Radiology: Cardiothoracic Imaging

ORIGINAL RESEARCH

Deep Learning Segmentation of Pectoralis Muscle Volume at CT and Comparison with Pectoralis Muscle Area in COPD

Daniel Genkin, BEng[1] • *Mustansir Verdawala*[1] • *Sophie É. Collins, PhD*[2] • *Wan C. Tan, MD*[3] • *Pei Z. Li, MSc*[4,5] • *Jean Bourbeau, MD*[4,5] • *Michael K. Stickland, PhD*[6] • *Dennis Jensen, PhD*[2,4,5,7] • *Miranda Kirby, PhD*[8]

Author affiliations, funding, and conflicts of interest are listed at the end of this article.

Radiology: Cardiothoracic Imaging 2026; 8(2):e250060 • https://doi.org/10.1148/ryct.250060 • Content codes: CA CH AI

Purpose: To develop a deep learning model for segmenting pectoralis muscle volume (PMV) at CT and evaluate the reproducibility, group differences, and associations of pectoralis muscle area (PMA) and PMV with chronic obstructive pulmonary disease (COPD)–related outcomes.

Materials and Methods: This study was a secondary analysis of the prospective Canadian Cohort Obstructive Lung Disease study (CanCOLD, data collected from November 2009 to July 2015). Randomly sampled CT scans from CanCOLD were used for model training, validation, and internal testing (n = 96, 16, and 32, respectively) and an external dataset for external testing (n = 32). A U-Net model was trained for PMV segmentation, and performance was assessed using the Dice similarity coefficient (DSC). PMA and PMV values were extracted from paired inspiration and expiration scans to assess segmentation reproducibility. Differences between individuals with or without COPD and associations with forced expiratory volume in 1 second (FEV_1), diffusing capacity of the lungs for carbon monoxide (DLCO), and peak oxygen uptake during exercise (VO_2) were reported.

Results: Individuals included those with (n = 634; mean age, 67.3 years ± 10.1 [SD]; 394 male participants) and without (n = 601; mean age, 65.8 years ± 9.6; 327 male participants) COPD. The model yielded DSCs of 0.94 ± 0.04, 0.93 ± 0.03, and 0.92 ± 0.04 in the training and validation, internal testing, and external testing datasets, respectively. Contrary to PMV (bias, 0.1 cm^3; P = .77), PMA showed bias between inspiration and expiration (bias, −2.7 cm^2; P < .001). Both PMA and PMV were reduced in patients with COPD (P < .05), but PMV was more strongly associated with FEV_1 (adjusted R^2 [R_{adj}^2], 0.609/0.598), DLCO (R_{adj}^2, 0.645/0.627), and VO_2 (R_{adj}^2, 0.680/0.666).

Conclusion: An accurate and generalizable CT-based deep learning model for pectoralis muscle segmentation was developed. Compared with PMA, PMV showed better reproducibility and stronger associations with COPD outcomes.

ClinicalTrials.gov identifier no. NCT00920348

Supplemental material is available for this article.

Chronic obstructive pulmonary disease (COPD) is a multisystem disease that affects organ systems in addition to the lungs. Individuals with COPD experience chronic low-grade inflammation that can trigger pathways of protein catabolism, leading to muscle atrophy (1). This progressive loss of muscle tissue, defined as *sarcopenia*, occurs in ~22% of patients with COPD (2) and has been recognized as a comorbid predictor of outcomes such as airflow obstruction (3), exacerbations (3), exercise intolerance (4), and mortality (5).

Some standard approaches for quantifying muscle mass include dual-energy absorptiometry and bioelectrical impedance assessments, but neither are typically performed during routine COPD care. However, chest CT scans are routinely performed in those with and at risk for COPD during persistent exacerbations (6), discordant symptomology (6), or lung cancer screening (6). Although chest CT scans are primarily used to evaluate COPD phenotypes such as emphysema or small airways disease, they can also visualize the respiratory musculature and are, therefore, useful in the quantification of sarcopenia. Identifying those with COPD who also have comorbid sarcopenia may have important clinical implications, as pulmonary rehabilitation can directly target muscle loss as a treatable condition and complement existing treatment strategies focused primarily on halting disease progression, and prediction models may be improved, as sarcopenia is an established risk factor for COPD morbidity and mortality (7–11).

A particular muscle of interest is the pectoralis, as it is in the complete field of view during chest CT scans, and its cross-sectional area has been validated against the dual-energy absorptiometry measurements in individuals exposed to tobacco (7). The established methods for extracting pectoralis muscle measurements from chest CT scans involves measurement of the two-dimensional (2D) pectoralis muscle area (PMA) on a single axial section containing the aortic arch (8). PMA measurements have been shown to quantitatively reflect sarcopenia in patients with COPD and are significantly associated with lung function (8–10), exercise capacity (8,11), and quality-of-life questionnaire scores (8,10). However, single-section measurements are susceptible to variability in section selection and anatomic differences, which can result in differences in the measured 2D PMA at different points during analysis. Furthermore, 2D muscle measurements fail to consider the three-dimensional (3D) structure of the muscle and, thus, fail to capture spatially heterogeneous muscle atrophy and adipose tissue infiltration.

Two recent studies leveraged deep learning methods to extract the entire 3D pectoralis muscle volume (PMV) (12,13). However, neither study externally tested their models, which is critical for generalizability and translation, nor did they compare 3D to 2D muscle measures.

Therefore, the objective of this study was to develop a deep learning model to automatically segment the 3D PMV from

Abbreviations

AIC = Akaike information criterion, BMI = body mass index, CanCOLD = Canadian Cohort Obstructive Lung Disease, COPD = chronic obstructive pulmonary disease, DLCO = diffusing capacity of the lungs for carbon monoxide, DSC = Dice similarity coefficient, FEV_1 = forced expiratory volume in 1 second, FVC = forced vital capacity, GOLD = Global Initiative for Obstructive Lung Disease, iNo = inhaled nitric oxide, PMA = pectoralis muscle area, PMV = pectoralis muscle volume, R_{adj}^2 = adjusted R^2, 3D = three dimensional, 2D = two dimensional, VO_2 = peak rate of oxygen consumption during exercise

Summary

A deep learning model was developed to accurately segment pectoralis muscle volume on CT scans, and three-dimensional measurements showed greater reproducibility and stronger associations with chronic obstructive pulmonary disease outcomes than two-dimensional measurements.

Key Points

- A deep learning pectoralis muscle volume (PMV) segmentation model yielded Dice similarity coefficients of 0.94 ± 0.04 (SD), 0.92 ± 0.04, 0.93 ± 0.03, and 0.92 ± 0.04 in the training, validation, internal test, and external test datasets, respectively.
- PMV measurements showed no evidence of bias between inspiratory and expiratory scans, contrary to two-dimensional pectoralis muscle area (PMA) measurements ($P < .001$).
- PMA and PMV measurements differentiated individuals with and without chronic obstructive pulmonary disease (COPD) ($P < .05$), but PMV was more strongly associated with COPD outcomes.

