

Semi-Automated Quantitative CT-Enhanced GAP Index for Prognostic Assessment in Idiopathic Pulmonary Fibrosis

Emine Bensus Erdemir

Istanbul University

Gorkem Durak

Northwestern University

Sule Gumurdu

Istanbul University

Deniz Saliha Sobayapan

Istanbul University

Muhammed Enes Tasci

Northwestern University

Alec Peltekian

Northwestern University

Sukru Mehmet Erturk

Istanbul University

Ulas Bagci

Northwestern University

Gülfer Okumus

Istanbul University

Zuleyha Bingol

kayazuleyha@yahoo.com

Istanbul University

Article

Keywords: GAP index, idiopathic pulmonary fibrosis, quantitative CT, artificial intelligence

Posted Date: September 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10831871/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Idiopathic pulmonary fibrosis (IPF) is marked by spatially heterogeneous fibrotic remodeling and highly variable clinical trajectories, which limit the precision of existing prognostic tools. Widely used indices, such as GAP (Gender–Age–Physiology), rely on physiological measurements and do not explicitly capture anatomical disease burden. Advances in artificial intelligence (AI)–assisted 3D quantitative Computed Tomography (CT) imaging enable objective, volumetric characterization of lung fibrosis and provide a foundation for radiology-integrated prognostic modeling.

Methods

We developed and evaluated an AI-assisted quantitative CT framework for prognostic assessment in patients with IPF. Whole-lung segmentation was performed using a pretrained deep learning model (LungMask), followed by semi-automated segmentation of fibrotic abnormalities with radiologist-guided refinement. Quantitative imaging biomarkers, including honeycombing volume, non-honeycomb fibrotic lung volume, preserved lung volume, and the pathological lung volume-to-total lung volume ratio, were derived from 3D CT analysis. This radiological parameter was integrated into the conventional GAP index to create a radiology-enhanced prognostic model. Prognostic performance was assessed using receiver operating characteristic (ROC) analysis, and Kaplan–Meier survival analysis.

Results

A total of 116 patients with IPF were included in the study (mean age: 64.2 ± 9.2 years; 69.8% male). During a median follow-up of 40.5 months, 44% of patients died. Patients who died had significantly lower baseline pulmonary function parameters and greater fibrotic lung involvement on AI-assisted quantitative CT analysis than survivors (all $p < 0.05$). The pathological lung volume-to-total lung volume ratio showed good prognostic performance for predicting mortality (AUC = 0.795, 95% CI: 0.713–0.877, $p < 0.001$). Integrating this radiological parameter into the conventional GAP index further improved prognostic performance (AUC = 0.805, 95% CI: 0.725–0.885, $p < 0.001$) compared with the GAP index alone (AUC = 0.712, 95% CI: 0.618–0.807). Kaplan–Meier analysis showed significant survival differences among modified GAP risk groups, with median survival of 100 months for patients with a score ≤ 6 and 48 months for those with a score ≥ 7 (log-rank $p < 0.001$).

Conclusions

This study presents an interpretable, AI-assisted quantitative CT imaging framework that integrates objective measures of fibrotic lung burden into an established clinical prognostic index. By unifying

volumetric CT biomarkers with physiological scoring, the proposed model improves prognostic precision while maintaining clinical usability. This work highlights the potential of AI-assisted quantitative imaging as a scalable, translational tool for personalized risk stratification in fibrotic lung disease.

Key points

- This study demonstrates that the pathological lung volume-to-total lung volume ratio derived from AI-assisted Quantitative Computed Tomography (QCT) is a strong predictor of mortality in IPF.
- Incorporating this radiological parameter into the GAP index significantly improves prognostic accuracy compared to the GAP index alone.
- These findings highlight that integrating radiological fibrotic burden with clinical indices enables more precise risk stratification and supports personalized management in IPF.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive, and ultimately fatal interstitial lung disease characterized by irreversible fibrotic remodeling of the lung parenchyma, with a median survival of approximately 2–5 years despite antifibrotic therapy (1, 2). The global incidence of IPF is estimated at 3–9 cases per 100,000 persons per year, while prevalence exceeds 20 per 100,000 in aging populations, reflecting a substantial and growing clinical and economic burden (3, 4). IPF predominantly affects older adults and is associated with frequent hospitalizations, increased healthcare utilization, and high annual per-patient healthcare costs compared with many other chronic respiratory diseases. Annual hospitalization rates are approximately 30–50%, and acute exacerbations are associated with mortality rates exceeding 50% within one year (5, 6). These clinical outcomes highlight the need for more accurate prognostic assessment tools to support individualized patient management.

The clinical course of IPF is highly heterogeneous, ranging from a relatively gradual decline to rapid progression, acute exacerbation, and death. As the disease advances, progressive physiological impairment leads to reduced exercise tolerance, limitations in daily activities, and marked deterioration in health-related quality of life (7). Given this variability and its substantial clinical consequences, accurate monitoring and risk stratification are essential for timely management and treatment planning. In routine follow-up, clinically meaningful disease progression is commonly defined by declines in pulmonary function parameters, particularly a $\geq 5\%$ decline in functional vital capacity (FVC) or a $\geq 10\%$ decline in diffusing capacity for carbon monoxide (DLCO) during follow-up, both of which are associated with worse outcomes in IPF (8, 13). Beyond these individual physiological measures, composite indices such as the Gender–Age–Physiology (GAP) score and Composite Physiologic Index (CPI) are widely used for staging and survival prediction in IPF (9, 10). However, physiological parameters and composite clinical indices primarily reflect functional impairment and do not directly quantify the underlying structural fibrotic burden.

To address this structural dimension of disease, CT imaging provides complementary structural information by depicting the extent and distribution of fibrotic lung involvement. CT enables assessment of fibrotic burden and longitudinal radiologic progression through features characteristic of the usual interstitial pneumonia (UIP) pattern, including reticulation, traction bronchiectasis, and honeycombing (11, 12). However, conventional visual or semi-quantitative CT assessment remains limited by interobserver variability and reduced sensitivity for detecting subtle or early disease progression, particularly over short follow-up intervals.

