İstanbul, Türkiye
[9]TTD 1st Vice President, Department of Chest Diseases, Gazi University Faculty of Medicine, Ankara, Türkiye
[10]TTD General Secretary, Department of Chest Diseases, Bursa Uludağ University Faculty of Medicine, Bursa, Türkiye
[11]TTD Board Member, Department of Chest Diseases, University of Health Sciences Türkiye, Gülhane Faculty of Medicine, Ankara, Türkiye
[12]TTD Board Member, Department of Chest Diseases, İstanbul University, İstanbul Faculty of Medicine, İstanbul, Türkiye
[13]TTD Board Member, Department of Chest Diseases, Ankara University Faculty of Medicine, Ankara, Türkiye
[14]TTD Board Member, Department of Chest Diseases, Başkent University Faculty of Medicine, Ankara, Türkiye
[15]TTD Board Member, Department of Chest Diseases, Trakya University Faculty of Medicine, Edirne, Türkiye
[16]TTD Board Member, Department of Chest Diseases and Intensive Care, Marmara University Faculty of Medicine, İstanbul, Türkiye
[17]TTD Board Member, Clinic of Chest Diseases, Mardin Training and Research Hospital, Mardin, Türkiye
[18]TTD Board Member, Department of Surgical Medical Sciences, Division of Thoracic Surgery, Bolu Abant İzzet Baysal University Faculty of Medicine, Bolu, Türkiye
[19]TTD Board Member, Department of Chest Diseases, Dokuz Eylül University Faculty of Medicine, İzmir, Türkiye

Cite this article as: Çetinkaya PD, Töreyin ZN, Yüce Özdemir H, et al. Turkish Thoracic Society Declaration on Peace and Health the importance of breathing in a world without war. *Thorac Res Pract.* 2026;27(1):68-74

ABSTRACT

War legitimizes the right to kill for its participants, equating victory with moral justification; however, this perceived legitimacy is largely constructed on the deaths of civilians and other living beings. Medicine, by its very nature, is grounded in the principles of preserving life, doing no harm, and adhering to scientific and ethical standards, and is therefore fundamentally opposed to the philosophy of war. Wars lead, directly and indirectly, to deaths, disability, displacement, poverty, and long-term public health disasters. Attacks on health facilities, restrictions on access to food and water, environmental destruction, and air pollution disproportionately affect vulnerable populations, particularly women and children. The psychological consequences of war often evolve into persistent neuropsychiatric conditions, including post-traumatic stress disorder, depression, and moral injury. Peace is not merely the absence of war; it is a prerequisite for health, environmental sustainability, and social well-being. The concept of positive peace strengthens health indicators through well-functioning institutions, equitable resource distribution, and the protection of human rights. Within this framework, health professionals and civil society organizations are critical actors in advocating for peace and in making human rights violations visible. The Turkish Thoracic Society regards peace as an indispensable condition for public and respiratory health, affirms its opposition to all wars—particularly the ongoing atrocities in Gaza—and reaffirms its commitment to sustained, health-based advocacy for peace.

KEYWORDS: Peace, war, conflict, displacement, respiratory health, human right, Gaza, peacebuilding, environment

Received: 09.01.2026

Accepted: 22.01.2026 ***Publication Date:*** 30.01.2026

Corresponding author: Assoc. Prof. Pelin Duru Çetinkaya, MD, e-mail: pelindurucetinkaya@hotmail.com

INTRODUCTION

War, by its nature, confers upon its participants a perceived legitimacy to kill. Not only among those directly involved, but also within dominant historical narratives that shape our present, "winning" and "victory" are often portrayed as forms of moral justification. Nevertheless, such notions of victory and moral justification are built on the deaths of countless individuals and other living beings on both sides. The violation of the right to life thus becomes normalized and, ultimately, legitimized. Medicine, by contrast, is fundamentally grounded in the preservation of life. It is both an art informed by scientific methodology and a moral stance rooted in the principle of doing no harm. The existential philosophies of war and of medicine are inherently incompatible. Wars are commonly driven by a combination of economic interests, geopolitical ambitions, control over resources (such as oil), and the subjugation of populations.[1] Cultural, religious, racial, and ideological differences are frequently exploited to persuade societies that war is inevitable or necessary. Once war begins, the values humanity cultivated over centuries can be rapidly eroded. Dominant historical discourses often present wars as unavoidable facts of the past or as natural components of human history, encouraging societies to perceive them as ordinary. The phrase "history is written by the victors" obscures massacres and civilian deaths committed by the victors, implicitly suggesting that the defeated deserved their fate. Meanwhile, the deaths, injuries, displacement, and impoverishment of countless individuals remain hidden behind narratives of success and triumph. Those who control economic power continue to cause death, disability, and disease among populations in pursuit of greater dominance.