Keywords

CT, Thorax, Lung, Volume Analysis, Chronic Obstructive Pulmonary Disease, Segmentation

chest CT scans and to evaluate the generalizability of the model on an external test set of unseen data. The secondary objectives were to quantitatively compare the PMA and PMV in terms of *(a)* reproducibility between paired inspiration and expiration CT scans, *(b)* differences between individuals with and without COPD, and *(c)* associations with common COPD outcomes, such as lung function, exercise capacity, and quality-of-life questionnaire scores.

Materials and Methods

Study Individuals

All individuals in the population-based Canadian Cohort Obstructive Lung Disease study (CanCOLD; ClinicalTrials.gov identifier no. NCT00920348) (14) were evaluated using data in the baseline visit collected from November 2009 to July 2015. Analysis of data from the CanCOLD study has resulted in more than 50 publications from other research groups. Although CT analysis of 2D muscle measurements in 90 of 1561 participants in the CanCOLD study has previously been described (9), this secondary analysis reports the development of 3D muscle measurements for all individuals in the CanCOLD study, independent of previous work. Informed written consent from all individuals and institutional review board approval were obtained at each participating site.

Briefly, the original CanCOLD study used random digit dialing to contact individuals >40 years of age from nine centers across six Canadian provinces. More than 1500 individuals with forced expiratory volume in 1 second to forced vital capacity ratio (FEV_1:FVC) values less than 0.7 were enrolled and categorized into one of four groups according to the Global Initiative for Obstructive Lung Disease (GOLD) criteria (6), as follows: *(a)* individuals without COPD who never smoked, *(b)* individuals without COPD who had a history of smoking, *(c)* individuals with mild COPD (GOLD 1), and *(d)* individuals with moderate-severe COPD (GOLD 2+). A flowchart detailing the selection criteria for this study is shown in Figure 1. For this secondary analysis, individuals who were missing inspiratory CT scans, spirometry data, or pack-year information or who had a preserved FEV_1:FVC ratio (FEV_1:FVC > 0.7) but an FEV_1 percent of the predicted value less than 80% (preserved ratio but impaired spirometry) were excluded from this study. Individuals who were missing expiratory CT scans were also excluded from the reproducibility study.

An external group of participants with COPD who were recruited to identify phenotypes that could predict a positive response to inhaled nitric oxide (iNO) were retrospectively analyzed as an external test group (ClinicalTrials.gov identifier no. NCT03679312) (15). Data collection for the external test group began in November 2019 and was completed by November 2023. The iNO study was a single-center randomized controlled trial, distinct from the CanCOLD study, involving individuals with mild to severe COPD. Individuals were recruited from community COPD clinics, with inclusion criteria requiring a COPD diagnosis and a smoking history of at least 10 pack-years. Participants with significant cardiovascular or metabolic disease, musculoskeletal impairments, or those taking oral corticosteroids or medications known to affect nitric oxide bioavailability were excluded. Informed written consent from all individuals in the iNO study and institutional review board approval was also obtained.

Pulmonary Function, Cardiopulmonary Exercise Test, and Questionnaire Scores

The pulmonary function, exercise testing results, and questionnaire scores used for this secondary analysis were sourced from the baseline CanCOLD visit. Pulmonary function testing was performed to measure postbronchodilator FEV_1 and FVC in liters according to American Thoracic Society guidelines (16,17). Similarly, the pulmonary diffusing capacity for carbon monoxide (DLCO) in mL/min/mm Hg was measured using standardized protocols (17). A symptom-limited incremental cardiopulmonary cycle exercise test was conducted to measure the peak rate of oxygen consumption (VO_2), as previously described (18). VO_2, a surrogate marker of aerobic capacity, is reported in liters per minute relative to normative values (18). Finally, the COPD assessment test (19), a quality-of-life questionnaire indicating the impact of COPD symptoms on individuals' everyday lives, was administered (scores ranged 0–40, with higher scores indicating greater impact).

CT Image Acquisition

The CT scans used for this secondary analysis were also sourced from the baseline CanCOLD visit. CanCOLD participant chest CT scans were acquired at each site using various CT scanners with individuals in the supine position at full inspi-