Quantitative computed tomography (QCT), particularly when combined with artificial intelligence (AI), enables objective, reproducible volumetric characterization of fibrotic and preserved lung tissue. Deep learning–based approaches enable voxel-level segmentation of complex radiologic patterns and have demonstrated improved sensitivity for detecting microstructural progression and predicting clinical outcomes (13, 14). However, the integration of AI-assisted imaging biomarkers with established clinical indices and their validation within clinically applicable prognostic models remains limited.

Therefore, this study aimed to develop and evaluate a composite prognostic model integrating AI-assisted QCT biomarkers with the GAP index to improve mortality risk prediction and risk stratification in patients with IPF.

Materials and Methods

Study Design and Patient Population

This is a retrospective, single-center observational study conducted in the Department of Pulmonology at the Istanbul Faculty of Medicine. Patients diagnosed with IPF between 2012 and 2024 were screened. IPF diagnosis was established by multidisciplinary discussion in accordance with current ATS/ERS/JRS/ALAT guidelines (1, 8).

Patients with baseline tests available at the time of diagnosis, such as pulmonary function tests (FVC and DLCO), were included in our study. Additionally, eligible patients were required to have baseline chest CT scans of sufficient quality for quantitative analysis and complete clinical data necessary to calculate established prognostic indices, such as the GAP score and the CPI. Only patients with at least 3 months of follow-up were considered for inclusion to ensure adequate longitudinal assessment.

Patients were excluded if they had chest CT images of insufficient quality for artificial intelligence–based quantitative analysis, or inadequate follow-up data that precluded reliable evaluation of disease progression and outcomes.

Initially, 256 patients with IPF were identified. Of these, 87 were excluded for missing baseline chest CT images, 39 for chest CT data that were technically incompatible with the AI-assisted quantitative analysis software, and 14 for unavailable pulmonary function test data at diagnosis. After applying these exclusion criteria, a final cohort of 116 patients was included in the analysis (Fig. 1).

Clinical and Physiological Assessment

Baseline demographic and clinical data were obtained from medical records. Spirometry was performed by experienced technicians and interpreted according to the ATS/ERS technical standards (2019 update) and the Global Lung Function Initiative (GLI-2012) reference equations (15,16). Spirometric measurements (forced expiratory volume in 1 second (FEV1), FVC, FEV1/FVC) and DLCO were recorded as percentages of predicted values. Spirometric tests were performed using a Vyntus Spiro PC (Vyaire Medical, Carefusion, Hoechberg, Germany) and interpreted in accordance with current guidelines (15). The 6-minute walk test (6MWT) was administered by a physiotherapist in a 50-meter corridor, following established protocols (17). At the beginning and end of the test, patients' fingertip oxygen saturation, pulse rate, blood pressure, dyspnea, and fatigue scores (Borg scale), and total walking distance were recorded. GAP score and the CPI were calculated based on the patients' baseline data at the time of diagnosis.

CT Scan Acquisition

All chest CT scans were performed in the supine position during inspiration. Scans were obtained with a 64-detector Aquilion scanner (Canon, Otawara, Japan). The scanning protocol included the following parameters: a tube voltage of 120 kV; a tube current-time product of 50–150 mAs; a pitch of 0.85–1.4; and a slice thickness and reconstruction interval of 1 mm. Intravenous contrast media were not administered.

AI-Assisted QCT Analysis

Baseline CT scans were analyzed with an AI-assisted 3D QCT framework designed to assess fibrotic lung disease. All CT images were acquired at full inspiration and resampled to isotropic voxel spacing to ensure spatial consistency across scans. Intensity normalization was applied to reduce inter-scan variability.

Lung Segmentation

Initial whole-lung segmentation was performed using a pretrained U-Net-based LungMask model (18). This model was selected for its robust delineation of lung boundaries, even when normal anatomy is distorted in advanced fibrotic lung disease. The resulting lung masks were visually inspected by a thoracic radiologist with 8 years of experience, and manual corrections were made as needed to remove extrapulmonary structures and correct boundary inaccuracies.

Fibrotic Pattern Segmentation

Within the segmented lung volumes, fibrotic abnormalities were identified using a semi-automated, threshold-based segmentation approach. After generating whole-lung masks, predefined Hounsfield unit (HU) ranges were applied within the lung parenchyma to produce a preliminary mask of higher-attenuation regions suspicious for non-honeycomb fibrotic abnormalities, including reticulation and traction bronchiectasis. Based on previously described QCT approaches for interstitial lung abnormalities and high-attenuation areas, voxels between -600 and -250 HU within the lung mask were assigned to the preliminary non-honeycomb fibrotic abnormality mask (19). Voxels outside the lung mask and in extrapulmonary structures were excluded. This algorithm-derived preliminary mask was visually reviewed by a thoracic radiologist with 8 years of experience, who refined the segmentation as needed to correct misclassifications, remove artifacts, and improve anatomical and pathological plausibility.

Honeycombing was manually segmented by the same thoracic radiologist using established morphologic CT features rather than attenuation thresholds alone. This approach was used because honeycombing cannot be reliably distinguished from surrounding fibrotic abnormalities using HU thresholds alone. The semi-automatically generated non-honeycomb fibrosis mask and the manually segmented honeycombing mask were then combined to produce the final pathological lung mask.

To ensure reliability, the final organ and pathology masks underwent a two-step quality-control process and were reviewed by a senior radiologist with 25 years of experience in thoracic radiology. This validation strategy balanced computational efficiency with clinical accuracy, minimizing errors common in fully automated pipelines (Fig. 3). The final segmentation maps were used for downstream volumetric analysis (Fig. 2).

Volumetric Quantification

Final voxel maps were used to compute absolute volumes (in milliliters) of honeycombing, non-honeycomb fibrotic lung, preserved lung parenchyma, and total lung volume. Volumetric measurements were automatically derived from three-dimensional reconstructions, enabling objective, reproducible quantification of fibrotic burden.