Advancements in technology have also transformed weapons and methods of warfare, amplifying both their destructive capacity and their short- and long-term adverse health effects.[2] War is inherently destructive to human and societal life, as it seeks to achieve objectives through the annihilation or permanent impairment of individuals. Parties to conflict frequently target health facilities and personnel or disrupt health service delivery, thereby obstructing the fundamental right to health.[3]

Main Points

- War vs. medicine: War legitimizes killing and equates victory with moral justification, whereas medicine is fundamentally committed to preserving life and adhering to ethical standards.
- Consequences of war: Direct and indirect harms include death, disability, displacement, poverty, long-term public health crises, environmental destruction, and psychological trauma, disproportionately affecting vulnerable groups such as women and children.
- Peace as a prerequisite: Peace is more than the absence of conflict; it is essential for health, environmental sustainability, and social well-being. Positive peace strengthens health outcomes through equitable institutions, resource distribution, and human rights protection.
- Role of health professionals: Health professionals and civil society play a critical role in advocating for peace and in highlighting human rights violations.
- Turkish Thoracic Society position: The Turkish Thoracic Society considers peace indispensable for public and respiratory health and actively opposes war, committing to health-based advocacy for peace, particularly in response to the situation in Gaza.

Health Impacts of War

Wars cause death both directly and indirectly, through injuries, diseases, disabilities, and the breakdown of health services. They lead to mass displacement and forced migration, with profound and cumulative long-term health consequences. Bombs and missiles directly take the lives of individuals, killing thousands of people who are not involved in hostilities. Approximately 80% of those killed in wars are civilians, with women, older adults, and children—among the most vulnerable populations—disproportionately affected.[4]

Children are particularly defenseless during armed conflict. Every child has the right to grow up with their parents in a healthy environment that provides homes, schools, playgrounds, and health facilities. The 1989 United Nations Convention on the Rights of the Child explicitly affirms children's rights to healthy development and access to the highest attainable standard of medical care, while emphasizing obligations under international humanitarian law to protect civilians during armed conflicts. In this context, article 37 clearly states that no child shall be subjected to unlawful or arbitrary deprivation of liberty.[5-7]

During armed conflict, homes, factories, and food- and water-storage facilities are destroyed, often by fire, depriving people of their right to safe shelter. Populations living in conflict zones are forced to endure unhealthy living conditions or displacement, and, due to inadequate housing face an increased risk of infectious diseases. Approximately 25% of those who survive bombings subsequently die from poor environmental conditions and infectious diseases. Individuals forced to flee conflict areas often live in refugee camps, where mortality rates increase ten- to twentyfold. While 13% of refugees lack access to primary health care services, this proportion is only 6% among host communities. Similarly, 9% of refugees have no access to a family physician, compared with 4% of the local population.[4]

During armed conflicts, hospitals and health facilities may be bombed, health workers lose their lives, and access to health care is severely disrupted. As a result, both individuals injured during attacks and the general population experience impaired access to routine health services. Access to medications and essential preventive health services is also interrupted. Restrictions on access to food further exacerbate hunger, malnutrition, and violence. These consequences disproportionately affect the poorest and most vulnerable segments of society. The destruction caused by armed conflict severely degrades air quality. Explosives, fuels, and the collapse of buildings lead to increased levels of particulate matter and other toxic substances that pose serious threats to respiratory health. For example, studies conducted in conflict zones such as Gaza, Syria, and Ukraine have demonstrated marked increases in chronic obstructive pulmonary disease, asthma, and acute respiratory infections during periods of war. In addition, the collapse of

health systems impedes early diagnosis and treatment of infectious diseases, such as tuberculosis and pneumonia. Among displaced populations, malnutrition, overcrowded living conditions, and air pollution further accelerate the spread of these diseases.[8,9] Various weapons and explosives used during armed conflict affect not only humans but all living organisms; beyond their immediate destructive effects, they contaminate the environment by releasing toxic metals and chemicals. By rendering agricultural land unusable, they also disrupt the food chain and threaten long-term food security.

Wars threaten not only physical health but also mental health. The loss of parents and children, limb amputations, destruction of living environments, persistent insecurity, and fear of death contribute to the development of a range of neuropsychiatric conditions, including post-traumatic stress disorder (PTSD), depression, anxiety, and moral injury. These psychological effects are not confined to the period of active conflict; rather, they result in sustained changes in emotional response systems and may persist for many years after the war has ended. While PTSD is a well-recognized and extensively studied condition, moral injury is a relatively new concept. Moral injury is defined as a profound cognitive and emotional response following events that violate an individual's deeply held moral values and beliefs. Potentially morally injurious events include situations in which individuals perpetrate acts that conflict with their moral convictions, fail to prevent such acts, witness them, or become aware of them. Moral injury has been associated not only with PTSD symptoms but also with emotional distress characterized by self-harm, self-defeating behaviors, and pervasive feelings of hopelessness. These neuropsychiatric consequences often continue for many years after the cessation of armed conflict.[10-13]