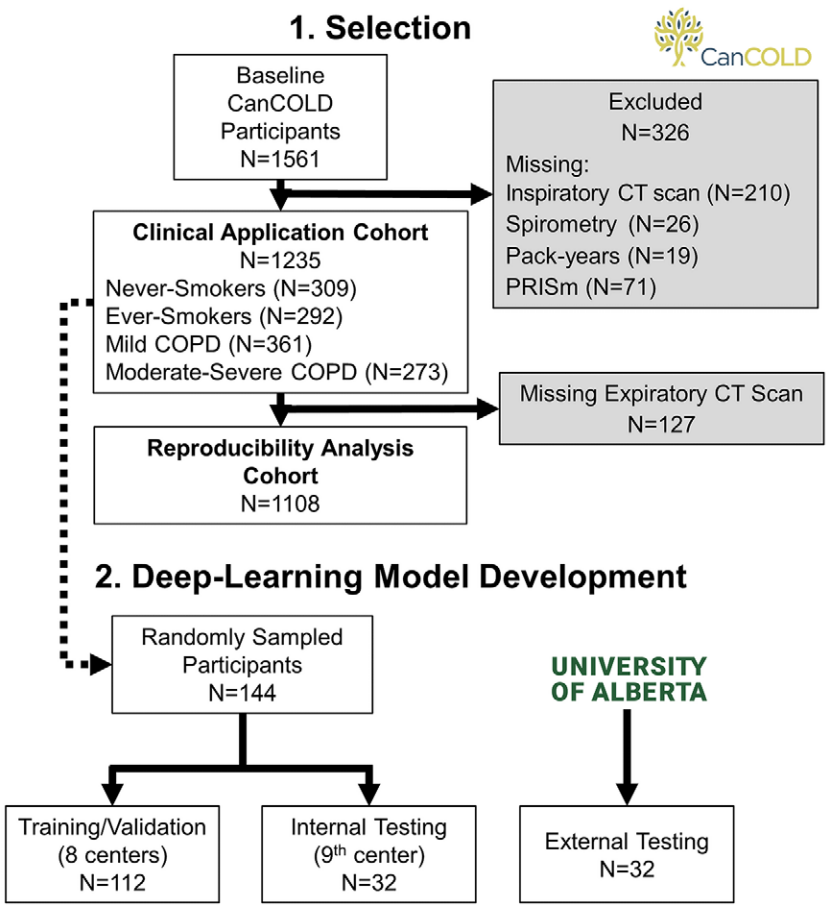

Figure 1: Participant consort flowchart. The baseline Canadian Cohort Obstructive Lung Disease (CanCOLD) study cohort included 1561 individuals, using data collected from November 2009 to July 2015. Individuals for whom inspiratory CT images could not be analyzed for any reason were excluded (n = 210). Individuals with no classification information, namely, no information related to spirometry (n = 26) or smoking status (n = 19), were excluded. Finally, individuals with a preserved forced expiratory volume in 1 second (FEV_1) to forced vital capacity (FVC) ratio (FEV_1:FVC > 0.7) but abnormally low FEV_1 percent of the predicted (FEV_1 %pred < 80%) value, also known as *preserved ratio but impaired spirometry* (PRISm) (n = 71), were also excluded. A total of 1235 individuals were included in the final clinical application analysis, and chronic obstructive pulmonary disease (COPD) status was spirometrically defined according to a postbronchodilator FEV_1:FVC value less than 0.70 (non-COPD: n = 601; COPD: n = 634). The individuals were further stratified into four groups: *(a)* non-COPD with no smoking history, *(b)* non-COPD with a smoking history (at risk), *(c)* mild COPD (Global Initiative for Chronic Obstructive Lung Disease [GOLD] 1: FEV_1 %pred ≥ 80%), and *(d)* moderate-severe COPD (GOLD 2+: FEV_1 %pred < 80%). The subset of individuals with paired expiratory CT images (n = 1108) was included for the reproducibility analysis. From the individuals selected for clinical analysis (n = 1235), 144 were randomly sampled (and sex matched) for model training and validation (n = 112, eight of nine CanCOLD centers) and internal testing (n = 32, ninth center from CanCOLD). Data from the inhaled nitric oxide study (n = 32, ClinicalTrials.gov identifier no. NCT03679312) collected between November 2019 to November 2023 were used to externally test the trained model.

ration and full expiration without the use of contrast agents (14). All scans were collected helically, with each participant's arms placed above their heads, and with a field of view containing the base to apex of both lungs. The imaging parameters during acquisition were 120 kVp, a 100–110-mA tube current, a 0.5-second gantry rotation time, a pitch of 1.25, a 1.0- or 1.25-mm section thickness, a 512 × 512 matrix reconstruction size, and 0.51–0.98-mm^2 pixel spacing. In the iNO study, full-inspiration chest CT scans were also acquired with individuals in the supine position, with arms raised above the head, and the following imaging protocol: 120 kVp, 220-mA tube current, a 0.5-second gantry rotation time, a pitch of 1.00, a 0.75-mm section thickness, a 512 × 512 matrix reconstruction size, and 0.56–0.80-mm^2 pixel spacing. A standard kernel was used for image reconstruction in both groups. A complete list of the acquisition parameters for both the inspiratory and expiratory CT scans stratified by site is shown in Tables S1 and S2.

Reference Standard Creation

All subsequent steps related to image processing and model training were performed in MATLAB R2024a (version 24.1.0; MathWorks). Data from individuals from eight of nine CanCOLD centers were used for deep learning model training and validation, whereas the unseen data from the ninth center were used as an internal test set for the model. Seven batches of 16 individuals each (112 total, balanced for COPD vs no COPD and male vs female individuals, with two individuals for each of the eight CanCOLD centers) were randomly sampled to ensure that the model was trained on a diverse representative dataset, minimizing the risk for bias across demographics, disease status, and geographic regions. In addition, 32 individuals were randomly sampled from the ninth CanCOLD center for internal testing (ensuring the same equal stratification as the training and validation data). Manual segmentation of the PMVs from the first two batches of the training and validation data (n = 32), the internal test data (n = 32), and the external test data from the iNO cohort (n = 32) was first performed by an image analyst (M.V., with 2 years of experience). An active learning approach was then implemented for reference standard segmentation on the remaining batches of the training and validation data. Initial model predictions on new batches of unannotated data were performed using intermediate deep learning models, which were iteratively trained on all available training data and then manually corrected by the image analyst, as described in further detail in Figure 2. This process continued until training and validation performance plateaued and there were at least 100 reference standard segmentations in the training and validation set, similar to previous work (12,13). To assess the reproducibility of the reference standard segmentations produced by the image analyst, the first batch of the training and validation set data (n = 16) was completely manually resegmented 2 weeks later by the same image analyst and another experienced image analyst (D.G., with 5 years of experience). Both image analysts were blinded to all individual information before any segmentations occurred.

Deep Learning Model Training

Once the complete reference standard set was created, the final deep learning model was trained using the training and