Pathological lung volume was defined as the sum of honeycombing and non-honeycomb fibrotic regions, including traction bronchiectasis and reticulation, as quantified by AI-assisted QCT segmentation. Preserved lung parenchyma was excluded from this definition. For each patient, the ratio of pathological lung volume to total lung volume was calculated as a normalized imaging biomarker.

Workflow Feasibility Assessment

To assess the practical feasibility of the semi-automated segmentation workflow, CT data from five randomly selected representative patients were reviewed by two independent expert radiologists with more than 4 years of thoracic radiology experience, using both a fully manual approach in 3D Slicer (20)

and the AI-assisted semi-automated method. For each method, the time required per patient for segmentation, correction, and final evaluation was recorded and compared descriptively.

Development of the Modified GAP Score

The optimal cutoff for mortality prediction was determined using receiver operating characteristic (ROC) curve analysis. Patients with values above this threshold received 1 additional point, which was incorporated into the conventional GAP index to generate a radiology-enhanced modified GAP score. Based on the total score, patients were stratified into two risk groups (≤ 6 vs ≥ 7) for survival analysis. The prognostic performance of the modified GAP score was then compared with that of the conventional GAP index using the ROC curve and Kaplan–Meier survival analyses.

Outcome Measures

The primary outcome was IPF-related mortality during follow-up. Survival time was calculated from the IPF diagnosis date to death or the last follow-up visit. Secondary outcomes included disease progression, defined as a $\geq 5\%$ decline in FVC or a $\geq 10\%$ decline in DLCO (8).

Statistical Analysis

Continuous variables were reported as mean $\pm$ standard deviation or as median (interquartile range), as appropriate. Categorical variables were reported as frequencies and percentages. Group comparisons were performed using Student's t test or Mann–Whitney U test for continuous variables and chi-square or Fisher's exact test for categorical variables.

Survival analysis was conducted using the Kaplan–Meier method, and between-group differences were assessed with the log-rank test. Prognostic performance was evaluated using ROC curve analysis, and areas under the curve (AUCs) were compared. Statistical analyses were conducted in SPSS version 27.0, and p-values < 0.05 were considered statistically significant.

Results

A total of 116 patients with IPF were included in the study. The mean age at diagnosis was 64.2 ± 9.2 years, and 81 patients (69.8%) were male. The median follow-up duration was 40.5 months (range: 3–177 months). Most patients had at least one comorbidity (77.6%), with hypertension being the most common. The majority of patients (84.5%, $N=98/116$) were receiving antifibrotic therapy, including pirfenidone (59.2%) and nintedanib (40.8%). Baseline demographic and clinical characteristics are summarized in Table 1.

During follow-up, 51 patients (44%) died. Patients in the mortality group were significantly older than survivors (66.1 ± 8.9 vs. 62.7 ± 9.2 years, $p=0.027$). Sex distribution, smoking status, and smoking exposure were similar between the groups. Hyperlipidemia and atrial fibrillation were more common in the group with mortality ($p=0.017$ and $p=0.027$, respectively). Mean baseline FVC was $70.2 \pm 19.1\%$

predicted in the mortality group versus $80.8 \pm 19.7\%$ in survivors ($p=0.004$), and mean baseline DLCO was $48.8 \pm 19.0\%$ versus $60.2 \pm 17.2\%$ ($p=0.001$). Similarly, baseline FEV₁ was significantly lower in patients who died ($75.6 \pm 20.4\%$ vs. $84.9 \pm 19.4\%$, $p=0.015$). The mean GAP score of the overall cohort was 3.26 ± 1.75 , and the mean CPI score was 42.3 ± 14.8 . Baseline GAP score at the time of diagnosis was higher in the non-survivor group than the survivors (3.98 ± 1.66 , 2.69 ± 1.61 $p<0.001$). Similarly, baseline CPI at the time of diagnosis was higher in the non-survivor group than the survivors (47.8 ± 14.0 , 38.1 ± 14.2 $p<0.001$).

During the first year of follow-up, an FVC decline of $\geq 10\%$ was more common among patients who died than among survivors (42.9% vs. 14.5% , $p=0.002$). Likewise, an FVC decline of $\geq 5\%$ occurred in 61.5% of the mortality group and 25.4% of survivors ($p<0.001$). A DLCO reduction of $\geq 10\%$ was also more common among patients who died (32.4% vs. 14.5% , $p=0.035$). Detailed pulmonary function test results are summarized in Table 2.

AI-assisted 3D QCT analysis revealed substantial inter-patient variability in lung parenchymal composition. The median honeycombing volume was 13.7 mL (mean: 84.1 ± 174.4 mL), whereas the median non-honeycomb fibrotic lung volume was 703.3 mL (mean: 797.4 ± 544.8 mL). The median preserved lung volume was 2638.7 mL (mean: 2848.7 ± 1478.7 mL). QCT-derived volumetric parameters are summarized in Table 3, and representative examples of CT segmentation and 3D reconstructions are shown in Figure 2.

During follow-up, patients who died had significantly higher honeycombing volume (134.3 ml $\pm$ 236.9 ml vs 44.8 ± 85.3), higher non-honeycomb fibrotic lung volume (1028 ± 555 vs 617 ± 465), and a higher pathological lung volume to total lung volume ratio (0.34 ± 0.20 vs 0.17 ± 0.14), along with significantly lower preserved lung volume (2110 ± 1087 vs 3429 ± 1493) and preserved lung ratio (0.62 ± 0.22 vs 0.82 ± 0.15) than survivors (all $p<0.05$). Clinical, functional, and quantitative CT parameters by mortality status are shown in Table 4.

The conventional GAP index showed moderate discriminative performance in predicting mortality (AUC = 0.712 , 95% CI: 0.618 – 0.807 , $p < 0.001$). At a cutoff of 3.5 , the GAP score had a sensitivity of 66.7% and a specificity of 67.7% for mortality prediction. Detailed diagnostic performance metrics are summarized in Table 5.