Throughout human history, wars have caused the deaths, disabilities, illnesses, and profound suffering of millions of people. Among these, the atrocities committed by the Nazis occupy a prominent place in collective memory, both because they occurred in the relatively recent past and because they have been extensively documented in research, archival records, literature, film, and works of art. The fact that such brutality could take place at the very center of Europe—often portrayed as the cradle of enlightenment and civilization, shaped by the contributions of countless scientists, philosophers, and artists—forced humanity into a profound confrontation with itself. This reckoning prompted the development of new legal frameworks and human rights protections, based on respect for inherent human dignity. Following the trials of Nazi leaders, genocide was formally defined and codified in international law as the gravest crime against humanity. While the horrors of the Nazi genocide remain deeply embedded in our collective memory, and despite the belief that humanity would never again witness such inhumanity, the world today faces a genocide unfolding before its eyes in Gaza. In the aftermath of the Nazi genocide, humanity faces the unsettling reality that the concepts and rights established to safeguard human life and dignity can once again be rendered meaningless. At a time when numerous wars and armed conflicts continue worldwide and within our region—each accompanied by its own human tragedies—the openly indiscriminate and relentless assault on the civilian population in Gaza demands particular and urgent attention.

The Humanitarian Toll of the Gaza Genocide

Since 8 October 2023, more than 70,000 people have reportedly been killed in Gaza, the majority of whom are civilians, approximately half of whom are women and children. Despite reports that a ceasefire came into effect on 10 October 2025, people continue to be killed daily. United Nations–supported experts have warned that Gaza is currently experiencing the worst-case famine scenario.

The Integrated Food Security Phase Classification (IPC) is an innovative, multi-stakeholder global initiative aimed at improving the analysis of food security and nutrition. It brings together 21 organizations and intergovernmental bodies, including governments, regional institutions, and international organizations. Through the IPC framework, the severity and extent of acute food insecurity, chronic food insecurity, and acute malnutrition are assessed impartially to inform urgent response and action planning.[14] According to the IPC, there is growing evidence that widespread hunger, malnutrition, and disease in Gaza are contributing to excess mortality, primarily from hunger and malnutrition. At least 30% of children have acute malnutrition. As a result of starvation alone or of the interaction between malnutrition and disease, two per 10,000 people and four per 10,000 children die each day.[15,16]

Children in Gaza are unable to adequately access health services due to ongoing conflict and attacks. Their health continues to deteriorate as a result of environmental damage caused by the war. The bombing of hospitals and the destruction or abandonment of health facilities have disrupted disease surveillance and control systems, and the continuity of care, further undermining the capacity to provide treatment and accelerating the breakdown of the health system.[17-19]

The war in Gaza has also imposed a substantial mental health burden on individuals, severely undermining overall societal well-being. High rates of mental health disorders related to violence, loss, and displacement have been documented in Gaza.[20] Factors such as forced displacement, the loss of family members, physical injuries, and the severe shortage of mental health care resources expose children to a continuous cycle of violence and trauma, profoundly affecting their psychological well-being.[21]

In Gaza, 161,600 housing units have been completely destroyed, 81,000 have been rendered uninhabitable, and an additional 194,000 have sustained partial damage. Approximately 88% of infrastructure has been damaged; schools and universities have been bombed; cultural heritage sites have been destroyed. Nearly 90,000 Palestinian university students in Gaza have been unable to attend university. More than 60% of schools, almost all universities, and numerous bookstores and libraries have been damaged or destroyed. Hundreds of teachers and academics—including university deans and leading Palestinian scholars—have been killed. These systematic attacks have completely eliminated the future educational prospects of children and young people in Gaza.[1] Moreover, the large quantities of bombs and explosives dropped daily on Gaza are expected to cause toxic contamination of the ecosystem, affect the lives of people and other organisms for generations to come, and render the region increasingly unsafe and unsuitable for sustained human life.

Civilians in Gaza, confronted with daily horrors and a profound sense of helplessness, are likely experiencing increasingly severe psychological distress. Children who are critically affected during key stages of development face a heightened risk of long-term mental and physical health problems. The psychological, cultural, and epigenetic dimensions of the intergenerational transmission of trauma may result in suffering that extends across generations. Addressing the mental health needs of these populations is therefore essential to mitigating the deep and lasting impacts of war.[22]

Moreover, even when they survive, many are left without protection or supervision and are forced to struggle to survive on their own; at the same time, they continue to die from hunger, disease, lack of access to treatment, or direct attacks. Both situations represent profound humanitarian failures, demonstrating that existing efforts remain fundamentally insufficient. When societies fail to implement effective prevention strategies and are unable to alter this reality, they contribute to the continuation of harm through inaction.[23]

However, international humanitarian law, as set out in the Geneva Conventions, clearly establishes that civilians and civilian objects must be protected and may not be targeted.

- Civilian populations and civilian objects, including homes, markets, and residential buildings, must not be targeted; only military objectives may be lawfully attacked.

- Hospitals, health facilities, ambulances, and health personnel are explicitly protected; attacks on facilities providing care to the sick and wounded are prohibited.

- Objects and infrastructure indispensable to the survival of the civilian population—such as food supplies, agricultural areas, drinking water installations, irrigation systems, and electricity networks—are protected. Deliberate and intentional attacks on such objects, including actions aimed at starving or depriving civilians of water, are prohibited.

- Schools and educational institutions—as civilian educational facilities—must not be targeted; students and educators are entitled to protection.