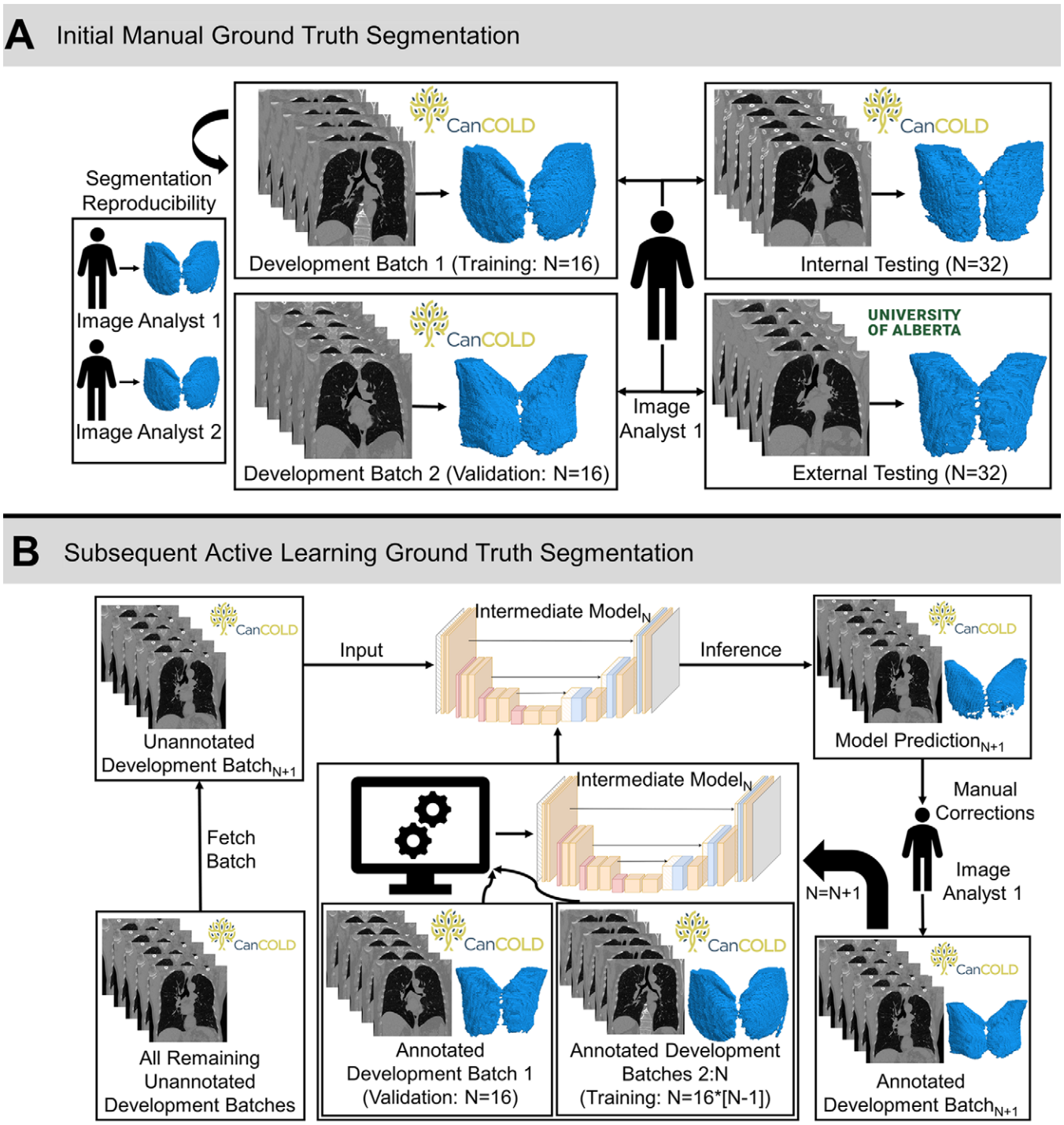

Figure 2: Complete process for creating reference standard pectoralis muscle volume segmentations. **(A)** Complete manual segmentation of pectoralis muscle volumes from the first two batches of training set data (n = 32), the complete internal testing data (n = 32), and all the external testing data from the inhaled nitric oxide cohort (n = 32; ClinicalTrials.gov identifier no. NCT03679312) was performed by an image analyst. The first batch of training set data (n = 16) was completely manually resegmented by the same research analyst and another experienced image analyst to evaluate reference standard segmentation reproducibility. **(B)** An active learning approach was implemented for reference standard segmentation of the five remaining batches of training set data by iteratively training intermediate deep learning models on all available training data as new batches were completed. First, an intermediate deep learning model was trained on the second batch of training set data (n = 16), with early stopping occurring once performance failed to improve for 2 epochs on the held-out first batch of training set data used for validation (n = 16). Preliminary predictions of the pectoralis muscle volumes for the third batch of training set data (n = 16) were then created by the first intermediate deep learning model and subsequently manually corrected by the same image analyst. The second and third batches of training set data were then combined (n = 32) and used to train the second intermediate deep learning model, using the same criteria for early termination on the same validation data (n = 16) followed by manual edits. The active learning process continued until all seven batches of training set data were segmented (n = 112) and the complete model training and validation reference standard set was created.

validation set data to segment the PMV completely automatically. First, the CT volumes and pectoralis segmentations were downsampled to a matrix size of 256 × 256 × 256 using nearest neighbor interpolation. A four-layer 3D U-Net architecture was created (20), and sevenfold cross-validation was implemented, with early stopping once the validation perfor-

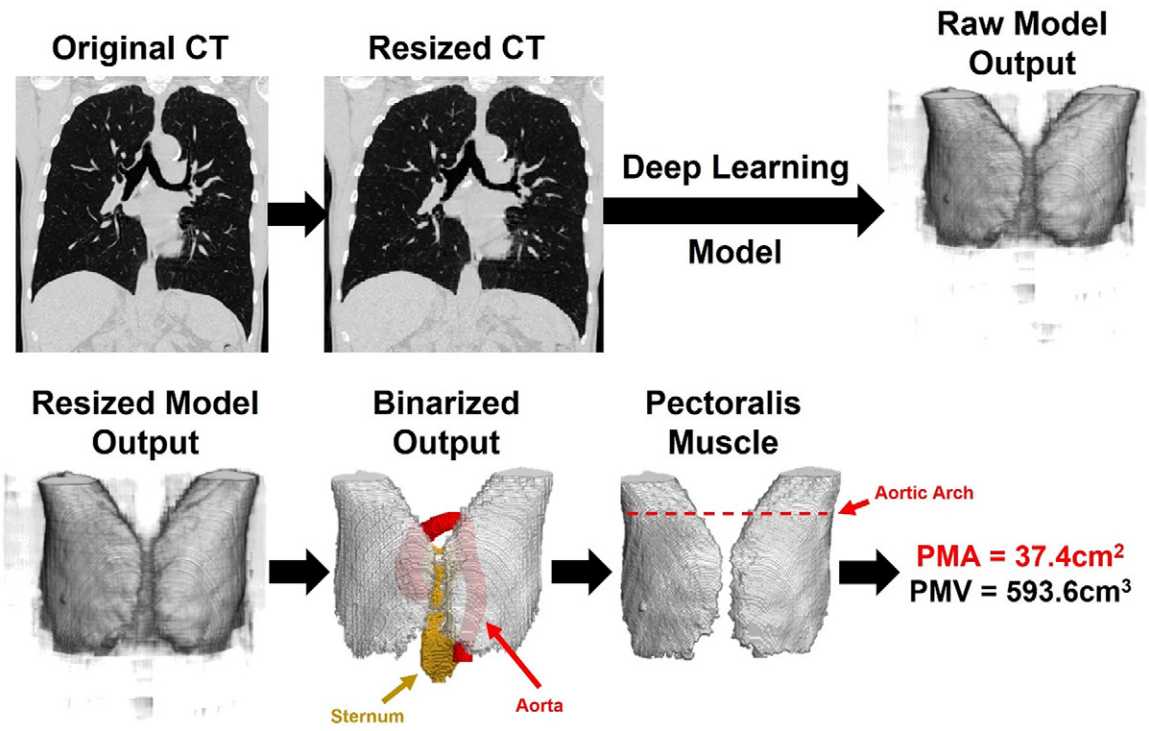

Figure 3: Complete pectoralis muscle extraction pipeline. The original CT image is resized to a matrix size of 256 × 256 × 256 using nearest neighbor interpolation and is fed to the developed deep learning model. The raw model outputs are then resized back to the original 1-mm^3 size using nearest neighbor interpolation and binarized at a threshold of 0.5. The largest connected component in direct contact with the sternum (segmented using TotalSegmentator) is maintained as the pectoralis, and the remaining components are removed. Morphologic operations (1. Open: sphere; radius, 2 mm) (2. Close: sphere, radius, 5 mm) are applied to the binarized volume, and all voxels outside the radiodensity range of −29 to 150 HU are removed. The aortic arch section is automatically determined as the topmost section of the aorta (segmented using TotalSegmentator). Finally, the two-dimensional pectoralis muscle area (PMA) is extracted as the number of voxels in the aortic arch section (in centimeters squared), while the three-dimensional pectoralis muscle volume (PMV) is extracted as the total number of voxels belonging to the pectoralis muscle segmentation (in cubic centimeters).

mance failed to improve for 2 epochs. Performance of the final model was evaluated based on the training and validation data, and the generalizability of the model was evaluated on the data from the ninth CanCOLD center (internal testing) and external iNO group. A complete list of model hyperparameters is available in Table S3. Model training was performed on an NVIDIA RTX 6000 48 GB GPU. The trained model is publicly available upon request at *https://drive.google.com/drive/folders/18y9eOAMVidULfvvJ62ujtPd7JbO7sYwi?usp = sharing*.