ROC analysis showed that the pathological lung volume-to-total lung volume ratio had good discriminative performance for predicting mortality (AUC = 0.795 , 95% CI: 0.713 – 0.877 , $p < 0.001$). Incorporating this radiological parameter into the conventional GAP index further improved prognostic performance (AUC = 0.805 , 95% CI: 0.725 – 0.885 , $p < 0.001$). Detailed ROC results are summarized in Table 5, and ROC curves comparing prognostic models are shown in Figure 4.

Cutoff-based diagnostic performance of the pathological lung volume-to-total lung volume ratio across thresholds is shown in Table 6. The pathological lung volume-to-total lung volume ratio showed high specificity at higher cutoffs, whereas lower thresholds yielded greater sensitivity for predicting mortality.

Integrating this radiological parameter into the conventional GAP index improved prognostic performance compared with the GAP index alone (AUC = 0.805 vs. 0.712, $p < 0.001$).

Kaplan–Meier survival analysis showed clear separation among risk groups defined by the radiology-enhanced modified GAP score. Patients with a modified GAP score ≤ 6 had a median survival of 100 months (95% CI: 74.0–126.0), whereas those with a score ≥ 7 had a significantly shorter median survival of 48 months (95% CI: 32.1–63.9) (log-rank $p < 0.001$). Survival curves are shown in Figure 5, and detailed results are summarized in Table 7.

Our exploratory workflow feasibility assessment for five representative patients, conducted with a fully manual evaluation, required approximately 110–120 minutes per patient, whereas the AI-assisted semi-automated workflow required approximately 15–20 minutes per patient.

Discussion

In this retrospective cohort study of patients with IPF, we evaluated whether AI-assisted 3D QCT–derived volumetric biomarkers could complement the GAP index in predicting the risk of ILD-related mortality. The pathological lung volume–to–total lung volume ratio was associated with mortality and showed discriminative performance. Incorporating this imaging biomarker into the GAP index yielded a higher AUC than the GAP index alone, suggesting that structural fibrotic burden may provide incremental prognostic information beyond physiology-based staging.

The clinical appeal of the GAP index lies in its simplicity. It uses routinely available demographic and physiological variables to provide a practical staging framework for mortality risk assessment in IPF (9). However, this simplicity is also a limitation, as the GAP index does not directly quantify the extent or distribution of structural fibrotic remodeling on CT. This may partly explain why prior studies have consistently shown substantial variability in survival among patients within the same GAP stage (9, 10). Our modified framework incorporates an interpretable QCT-derived measure of structural disease burden into the GAP index. This approach preserves the clinical familiarity of GAP while allowing the score to incorporate both physiological impairment and CT-defined structural fibrotic burden. In our cohort, patients with higher modified GAP scores had shorter median survival, supporting the potential value of QCT-derived structural disease assessment for refining mortality risk stratification. Such information may help identify patients who warrant closer monitoring, timely consideration of lung transplantation referral, or enrollment in a clinical trial.

Our findings should be interpreted in the context of prior studies evaluating CT-derived biomarkers for prognostication in IPF. Ley et al. evaluated semi-quantitative QCT fibrosis scoring within the GAP framework in a two-cohort study of 348 patients with IPF (21). In that study, CT fibrosis and emphysema scores were visually labeled by radiologists across six lung zones and averaged to generate overall QCT scores. The authors proposed a QCT-GAP model in which DLCO was replaced with the QCT fibrosis score, and it performed comparably to the original GAP model. However, simply adding the QCT fibrosis score to the complete GAP model did not improve performance. In contrast, our study used a radiologist-

supervised, AI-assisted 3D QCT workflow to derive a volumetric measure of fibrotic lung burden rather than relying on visual zonal scoring. We also did not replace a physiological component of GAP; instead, we incorporated the pathological lung volume-to-total lung volume ratio as an additional imaging variable to capture CT-defined structural disease burden alongside conventional demographic and physiological variables.

Other quantitative CT studies further support the prognostic relevance of imaging-derived structural biomarkers in IPF. Studies using CALIPER, a computer-based CT analysis tool that quantifies parenchymal abnormalities and pulmonary vascular structures, provided important evidence that imaging-derived measures are associated with clinically meaningful outcomes. In these studies, parameters such as pulmonary vessel volume, honeycombing, and vessel structures predicted mortality or functional decline and provided prognostic information beyond conventional visual CT assessment (13, 22). Salisbury et al. extended this concept by showing that automated quantification of ground-glass reticular abnormalities was associated with disease progression even after adjustment for GAP stage and smoking status, supporting the complementary value of CT-derived biomarkers within clinical risk frameworks (23). Texture-based quantitative CT studies have also shown that both baseline fibrotic disease burden and longitudinal changes in fibrosis scores are associated with survival or short-term clinical progression, suggesting that computational CT analysis can capture both disease extent and temporal progression (24, 25). Together, these studies indicate that CT-derived structural biomarkers can provide prognostic information that is not fully reflected by pulmonary function parameters alone. Our study builds on this evidence by emphasizing clinical integration through a single, interpretable volumetric biomarker incorporated into the GAP index, rather than by developing a standalone imaging model or by directly comparing multiple CT feature extraction methods.

The AI-assisted semi-automated workflow also has potential practical relevance for clinical translation (27). Manual or visual assessment of the extent of fibrotic disease on CT is time-consuming and difficult to implement in routine practice, particularly when the whole-lung disease burden needs to be evaluated (26, 27). Our workflow does not eliminate expert radiologist input; rather, it combines AI-assisted lung segmentation with threshold-based preliminary fibrosis masks, which radiologists can review and refine. In our exploratory feasibility assessment, this semi-automated approach significantly reduced the average evaluation time compared with a fully manual assessment. This suggests that radiologist-supervised AI assistance improves the practicality of QCT-derived biomarkers in clinical and research settings (28).