- Places of worship and sacred sites are protected under international law.

- Cultural heritage and historical property are protected under the 1954 Hague Convention and its Protocols; intentional attacks on cultural property during armed conflict are prohibited.[24]

Unlawful attacks, including deliberate attacks against civilian targets and indiscriminate or disproportionate attacks, must be investigated as war crimes by international criminal mechanisms or national courts.[25] Numerous reports document that the Israeli government has committed war crimes by bombing hospitals, schools, and civilian-populated areas, thereby depriving innocent civilians, including children and women, in Gaza of their right to life. For a ceasefire to evolve into lasting peace, the government of Israel must be held accountable and tried for the crime of genocide.

The Concept of Peace

Health is not merely the absence of disease or disability; it is a state of complete physical, mental, and social well-being.[26] This definition clearly demonstrates how closely the concept of peace is intertwined with health. Peace is not simply the absence of war; it is a prerequisite for a healthy and sustainable life.

Interstate and intrastate wars, economic crises, human rights violations, social injustices, and violence arising from adverse conditions such as disease, injury, and death constitute a reality faced by societies worldwide today. Conversely, according to the understanding articulated by Galtung[27], the founder of peace studies as an academic discipline, the absence of such a broad spectrum of violence itself constitutes peace.

Peace enables societies to strengthen their health systems, expand preventive health services, and implement effective environmental policies. Routine immunization programs, early detection initiatives, and tobacco control strategies are sustainable under peaceful conditions. Thus, peace is not merely a state of non-violence but a way of life that enables the sustainability of health, environmental balance, and societal resilience.

The right of every individual to attain the highest possible standard of health, without discrimination based on race, religion, political belief, economic status, or social condition, is a fundamental human right. Health represents a shared objective that unites all people, regardless of political views, gender, ethnicity, or religion. These simple yet powerful rights-based approach, which underpins Universal Health Coverage is also recognized by the World Health Organization as a fundamental condition for peace.[28]

Negative peace is defined as the absence of war and direct violence; however, it does not necessarily reflect a society's levels of stability or social cohesion.

Positive peace, by contrast, encompasses not only the absence of violence but also the attitudes, institutions, and structures that enable peace, prosperity, and sustainable development. It is characterized by a more durable form of peace, built upon sustained investments in economic development and institutions and on social norms that support peaceful coexistence. Positive peace can be used to assess a society's resilience, its capacity to absorb shocks without descending into conflict or relapsing into violence.[29]

The positive peace index (PPI) developed by the Institute for Economics & Peace (IEP) is based on eight key pillars that measure these dimensions: well-functioning government, low levels of corruption, a sound business environment, equitable distribution of resources, acceptance of the rights of others, free flow of information, high levels of human capital, and good relations with neighboring countries. Improvements in positive peace not only strengthen peace itself but are also associated with a range of desirable societal outcomes, including higher GDP growth, improved measures of well-being, greater resilience, and more cohesive societies.[30]

According to the IEP's 2024 Positive Peace Report, Türkiye experienced a 7.8% decline in its PPI score between 2013 and 2022, making it the country with the largest decline in the European region. This decline was particularly pronounced in the pillars of "well-functioning government" and "low levels of corruption." The attempted coup in 2016 and the economic crisis following 2018 weakened societal resilience, and in the 2024 global ranking, Türkiye placed 91st with a PPI score of 3.[29,31]

Positive peace possesses the capacity to anticipate, prevent, and sustain stability. It allows societies to foresee vulnerabilities to crises, serves as a tool for the prevention of violence, and offers long-term, cost-effective solutions. A well-functioning government ensures accessible and high-quality health services; equitable resource distribution guarantees equal service provision; and the free flow of information supports health literacy.[2]

Environmental Peace

Peace is not limited to relationships among human beings; it also encompasses relationships between human beings and the environment. Environmental peace refers to a mode of living that is harmonious with nature and grounded in sustainability. Threats such as air pollution, climate change, and environmental degradation are increasingly recognized as the non-war threats of the modern era.[2]

Reducing carbon emissions, expanding the use of renewable energy, and implementing clean air policies are not only environmental priorities but also health-centered strategies essential for sustaining peace at the global level. Lung health is one of the most direct and measurable indicators of environmental peace, because respiratory health is strongly influenced by air quality, ecological balance, and climate stability.

The Role of Civil Society Organizations in Peacebuilding

Conflict and peace are phenomena that have coexisted throughout human history—evolving, transforming, and at times replacing one another—while continually being examined and reinterpreted. Peace represents one of humanity's most fundamental aspirations: a condition that societies seek to establish at every level and across all domains. Peace studies, as a field, emerged particularly in the aftermath of world wars, driven by efforts of international organizations, academic institutions, and civil society actors to prevent the recurrence of such catastrophes.

Civil society organizations (CSOs) are key actors in the defense of human rights and the redress of human rights violations. They raise their voices against abuses, advocate for the rights of victims, and contribute to the protection and promotion of human rights standards at both national and international levels. The contributions of CSOs to peacebuilding are realized through multiple functions, including protection, monitoring, advocacy, socialization, social cohesion, facilitation, and service provision. While these functions may involve engagement with elite-level actors, CSOs primarily operate through interaction with grassroots communities and mid-level actors.