Postprocessing and Measurement Extraction

After model training, the pectoralis segmentations were automatically extracted. The complete pectoralis muscle extraction pipeline with postprocessing steps is described in Figure 3. Briefly, model outputs were binarized at a threshold of 0.5, and the largest connected component in contact with the sternum (automatically segmented using TotalSegmentator [21]) was retained as the pectoralis after the removal of voxels outside the standard muscle radiodensity range of –29 to 150 HU (22). The aorta was also automatically segmented using TotalSegmentator, and its area on the topmost axial section was defined as the aortic arch. The PMA was defined as the number of pectoralis pixels in the aortic arch section in centimeters squared, whereas the PMV was defined as the total number of pectoralis voxels in cubic centimeters. Finally, using the complete pipeline, the PMA and PMV were extracted from all eligible baseline inspiratory (n = 1235) and expiratory (n = 1108) CT scans in the CanCOLD dataset for clinical application of the measurements.

Statistical Analysis

Statistical analyses were performed using GraphPad Prism (version 10.2.3; GraphPad Software) and SPSS (version 29.0.2). Differences between the no-COPD and COPD groups as well as the training and validation, internal testing, and external testing groups were evaluated using one-way analysis of variance or the Fisher exact test with Tukey correction for multiple comparisons. The overlap between the automated pectoralis volume segmentation and the reference standard segmentations was quantified using the Dice similarity coefficient (DSC), Jaccard index, absolute error (cm^2/cm^3), and relative error (%) (9).

DSC values were categorized as excellent (>0.90), very good (0.80–0.90), good (0.70–0.79), or poor (<0.70). Measurement reproducibility was strictly evaluated in the subset of individuals with inspiration and expiration measurements using Bland-Altman analysis, the coefficient of variation, the intraclass correlation coefficient (absolute agreement, single rater case), and the Spearman correlation coefficient. Formal comparisons between the PMA and PMV reproducibility metrics were performed using random bootstrapped resampling (n = 500) and paired t tests. All further analyses were conducted on inspiratory measurements. Analysis of

covariance was used to compare PMA and PMV between individuals with and without COPD, as well as among the four study groups. The covariates in the analysis of covariance model included age, sex, height, body mass index (BMI), pack-years, smoking status, history of heart disease, systemic hypertension, diabetes mellitus, and ConCOLD study center. Spearman correlation coefficients were used to determine the associations of PMA and PMV with demographics (age, height, BMI, and pack-years). Univariable and multivariable linear regression models were constructed to determine associations of the independent PMA and PMV variables with the FEV_1, VO_2, DLCO, and CT scores. The multivariable linear regression models were also similarly adjusted for the same potential confounding variables of age, sex, height, BMI, pack-years, smoking status, history of heart disease, systemic hypertension, diabetes mellitus, and ConCOLD study center. Generalized additive model analysis with smoothing splines indicated that the associations were primarily linear with minor nonlinear components, supporting the use of linear regression, as shown in Figure S1. A base model with the aforementioned covariates was created, and the adjusted R^2 (R_{adj}^2) and Akaike information criterion (AIC) were used to evaluate model fit. PMA and PMV were separately and sequentially added to the model, and the F statistic change, R_{adj}^2, AIC, and standardized β coefficients are reported. In accordance with suggested interpretation guidelines, a ΔAIC greater than 2 indicated moderate evidence of improved fit, and a ΔAIC greater than 4 indicated strong evidence of improved model fit (23). $P < .05$ was considered to indicate a significant difference.

Results

Group Characteristics

A total of 1235 individuals from the CanCOLD project were considered eligible for this study (634 individuals with COPD: mean age, 67.3 years ± 10.1 [SD], 394 male individuals; 601 without COPD: mean age, 65.8 years ± 9.6, 327 male individuals), including 1108 with paired inspiratory and expiratory CT scans for the reproducibility analysis (572 patients with COPD: mean age, 67.1 years ± 10.1, 358 male individuals; 536 without COPD: mean age, 65.7 years ± 9.7, 290 male individuals), and 32 individuals from the iNO study were considered eligible and included in the external test group (31 individuals with COPD: mean age, 68.6 years ± 7.4; 16 male individuals), as described in Figure 1. Complete demographics, pulmonary function scores, exercise capacities, and quality-of-life questionnaire scores for all eligible individuals stratified by COPD status are listed in Table 1. Compared with the individuals without COPD, those with COPD were older ($P = .009$), more likely to be male ($P = .006$), and taller ($P < .001$). In addition, the individuals with COPD had worse FEV_1 ($P < .001$), VO_2 ($P < .001$), DLCO ($P = .001$), and COPD assessment scores ($P < .001$). These characteristics are also reported separately for the individuals in the training and validation, internal test, and external test sets in Table 1. Individuals from the training, validation, internal test, and external test sets were well matched for age and sex, although the external test set individuals had worse pulmonary function ($P < .001$). Characteristics stratified by the four study groups and among the individuals included in the reproducibility analysis are presented in Tables S4 and S5.

Reference Standard Segmentation Reproducibility

The intraobserver reproducibility of the reference standard segmentations on the first batch of training and validation data, as measured with DSC, was excellent (0.95 ± 0.02), while the interobserver reproducibility was very good (0.88 ± 0.04), indicating that the reference standard PMVs on which the final model was trained were consistent and reliable. These DSC values corresponded to intra- and interobserver measurement errors of 22.9 cm^3 ± 13.3 (4.4% ± 2.6) and 66.8 cm^3 ± 28.3 (12.9% ± 5.8), respectively, as summarized in Table 2.

Deep Learning Model Performance

Model performance for each incremental batch added during the active learning process is described in Figure S2. The DSC plateaued after the third batch, indicating that adding more data to the training set would not have markedly altered model performance. In total, the active learning process aided in the successful creation of 112 reference standard segmentations (~16 000 sections) for model training and validation. Table S6 describes the model performance for each fold during the cross-validation training of the final model. The average training time and number of epochs until conversion were 166.7 minutes ± 83.8 and 62.4 epochs ± 10.8, respectively, for the seven folds.