Limitations

This single-center retrospective study has several limitations. First, the relatively small sample size and single-center design may limit the generalizability of the findings and introduce selection bias. Another important limitation is that the proposed QCT workflow is semi-automated rather than fully automated, as no automated tools exist for segmenting and characterizing fibrotic lung tissue. Although AI-assisted segmentation substantially reduced the time required for quantitative image analysis, expert radiologist

oversight and manual refinement were still necessary to ensure accurate classification of fibrotic regions and volumetric assessment. Nevertheless, given the limited number of studies investigating AI-assisted quantification of volumetric fibrosis in IPF, the ability to combine detailed clinical data with semi-automated 3D quantitative imaging analysis is a notable strength and contributes to the originality of the present study.

Conclusion

In conclusion, this retrospective cohort study showed that a radiologist-supervised, AI-assisted QCT biomarker of fibrotic lung burden can complement the conventional GAP index in predicting mortality risk in patients with IPF. The proposed concept is strong and clinically intuitive: GAP captures physiology, whereas CT captures structural fibrotic burden. The reported improvement in AUC from GAP alone to modified GAP is promising. The clinical relevance of this approach lies in its simplicity, achieved by incorporating a single, biologically meaningful volumetric imaging marker into an established clinical score, and in its ability to capture structural disease burden that is not fully reflected by pulmonary function parameters or physiology-based indices alone. This multidisciplinary approach may help better identify patients who could benefit from closer monitoring, timely referral for lung transplantation evaluation, or more appropriate selection for clinical trials. Further prospective multicenter studies with external validation are needed to confirm the generalizability of these findings and to determine how radiology-enhanced prognostic models can be integrated into clinical decision-making.

Declarations

Ethics approval and consent to participate

This study was approved by the Clinical Research Ethics Committee of Istanbul Faculty of Medicine, Istanbul University (Decision No: 17; Date: 22 August 2025). An official approval letter was issued on 16 June 2025 (File No: E-29624016-050.99-3364824), confirming that the study was evaluated in the ethics committee meeting held on 30 May 2025 (Meeting No: 11). The official electronically signed decision document is dated 11 September 2025 (No: 3577699).

Given the retrospective design, the committee waived the requirement for informed consent. The study was conducted in accordance with the Declaration of Helsinki.

Authors' contributions

All authors contributed to the study design, data collection, analysis, and manuscript preparation. Prof. Dr. Zuleyha Bingol supervised the project. Dr. Emine Bensus Erdemir performed the QCT analysis, data interpretation, and drafted the first version of the manuscript. All authors critically revised and approved the final manuscript.

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflict of Interest: The authors have no relevant financial or non-financial interests to disclose.

Clinical trial number: not applicable.

Funding: No funding

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

References

1. Raghu, G. et al. Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am. J. Respir Crit. Care Med.* **198** (5), e44–e68 (2018).
2. Raghu, G. et al. An official ATS/ERS/JRS/ALAT clinical practice guideline: treatment of idiopathic pulmonary fibrosis. *Am. J. Respir Crit. Care Med.* **192** (2), e3–e19 (2015).
3. Hutchinson, J., Fogarty, A., Hubbard, R. & McKeever, T. Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *Eur. Respir J.* **46** (3), 795–806 (2015).
4. Lederer, D. J. & Martinez, F. J. Idiopathic pulmonary fibrosis. *N Engl. J. Med.* **378** (19), 1811–1823 (2018).
5. Raimundo, K. et al. Clinical and economic burden of idiopathic pulmonary fibrosis. *Clinicoecon Outcomes Res.* **8**, 483–491 (2016).
6. Collard, H. R. et al. Acute exacerbation of idiopathic pulmonary fibrosis. An international working group report. *Am. J. Respir Crit. Care Med.* **194** (3), 265–275 (2016).
7. Swigris, J. J. et al. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. *Respir Med.* **104** (2), 296–304 (2010).
8. Raghu, G. et al. Idiopathic pulmonary fibrosis (an update) and progressive pulmonary fibrosis in adults: an official ATS/ERS/JRS/ALAT clinical practice guideline. *Am. J. Respir Crit. Care Med.* **205** (9), e18–e47 (2022).
9. Ley, B. et al. A multidimensional index and staging system for idiopathic pulmonary fibrosis. *Ann. Intern. Med.* **156** (10), 684–691 (2012).
10. Wells, A. U. et al. Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *Am. J. Respir Crit. Care Med.* **167** (7), 962–969 (2003).
11. Devaraj, A., Wells, A. U. & Hansell, D. M. Imaging: how to recognise idiopathic pulmonary fibrosis. *Eur. Respir Rev.* **23** (132), 215–219 (2014).
12. Lynch, D. A. et al. Diagnostic criteria for idiopathic pulmonary fibrosis: a Fleischner Society White Paper. *Lancet Respir Med.* **6** (2), 138–153 (2018).
13. du Bois, R. M. et al. Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. *Am. J. Respir Crit. Care Med.* **184** (12), 1382–

1389 (2011).