The selection, feasibility, and effectiveness of peacebuilding functions undertaken by CSOs vary depending on the specific context, situation, and society. These functions are also shaped by factors such as the level of violence, the stance of state authorities, media influence, diaspora support, and the positions of external political actors and forces.

Depending on their capacities, CSOs may implement one, several, or, in rare cases, all of these functions during different phases of peacebuilding. Advocacy activities are among the most common forms of civil society engagement and include mobilizing the public to prevent conflict, demanding regime change, protesting human rights violations, condemning structural violence, and drawing global attention to coups violence, massacres, genocide, or attacks in specific regions. Such efforts often take the form of mass demonstrations, public gatherings, protests, and non-violent civic actions.

Another form of civil society advocacy involves lobbying efforts aimed at influencing peace negotiations and safeguarding the rights of affected parties. The ability of CSOs to carry out these functions is closely linked to contextual factors, including governmental attitudes, the intensity of violence, media dynamics, donor engagement, diaspora involvement, and the influence of external actors.

The number of CSOs capable of performing these peacebuilding functions—many of which require a high degree of specialization and professionalization—has increased internationally in recent years. However, in Türkiye, the current distribution and capacity of CSOs are insufficient to provide the infrastructure needed to support peacebuilding effectively.[32]

The Role of Health Professionals in Promoting Peace for Health

Health professionals—particularly physicians—whose practice is grounded in the principles of 'do no harm' and of preserving life, and who serve as trusted representatives of autonomy and ethical responsibility, occupy a unique and dedicated role in promoting peace. Because the foundational philosophy of medicine centers on disease prevention and the avoidance of preventable deaths, health professionals must consistently oppose war and other forms of armed conflict. This ethical stance renders the promotion of peace not merely an option, but a professional duty for health workers. At times, health professionals may also serve as mediators in conflict settings.

The responsibility of physicians extends beyond treating those injured in war; it fundamentally includes advocating for peace and opposing wars and conflicts that threaten human life and health. Much like their role in combating tobacco use, health professionals must actively engage in advocacy for peace. Such advocacy should not be limited to the pursuit of negative peace—the mere absence of armed conflict—but also encompass broader commitments to social well-being, human rights, democratic governance, improved living standards, and societal welfare. Moreover, health professionals can contribute to peace by documenting and reporting human rights violations, thereby supporting efforts to bring such abuses to an end.[33,34]

Despite their central role in safeguarding human health, physicians rarely advocate at the global level to prevent or end wars. To fulfill such a role effectively, the health professions must fully recognize the extensive and multifaceted impacts of war on population health.[1] Therefore, medical education and training curricula are of critical importance in preparing future health professionals to engage meaningfully in peacebuilding. Peace education and human rights–based approaches should therefore be integrated into health education programs.[35]

In addition, given the profound effects of war on health—and particularly its specific consequences for the respiratory system, including chemical exposures, dust, air pollution, and the effects of psychological trauma on respiratory health—the discipline of pulmonology should incorporate a mandatory curriculum component on "lung health in conflict and war settings."

Caglayan[36], in the article titled *"How Can Health Be More Effective in Peace Work in Türkiye: Introducing Peace Through Health,"* examines how the peace through health (PtH) approach can be applied in Türkiye, how the health sector can play a more effective role in peacebuilding, and how this framework may be integrated into university curricula. Since the 1980s, the PtH concept has emphasized the interdependence of health and peace, advocating for the active involvement of health professionals in peace processes and highlighting the potential of health services to contribute to peace in contexts such as war, migration, pandemics, and disasters. Caglayan[36] notes that, despite Türkiye's strategic regional position as a potential key actor for PtH, theoretical work—including academic publications and educational curricula—remains very limited. The absence of PtH courses in universities underscores the need for interdisciplinary collaboration among medicine, law, social sciences, and academia to institutionalize PtH and integrate it into medical education. Strengthening Türkiye's contribution to regional PtH diplomacy is also emphasized.[36]

Wars are among the most devastating threats to public health. They kill, injure, displace, and profoundly affect millions of people. Interstate wars fuel civil wars and perpetuate cycles of violence. Peacebuilding is not merely a political or diplomatic endeavor; it is a complex, long-term, and multidimensional process that encompasses cultural, environmental, social, and individual transformation. Peacebuilding is a critical approach to the creation of a culture of peace. Ultimately, peacebuilding is about the continuity of life on Earth. Preventing violence and achieving universal peace are imperatives.[37]

The peace agenda must not remain confined to the monopoly of the United Nations Security Council; rather, the process of building consensus should serve as a meeting point for the global community. A human rights–based approach alone will be insufficient to eliminate wars, given the limited efforts currently devoted to the protection of human rights. Therefore, the global development agenda must commit to ending all wars by 2030 and should explicitly include goals aimed at de-escalation and deterrence of armed conflict.[38] Immediate action is required to achieve a world without war.