The segmentation accuracy for the seven folds is also described in Table 2. Average model performance, as measured with DSC, was 0.94 ± 0.04, 0.92 ± 0.04, 0.93 ± 0.03, and 0.92 ± 0.04 in the training, validation, internal test, and external test data, respectively. These corresponded to measurement errors of 22.0 cm^3 ± 31.8 (4.4% ± 5.9), 28.0 cm^3 ± 29.5 (5.6% ± 5.7), 36.9 cm^3 ± 28.4 (6.7% ± 5.6), and 38.7 cm^3 ± 27.4 (8.2% ± 6.1), respectively. An example of automated pectoralis segmentation highlighting PMA on the axial section containing the aortic arch and the entire PMV for a healthy male individual who never smoked and a male individual with severe COPD is shown in Figure 4. A reduction in the muscle area and volume is readily apparent in the participant with severe COPD compared with the healthy participant who had never smoked.

Pectoralis Measurement Reproducibility

The Bland-Altman plots generated using the inspiratory and expiratory PMA and PMV are shown in Figure 5. There was a high degree of agreement and no evidence of a bias in the PMV measurements between the two CT scans (bias, 0.1 cm^3 [95% CI: −48.0, 48.3]; $P = .77$). However, the PMA was greater at expiration than at inspiration (bias, −2.7 cm^2 [95% CI: −11.3, 5.9]; $P < .001$). In addition, the coefficient of variation between the inspiration and expiration measurements was lower, whereas the intraclass correlation coefficient and Spearman correlation coefficient were greater for PMV than for PMA ($P < .001$) (Table 3).

Clinical Validation of Pectoralis Measurements

Individuals with COPD had lower PMA (no COPD, 36.9 cm^2 vs COPD, 34.7 cm^2; $P < .001$) and PMV (no COPD, 508.2 cm^3 vs COPD, 488.0 cm^3; $P = .004$) after adjustment for covariates, as shown in Figure 6. Furthermore, compared with all the individuals in the other groups, the individuals in the moderate-severe

Table 1: Demographics, Pulmonary Function, and Cardiopulmonary Exercise Scores of All Study Individuals and Those Only Included in Training and Evaluation of Deep Learning Model

	Clinical Application (n = 1235)		Model Development (n = 176)		
Parameter	No COPD (n = 601)	COPD (n = 634)	Training and Validation (n = 112)	Internal Test (n = 32)	External Test (n = 32)
Age (y)	65.8 ± 9.6	67.3 ± 10.1	65.5 ± 9.1	68.0 ± 10.6	68.6 ± 7.4
Female sex	274 (45.6)	240 (37.9)	54 (48.6)	15 (46.9)	16 (50)
Height (cm)	167.6 ± 9.4	169.5 ± 9.7	167.8 ± 9.6	168.4 ± 8.8	168.4 ± 9.2
BMI (kg/m^2)	27.5 ± 4.9	27.5 ± 5.3	28.1 ± 4.8	29.7 ± 5.7	27.9 ± 5.6
Does not smoke	254 (42.3)	177 (27.9)	34 (30.4)	10 (31.2)	NA
Smoking habit (pack-years)	22.2 ± 19.8	33.9 ± 23.5	22.1 ± 23.7	21.8 ± 29.1	NA
HDHTDM	204 (33.9)	259 (40.9)	50 (44.6)	11 (34.4)	NA
COPD	0 (0)	634 (100.0)	65 (58.0)	16 (50.0)	31 (96.9)
GOLD 1	0 (0)	361 (56.9)	29 (25.9)	9 (28.1)	10 (31.3)
GOLD 2	0 (0)	236 (37.2)	31 (29.5)	5 (15.6)	13 (40.6)
GOLD 3	0 (0)	37 (5.8)	5 (4.8)	2 (6.3)	7 (21.9)
GOLD 4	0 (0)	0 (0)	0 (0)	0 (0)	1 (3.1)
FEV_1 (L)	2.9 ± 0.7	2.4 ± 0.8	2.4 ± 0.8	2.5 ± 0.9	1.7 ± 0.5
FEV_1 (%pred)	103.7 ± 14.0	82.6 ± 18.8	86.5 ± 21.2	91.8 ± 24.2	67.1 ± 17.0
FVC (L)	3.7 ± 1.0	3.8 ± 1.1	3.6 ± 1.0	3.8 ± 1.1	3.4 ± 0.9
FVC (%pred)	104.2 ± 14.3	103.3 ± 18.1	100.1 ± 17.6	106.9 ± 15.2	107.7 ± 19.0
FEV_1:FVC ratio (%)	77.2 ± 4.6	61.3 ± 8.0	66.4 ± 9.6	66.0 ± 14.7	50.0 ± 9.2
DLco (mL/min/mm Hg)	22.4 ± 6.4	20.8 ± 7.1	20.4 ± 6.5	24.0 ± 7.4	NA
DLco (%pred)	99.8 ± 19.9	90.1 ± 22.6	90.8 ± 20.9	107.4 ± 26.7	NA
VO_2 (L/min)	1.8 ± 0.7	1.7 ± 0.6	1.6 ± 0.6	1.7 ± 0.7	NA
VO_2 (%pred)	94.9 ± 24.8	85.1 ± 22.8	85.0 ± 23.1	89.0 ± 16.8	NA
COPD assessment test score (N)	5.7 ± 4.3	8.1 ± 6.5	8.0 ± 6.5	8.1 ± 6.1	NA

Note.—Continuous data are reported as means ± SDs, and categorical data are reported as numbers of individuals, with percentages of total individuals in parentheses. Forced expiratory volume in 1 second (FEV_1) percent of the predicted (%pred) values were calculated using reference equations from reference 35. Diffusion capacity for carbon monoxide (DLco) %pred values were calculated using reference equations from reference 36. Peak rate of oxygen consumption on symptom-limited incremental cardiopulmonary cycle exercise testing (VO_2) %pred values were calculated using reference equations from reference 37. The training and validation data are sourced from eight of nine centers from the CanCOLD study (n = 112). The internal test data are sourced from the ninth center from CanCOLD (n = 32). The external test data are sourced from the inhaled nitric oxide (iNO) study (n = 32) (ClinicalTrials.gov identifier no. NCT03679312). BMI = body mass index (calculated as weight in kilograms divided by height in meters squared), COPD = chronic obstructive pulmonary disease, FVC = forced vital capacity, GOLD = Global Initiative for Chronic Obstructive Pulmonary Disease, HDHTDM = heart disease, systemic hypertension, or diabetes mellitus, NA = data not available.