14. Walsh, S. L. F., Calandriello, L., Silva, M. & Sverzellati, N. Deep learning for classifying fibrotic lung disease on high-resolution computed tomography. *Eur. Respir J.* **52** (5), 1800984 (2018).
15. Humphries, S. M. et al. Quantitative high-resolution computed tomography fibrosis score: performance characteristics in idiopathic pulmonary fibrosis. *Eur. Respir J.* **52** (3), 1801384 (2018).
16. Graham, B. L. et al. 2017 ERS/ATS standards for single-breath carbon monoxide uptake in the lung. *Eur. Respir J.* **49** (1), 1600016 (2017).
17. Crapo, R. O. et al. Standardization of spirometry: 1994 update. *Am. J. Respir Crit. Care Med.* **152** (3), 1107–1136 (1995).
18. 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.* **166** (1), 111–117 (2002).
19. Hofmanninger, J. et al. LungMask: automated lung segmentation for thoracic CT. Available from: <https://github.com/JoHof/lungmask>
20. Podolanczuk, A. J. et al. High attenuation areas on chest computed tomography in community-dwelling adults: the MESA study. *Eur. Respir J.* **48** (5), 1442–1452 (2016).
21. Fedorov, A. et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging.* **30** (9), 1323–1341 (2012).
22. Ley, B. et al. Idiopathic pulmonary fibrosis: CT and risk of death. *Radiology* **273** (2), 570–579 (2014).
23. Jacob, J. et al. Predicting outcomes in idiopathic pulmonary fibrosis using automated computed tomographic analysis. *Am. J. Respir Crit. Care Med.* **198** (6), 767–776 (2018).
24. Salisbury, M. L. et al. Idiopathic pulmonary fibrosis: the association between the adaptive multiple features method and fibrosis outcomes. *Am. J. Respir Crit. Care Med.* **195** (7), 921–929 (2017).
25. Park, H. J., Lee, S. M., Song, J. W., Kim, N. & Seo, J. B. Texture-based automated quantification of regional patterns on CT in idiopathic pulmonary fibrosis: relationship with pulmonary function test results and survival. *Korean J. Radiol.* **20** (5), 737–747 (2019).
26. Kim, G. H. J. et al. Prediction of idiopathic pulmonary fibrosis progression using early quantitative changes on CT imaging for a short term of clinical 18-24-month follow-ups. *Eur. Radiol.* **30** (2), 726–734 (2020).
27. Robbie, H., Daccord, C., Chua, F. & Devaraj, A. Evaluating disease severity in idiopathic pulmonary fibrosis. *Eur. Respir Rev.* **26** (145), 170051 (2017).
28. Zhang, T. et al. What role does artificial intelligence-driven quantitative computed tomography play in the management of idiopathic pulmonary fibrosis? *Quant. Imaging Med. Surg.* **15** (8), 5601–5615 (2025).

Tables

Tables 1 to 7 are available in the Supplementary Files section.

Figures

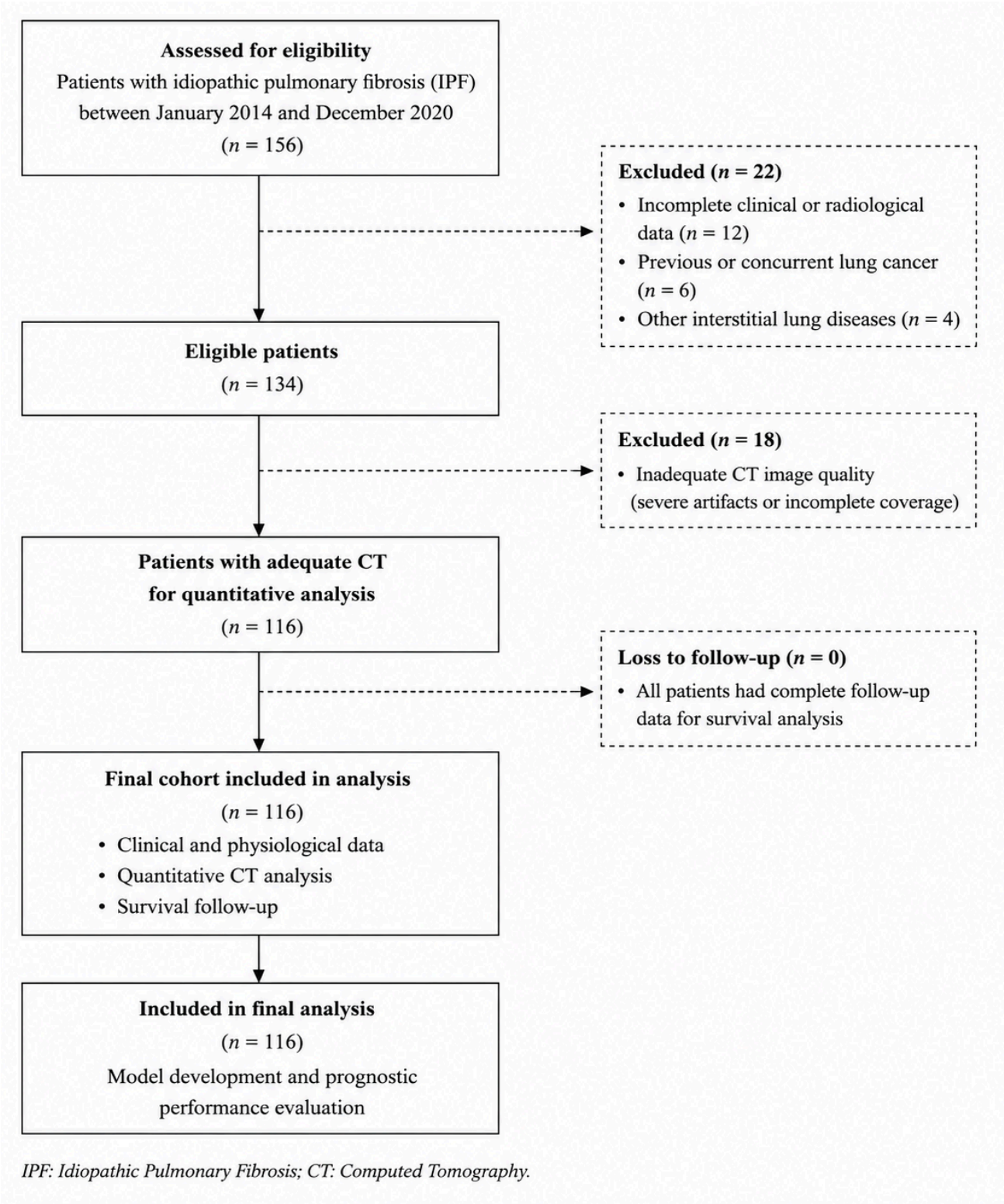


Figure 1

Flow chart of patient selection and study inclusion process.