Turkish Thoracic Society Declaration on Peace and Health

The Advocacy Committee of the Turkish Thoracic Society (TTS) organized a meeting entitled "Peace and Lung Health" on 1 September 2025, in observance of the International Day of Peace. Following this meeting, the Committee recognized the need to formally articulate the values advocated by the TTS regarding peace, health, and lung health. Accordingly, the TTS declares the following:

• The TTS recognizes peace as a fundamental prerequisite for public health.

• The society advocates for the integration of peace policies with public health strategies, given their positive impact on the right to life and on overall societal well-being.

• The society will actively and consistently advocate, within scientific platforms and the public sphere, that peace is a public health issue and a prerequisite for health, particularly for lung health.

• To enable physicians trained in our country to take a clear stance in favor of peace and health, the Society affirms that the health impacts of war—especially its consequences for the respiratory system, including chemical exposures, dust, air pollution, and psychological trauma and its effects—should be integrated into both pre-specialty and postgraduate medical education programs.

• The society will encourage and support all research initiatives related to peace and lung health.

• The society will promote activities aimed at raising societal awareness by using real patient experiences and case presentations from current and past war settings as educational materials.

• The right to life is inviolable. Accordingly, the Society declares its opposition to all forms of genocide—particularly in Gaza—and reaffirms its commitment to intensify its efforts to protect the right of all people to a healthy life.

Acknowledgements

This article was translated with the support of artificial intelligence tools.

Footnotes

Authorship Contributions

Concept: P.D.Ç., Z.N.T., H.Y.Ö., Y.Y., O.K., A.Y., H.C.Ç., Design: P.D.Ç., Z.N.T., H.Y.Ö., Y.Y., O.K., A.Y., H.C.Ç., Literature Search: H.C.Ç., P.D.Ç., Z.N.T., H.Y.Ö., Writing: P.D.Ç., Z.N.T., H.Y.Ö., Y.Y., O.K., A.Y., H.C.Ç., Ç.Ç., N.K., E.D., N.Ö., Z.B., A.G.K., D.E., E.Y., M.B.B., B.Ç., H.K., E.B.