Table 2: Deep Learning Model Performance for Pectoralis Muscle Volume Segmentation

Parameter	DSC	Jaccard Index	Absolute PMV Error (cm^3)	Relative PMV Error (%)
Reproducibility				
Intraobserver (n = 16)	0.95 ± 0.02	0.91 ± 0.04	22.9 ± 13.3	4.4 ± 2.6
Interobserver (n = 16)	0.88 ± 0.04	0.95 ± 0.02	66.8 ± 28.3	12.9 ± 5.8
Deep learning model				
Training (n = 96)	0.94 ± 0.04	0.88 ± 0.06	22.0 ± 31.8	4.4 ± 5.9
Validation (n = 16)	0.92 ± 0.04	0.86 ± 0.06	28.0 ± 29.5	5.6 ± 5.7
Internal test (n = 32)	0.93 ± 0.03	0.87 ± 0.05	36.9 ± 28.4	6.7 ± 5.6
External test (n = 32)	0.92 ± 0.04	0.85 ± 0.06	38.7 ± 27.4	8.2 ± 6.1

Note.—Data are means ± SDs. Reference standard segmentation reproducibility is reported both within the same human observer and between human observers strictly using the first batch of training and validation data segmented (n = 16). The performance of the deep learning model is reported after training the final model in comparison to the entire set of reference standard segmentations after postprocessing and is the average performance over all seven folds. DSC = Dice similarity coefficient, PMV = pectoralis muscle volume.

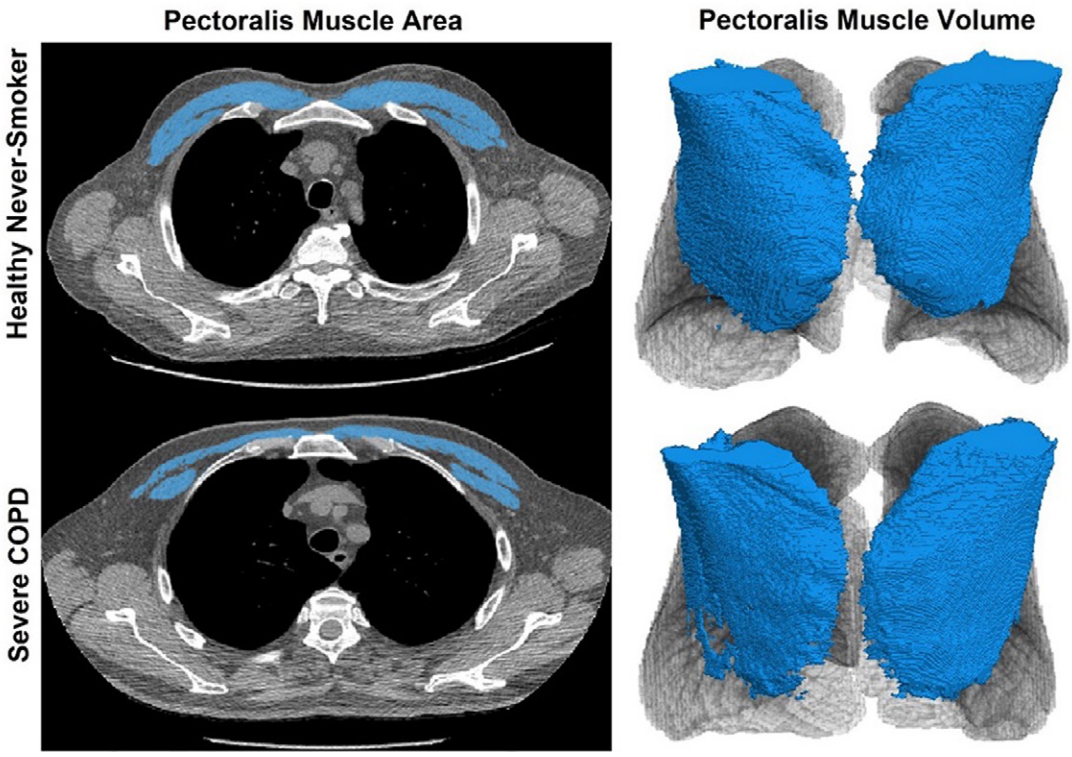

Figure 4: Images show the pectoralis muscles of a healthy male individual who never smoked (age, 66 years; height, 178 cm; body mass index [BMI, calculated as weight in kilograms divided by height in meters squared], 28.4; number of cigarette pack-years, 0; forced expiratory volume in 1 second [FEV_1], 97.6% predicted; FEV_1:forced vital capacity [FVC] ratio, 0.71; pectoralis muscle area [PMA], 59.4 cm^2; pectoralis muscle volume [PMV], 764.0 cm^3) and a male individual with a smoking history and chronic obstructive pulmonary disorder (COPD) (age, 66 years; height, 178 cm; BMI, 27.5; number of cigarette pack-years, 43.2, FEV_1, 48% predicted; FEV_1:FVC, 0.56; PMA, 35.0 cm^2; PMV, 480.8 cm^3) from the Canadian Cohort Obstructive Lung Disease (ie, CanCOLD) study. The CT image is shown in the axial plane. The PMV is automatically extracted using the developed deep learning model and overlayed onto the lungs for visual clarity.

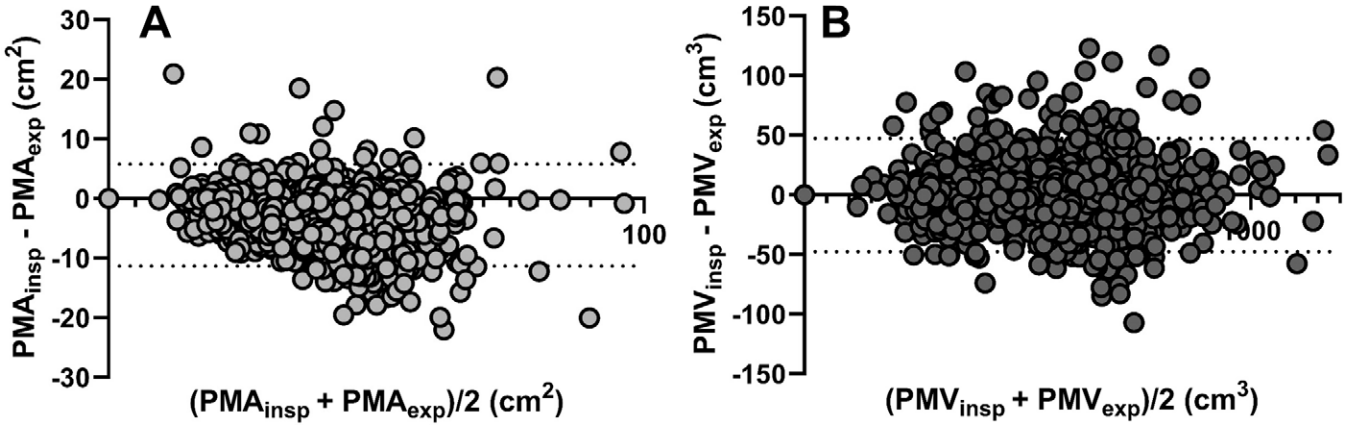

Figure 5: Bland-Altman plots show the relationships between inspiratory (insp) versus expiratory (exp) **(A)** pectoralis muscle area (PMA; bias, −2.7 cm^2 [95% CI: −11.3, 5.9]; $P < .001$) and **(B)** pectoralis muscle volume (PMV; bias, 0.1 cm^3 [95% CI: −48.0, 48.3]; $P = .77$) measurements among individuals from the Canadian Cohort Obstructive Lung Disease (ie, CanCOLD) study with both inspiratory and expiratory CT scans ($n = 1108$).