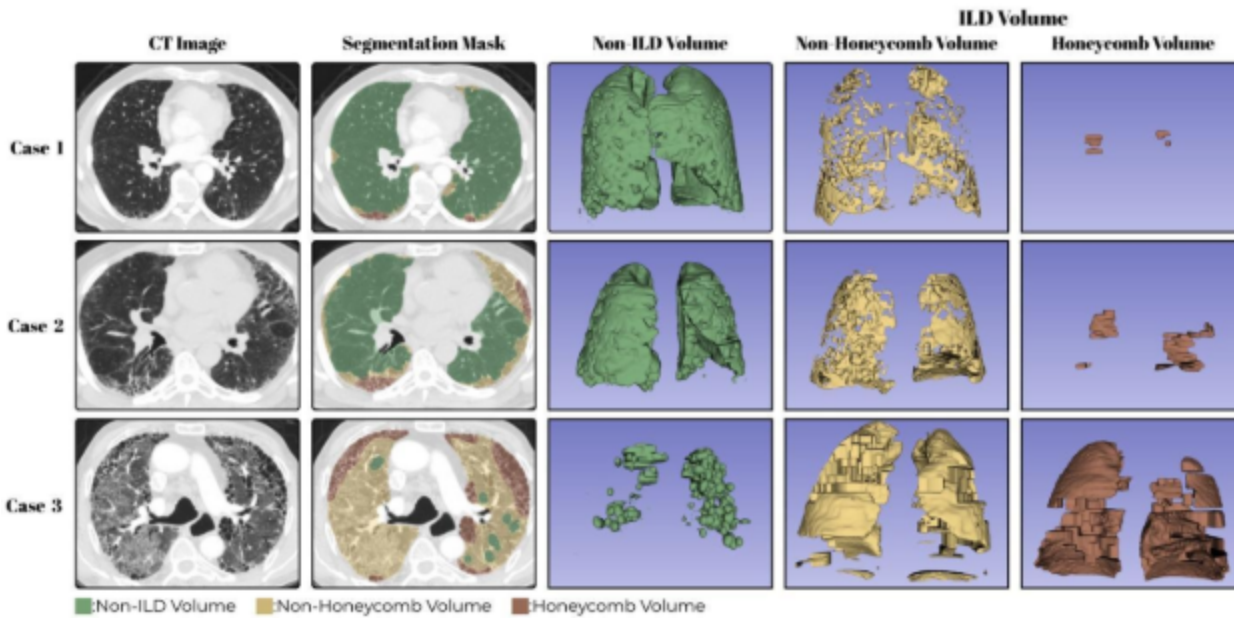


Figure 2

Representative chest Computed Tomography (CT) scans and segmentation outputs from three patients with IPF, showing mild, moderate, and severe honeycombing (green = healthy lung; yellow = non-honeycomb fibrosis; brown = honeycombing).

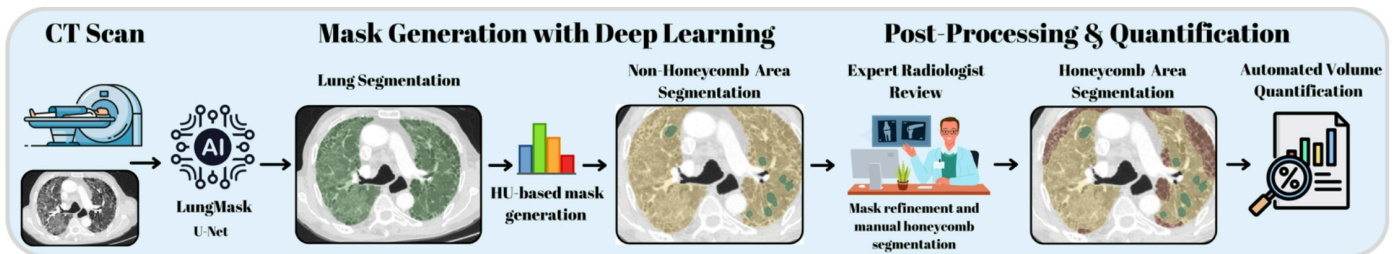


Figure 3

Workflow of AI-assisted Quantitative Computed Tomography (QCT) analysis, including preprocessing, deep learning-based segmentation, radiologist validation, and volumetric quantification of lung regions.