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. Jayasinghe S. The 12 dimensions of health impacts of war (the 12-D framework): a novel framework to conceptualise impacts of war on social and environmental determinants of health and public health. *BMJ Glob Health*. 2024;9(5):e014749. **[Crossref]**
2. Zohny M, Elgailani OSA, Al-Hamody M. IFMSA Policy Document Peace and Health [Internet]. 2021. Accessed January 22, 2026. **[Crossref]**
3. Rushton S, Devkota B. Choosing not to weaponize healthcare: politics and health service delivery during Nepal's civil war, 1996-2006. *Med Confl Surviv*. 2020;36(3):212-231. **[Crossref]**
4. Türk Tabipleri Birliği Merkez Konseyi. Savaş ve Hekimlik Olmaz; ya Savaş ya Hekimlik, Hekimligin Ön Koşulu "Barış"tır. Toplum ve Hekim. 1994;63(9)7-9. **[Crossref]**
5. UNICEF. Convention on the Rights of the Child. Accessed October 14, 2025. **[Crossref]**
6. Kadir A, Shenoda S, Goldhagen J, Pitterman S; Section On International Child Health. The effects of armed conflict on children. *Pediatrics*. 2018;142(6):e20182586. **[Crossref]**
7. ISSOP. Beirut Declaration on Prohibition of Harm to Children in Armed Conflict [Internet]. Accessed January 26, 2026. **[Crossref]**
8. Sahloul MZ, Monla-Hassan J, Sankari A, et al. War is the enemy of health. Pulmonary, critical care, and sleep medicine in War-Torn Syria. *Ann Am Thorac Soc*. 2016;13(2):147-155. **[Crossref]**
9. Szema A, Mirsaidi N, Patel B, et al. Proposed Iraq/Afghanistan war-lung injury (IAW-LI) clinical practice recommendations: National Academy of Sciences' Institute of Medicine Burn Pits Workshop. *Am J Mens Health*. 2017;11(6):1653-1663. **[Crossref]**
10. Jain N, Prasad S, Czárth ZC, et al. War psychiatry: identifying and managing the neuropsychiatric consequences of armed conflicts. *J Prim Care Community Health*. 2022;13:21501319221106625. **[Crossref]**
11. Fleming WH. Moral injury and the absurd: the suffering of moral paradox. *J Relig Health*. 2021;60(5):3012-3033. **[Crossref]**
12. Williamson V, Murphy D, Phelps A, Forbes D, Greenberg N. Moral injury: the effect on mental health and implications for treatment. *Lancet Psychiatry*. 2021;8(6):453-455. **[Crossref]**
13. Griffin BJ, Purcell N, Burkman K, et al. Moral injury: an integrative review. *J Trauma Stress*. 2019;32(3):350-362. **[Crossref]**
14. IPC. Integrated Food Security Phase Classification. Understanding the IPC Scales [Internet]. Accessed January 22, 2026. **[Crossref]**
15. BBC. Israel and the Palestinians: history of the conflict explained. Accessed November 14, 2025. **[Crossref]**
16. BBC. Famine 'currently playing out' in Gaza, UN-backed experts warn. Accessed November 14, 2025. **[Crossref]**
17. UNICEF. 25 years of children and armed conflict. 2022. Accessed October 9, 2025. **[Crossref]**
18. UNICEF. Barely a drop to drink': children in the Gaza Strip do not access 90 per cent of their normal water use. 2023. Accessed October 9, 2025. **[Crossref]**
19. World Health Organization. Women and newborns bearing the brunt of the conflict in Gaza, UN agencies warn. 2023. Accessed October 9, 2025. **[Crossref]**
20. Aqtam I. A narrative review of mental health and psychosocial impact of the war in Gaza. *East Mediterr Health J*. 2025;31(2):89-96. **[Crossref]**
21. Latif S, Muaileq EAA, Chaabo H, et al. Cases of trauma due to war and violence among children in Gaza. *East Mediterr Health J*. 2025;31(4):210-215. **[Crossref]**
22. Javanbakht A. Gaza's hidden crisis: adults, children, and generations of psychological torment to come. *Eur J Psychotraumatol*. 2024;15(1):2416824. **[Crossref]**
23. İnanç BB. Peace keeps us alive; Child health in Gaza. *Anatol Clin*. 2024;29(Special Issue on Gaza): 33-37. **[Crossref]**
24. ICRC. International Humanitarian Law Databases. Accessed October 9, 2025. **[Crossref]**
25. Support the Guardian. Can hospitals be military targets? What international law says. Accessed October 10, 2025. **[Crossref]**
26. World Health Organization. WHO remains firmly committed to the principles set out in the preamble to the Constitution. Accessed October 10, 2025. **[Crossref]**
27. Galtung J. Violence, peace, and peace research. *Journal of Peace Research*. 1969;6(3):167-191. **[Crossref]**
28. World Health Organization. Health and peace initiative. Accessed October 10, 2025. **[Crossref]**
29. Positive Peace. Our understanding of the term 'peace' has evolved significantly over the last 2,500 years. Accessed October 10, 2025. **[Crossref]**
30. Institute for Economics & Peace. Positive Peace is transformational in that it is a cross-cutting facilitator for progress. Accessed October 10, 2025. **[Crossref]**
31. Positive Peace Report 2024. Accessed October 10, 2025. **[Crossref]**
32. İşbilir AC. An evaluation on the peacebuilding functions of civil society organizations and the peace building capacity of Turkish NGOs. Türkiye Ortadoğu Çalışmaları Dergisi. 2021;8(2)85-126. **[Crossref]**
33. Burkle FM Jr, Erickson T, von Schreeb J, et al. The solidarity and health neutrality of physicians in war & peace. *PLoS Curr*. 2017;9:ecurrents.dis.1a1e352febd595087cbeb83753d93a4c. **[Crossref]**
34. Galtung J. Health as a Bridge for Peace in the Context of Humanitarian Action in Complex Emergency Situations. Prepared for the World Health Organization, Geneva Division of Emergency and Humanitarian Action. 1997. Accessed October 15, 2025. **[Crossref]**
35. World Medical Association. WMA resolution on the inclusion of medical ethics and human rights in the curriculum of medical schools world-wide. Accessed October 15, 2025. **[Crossref]**
36. Caglayan S. How can health be more effective in peace works in Turkey: introducing peace through health. *Med Confl Surviv*. 2024;40(2):182-200. **[Crossref]**
37. İnandı T, Altunöz E. Barış İnşası: Kavramsal Bir Bakış. Sosyal Bilimler ve Sağlık Bülteni. 2025(15):72-78. **[Crossref]**
38. Jayasinghe S. The post-millennium development goals agenda: include 'end to all wars' as a public health goal! *Glob Health Promot*. 2014;21(3):29-32. **[Crossref]**

Thorac Res Pract. 2026;27(1):75-76 **DOI:** 10.4274/ThoracResPract.2025.2025-5-4

Letter to the Editor

The Necessity of Bronchiectasis Registries - The Turkish Registry of Bronchiectasis

Deniz Kızılırmak[1], Miguel Angel Martinez-Garcia[2], Sedat Çiçek[3], Ayşın Şakar Coşkun[1], Oğuz Kılınç[4], Ebru Çakır Edis[5]

[1]Department of Chest Diseases, Manisa Celal Bayar University Faculty of Medicine, Manisa, Türkiye
[2]Department of Chest Diseases, Hospital Universitario y Politécnico La Fe, Valencia, Spain; Centro de Investigación Biomédica en Red (CIBERES), Instituto de Salud Carlos III, Madrid, Spain
[3]Clinic of Chest Diseases, Karaman Training and Research Hospital, Karaman, Türkiye
[4]Department of Chest Diseases, İzmir University of Economics Faculty of Medicine, İzmir, Türkiye
[5]Department of Chest Diseases, Trakya University Faculty of Medicine, Edirne, Türkiye

Cite this article as: Kızılırmak D, Martinez-Garcia MA, Çiçek S, Şakar Coşkun A, Kılınç O, Çakır Edis E. The necessity of bronchiectasis registries - the Turkish registry of bronchiectasis. *Thorac Res Pract.* 2026;27(1):75-762026;27(1):75-76

KEYWORDS

Bronchiectasis, bronchiectasis database, national registry

Received: 28.05.2025 ***Revision Requested:*** 15.11.2025 ***Last Revision Received:*** 22.11.2025

Accepted: 15.12.2025 ***Epub:*** 20.01.2026 ***Publication Date:*** 30.01.2026

Dear Editor,

Bronchiectasis demonstrates a significant discrepancy between its high prevalence and the limited evidence base. Patient registries, both national and international, are key to collecting comprehensive data. However, the cost-effectiveness of this approach remains a subject of ongoing debate.[1]

In the opinion of the authors of this editorial, despite the significant costs and effort involved in establishing such registries, they are justified—provided that certain key characteristics are met. The registry must be sufficiently comprehensive in capturing relevant variables to avoid missing critical data that could impede meaningful research, particularly in areas such as therapeutic interventions, comorbidities, microbiological findings, and prognostic factors.

Given the geographic diversity of bronchiectasis, registries should reflect regional or national characteristics. Currently, 15 bronchiectasis registries are active worldwide. The largest registries are the European (EMBARC, ~20,000 patients), the Chinese (~16,000 patients), and the U.S. (~8,000 patients). Several countries, including Spain, India, Germany, Italy, Australia, South Korea, Japan, Türkiye, and the U.K. (closed in 2023), have data on over 1,000 patients. Others, such as Argentina, New Zealand, Canada, and Brazil, are still in the early stages.[2] In Spain, an earlier registry (2002–2011) included 2,123 patients from 36 centers,[3] while the ongoing RIBRON registry (since 2015) has data on 2,631 patients from 42 centers with over 7 years of follow-up.[4]

A significant advancement in this field in Türkiye has been the development of the "Turkish Adult Bronchiectasis Database" (TEBVEB), established with support from the Turkish Thoracic Society (Figure 1).[5] TEBVEB is a web-based registry designed to collect detailed demographic, clinical, and microbiological data on Turkish patients with bronchiectasis. Between March 2019 and January 2022, 1,035 adult patients with bronchiectasis were enrolled at 25 centers.

This study is the first multicenter prospective investigation of the sociodemographic and clinical characteristics of patients with bronchiectasis in Türkiye. The findings suggest that patients in Türkiye are diagnosed at a younger age compared to

Corresponding author: Deniz Kızılırmak, MD, e-mail: dr_dkizilirmak@yahoo.com

Figure 1. Provinces where the study centers are located[5]

their European counterparts, and post-infectious bronchiectasis remains common. Additionally, comorbid COPD emerges as a major determinant of mortality. The lower-than-expected prevalence of tuberculosis indicates the effectiveness of infection control programs in the country.

The Turkish Bronchiectasis Database provides a robust foundation for national research on bronchiectasis in Türkiye and has contributed significantly to the development of national health policies.

National registries can inform health policy, particularly regarding infection prevention and the appropriate use of antibiotics and corticosteroids. Developing specialized protocols is also key to ensuring optimal patient management.

Acknowledgements

We extend our sincere gratitude to the researchers who have contributed to the advancement of bronchiectasis research and the development of national and international bronchiectasis registries.

Footnotes

Authorship Contributions

Concept: D.K., M.A.M-G., A.Ş.C., O.K., E.Ç.E., Design: D.K., M.A.M-G., S.Ç., A.Ş.C., O.K., Literature Search: D.K., M.A.M-G., E.Ç.E., Writing: D.K., M.A.M-G., S.Ç.

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. Gliklich RE, Leavy MB, Dreyer NA, eds. Registries for evaluating patient outcomes: a user's guide [Internet]. 4th ed. Rockville (MD): Agency for Healthcare Research and Quality (US); 2020 Sep. Chapter 1, Patient registries. **[Crossref]**
2. Oscullo G, Bekki A, de la Rosa D, Martinez-García MA. Registries for bronchiectasis in the world: an opportunity for international collaboration. *Int J Tuberc Lung Dis*. 2025;29(5):199-201. **[Crossref]**
3. Martínez-García MÁ, Oscullo G, Gómez-Olivas JD, et al. Bronchiectasis: changes in the characterization of patients during 20 years of follow-up. Data from the Spanish Bronchiectasis Registries. *Arch Bronconeumol*. 2023;59(10):688-690. English, Spanish. **[Crossref]**
4. de la Rosa-Carrillo D, Máiz-Carro L, Martínez-García MÁ. What Have we learned about bronchiectasis from RIBRON (Spanish Bronchiectasis Registry)? *Arch Bronconeumol*. 2023 Oct;59(10):625-626. English, Spanish. **[Crossref]**
5. Çakır Edis E, Çilli A, Kızılırmak D, et al. Bronchiectasis in Türkiye: data from a multicenter registry (Turkish Adult Bronchiectasis Database). *Balkan Med J*. 2024;41(3):206-212. **[Crossref]**