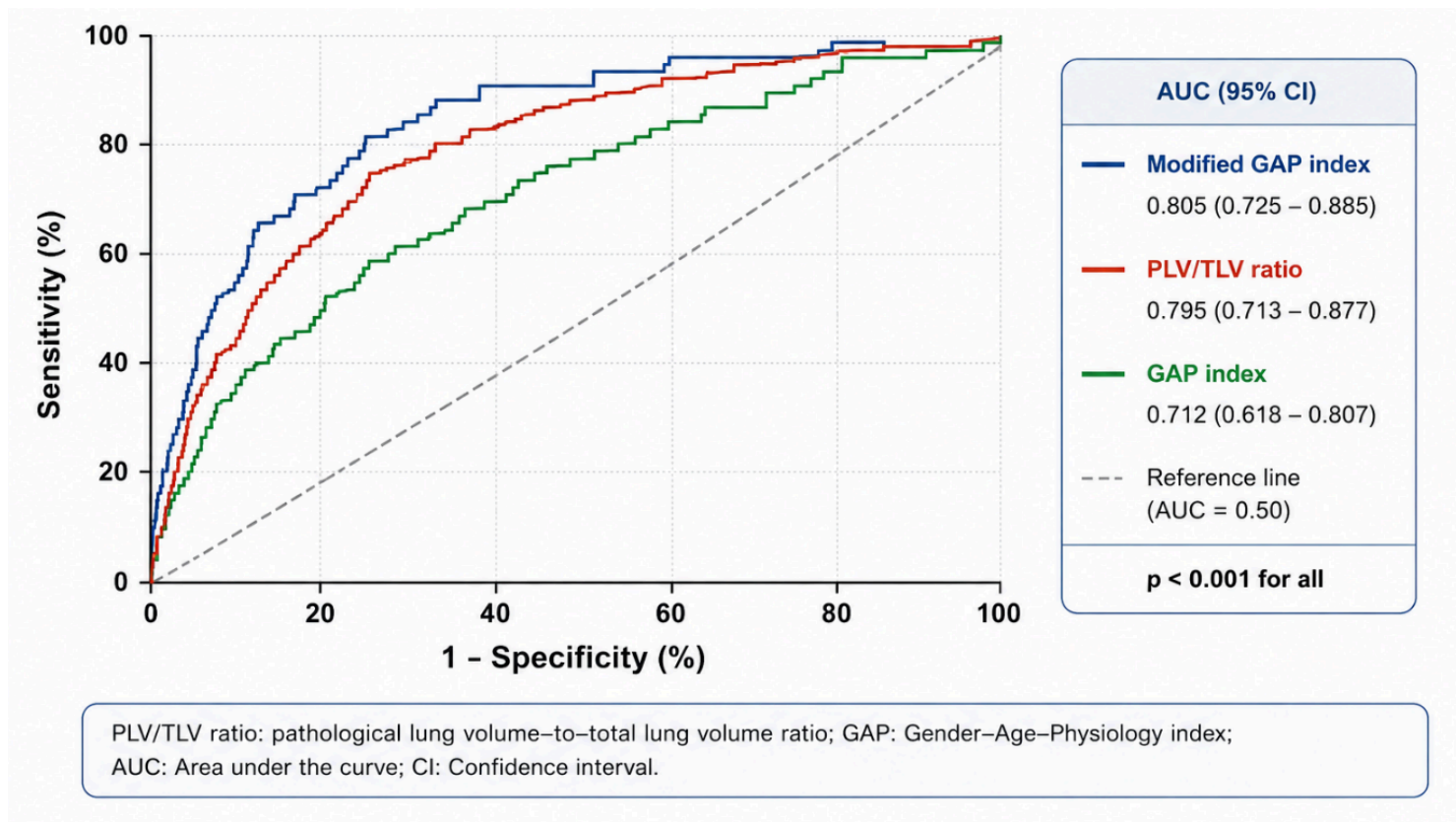


Figure 4

Receiver operating characteristic (ROC) curves comparing the prognostic performance of the pathological lung volume–to–total lung volume ratio and the modified GAP index for mortality prediction in idiopathic pulmonary fibrosis.

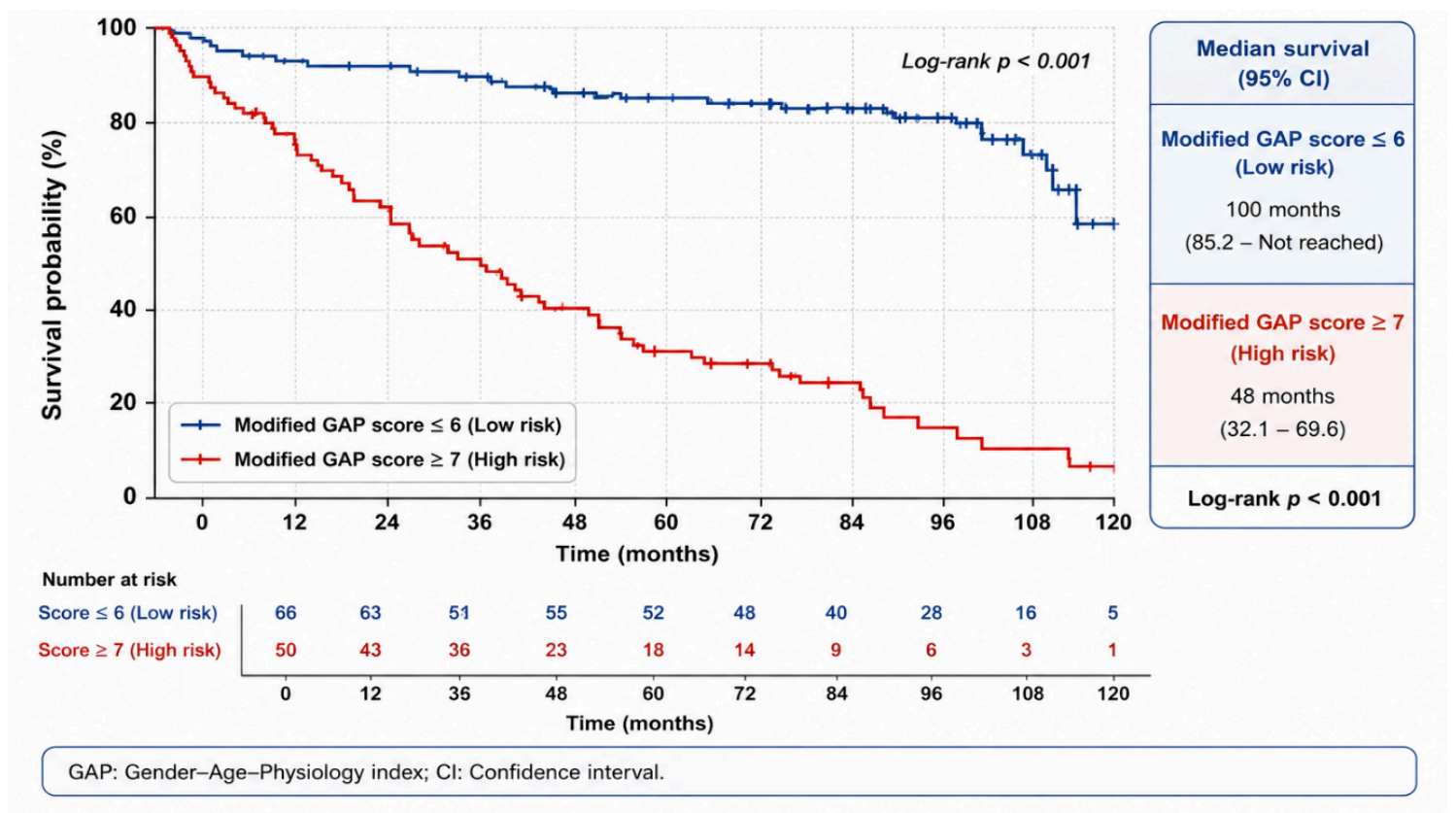


Figure 5

Kaplan–Meier survival curves according to the radiology-enhanced modified GAP score

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Tables.docx](#)