COPD group had lower PMA and PMV ($P < .05$), as shown in Figure S3. Unadjusted differences are shown in Figure S4.

Spearman correlations between PMA, PMV, and common demographics are shown in Figure S5. PMV was more strongly correlated with age (PMA: ρ, −0.16 [$P < .001$]; PMV: ρ, −0.25 [$P < .001$]) and height (PMA: ρ, 0.61 [$P < .001$]; PMV: ρ, 0.75 [$P < .001$]) than PMA, but neither was correlated with pack-years (PMA: ρ, −0.04 [$P = .13$]; PMV: ρ, −0.02 [$P = .40$]).

Finally, Table 4 shows the univariable and multivariable associations of PMA and PMV with FEV_1, VO_2, DLCO, and COPD assessment scores in individuals with COPD. The associations for all individuals are shown in Table S7. In all univariable models, the PMV yielded higher R_{adj}^2 values and smaller AIC scores than the PMA did. Both PMA and PMV remained significant independent predictors ($P < .001$) when added to the base models (including covariates only), resulting in a change in the F statistic ($P < .001$), an improved R_{adj}^2, and a smaller AIC. However, base models that included the PMV yielded higher R_{adj}^2 and smaller AIC scores than those that included the PMA. For all outcomes except CT, the ΔAIC between the PMV and

Table 3: Reproducibility Analysis between Inspiration and Expiration Pectoralis Muscle Measurements

Parameter	Coefficient of Variation (%)	Intraclass Correlation Coefficient	Spearman Correlation Coefficient
PMA	12.0	0.92 (0.82, 0.96)	0.95
PMV	4.8*	0.99 (0.99, 0.99)*	0.99*

Note.—Values are calculated as inspiratory minus expiratory pectoralis muscle measurements for individuals enrolled in the Canadian Cohort Obstructive Lung Disease (ie, CanCOLD) study who underwent both scans (n = 1108). The intraclass correlation values are reported with 95% CIs in parentheses. PMA = pectoralis muscle area, PMV = pectoralis muscle volume.
* The PMV values are significantly different (P < .001) from the PMA values according to a paired t test analysis following n = 500 bootstrapping with random sampling.

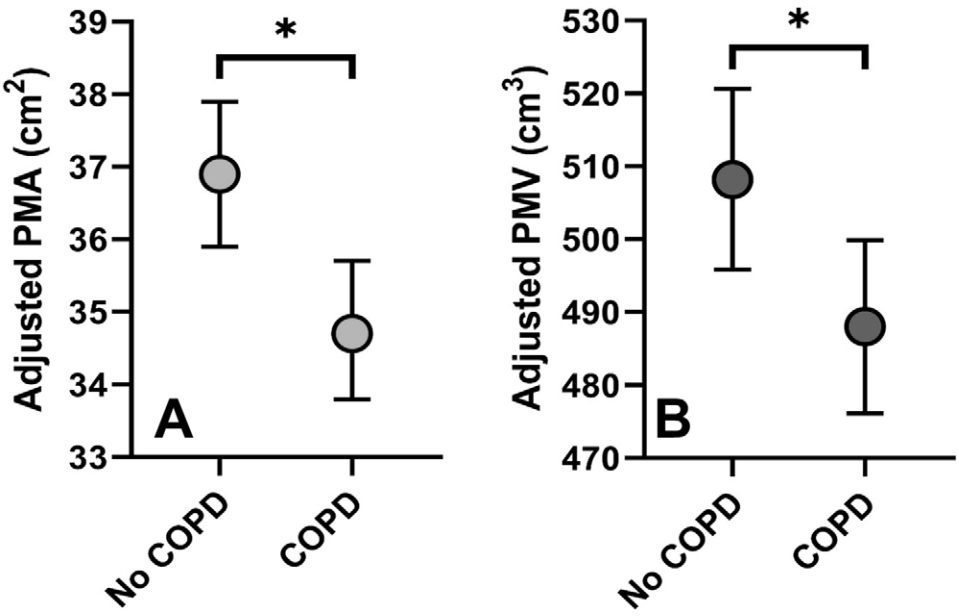

Figure 6: Plots show adjusted differences between **(A)** pectoralis muscle area (PMA) (no chronic obstructive pulmonary disorder [COPD] mean, 36.9 cm² [95% CI: 35.9, 37.9] vs COPD mean, 34.7 cm² [95% CI: 35.7, 33.8]; P < .001) and **(B)** pectoralis muscle volume (PMV) (no COPD mean, 508.2 cm³ [95% CI: 495.8, 520.6] vs COPD mean, 488.0 cm³ [95% CI: 476.1, 499.9]; P = .004) measurements stratified by the absence or presence of COPD using all the inspiratory CT scans of the eligible individuals in the Canadian Cohort Obstructive Lung Disease (ie, CanCOLD) study (n = 1235). The measurements shown are estimated marginal means adjusted for age, sex, height, body mass index, center, smoking status, cigarette pack-years, and the presence of comorbidities (heart disease, hypertension, and diabetes mellitus). * = differences are significantly different between groups (P < .05).

PMA models exceeded 4 units, which is considered a meaningful improvement.

Discussion

In this study, a highly accurate deep learning model that automatically extracts the 3D PMV from chest CT scans was developed (training set DSC: 0.94 ± 0.04; validation set DSC: 0.92 ± 0.04). The deep learning model maintained excellent segmentation performance in unseen hold-out CanCOLD center and unseen external test datasets (internal test set DSC: 0.93 ± 0.03; external test set DSC: 0.92 ± 0.04). PMV (bias, 0.1 cm³ [95% CI: −48.0, 48.3]; P = .77) was also shown to be more reproducible than PMA (bias, −2.7 cm² [95% CI: −11.3, 5.9]; P < .001) between chest CT scans taken in different lung inflation states (inspiration vs expiration). Furthermore, the PMV was lower in individuals with COPD than in those without COPD (no COPD: 508.2 cm³ vs COPD: 488.0 cm³, P = .004) and was more strongly associated with important clinical outcome measures than the PMA was (FEV_1: R_{adj}^2, 0.609 vs 0.598; D_{LCO}: R_{adj}^2, 0.645 vs 0.627; and VO_2: R_{adj}^2, 0.680 vs 0.666). These findings highlight the potential of PMV as a more reliable and clinically meaningful imaging measure for assessing muscle mass in patients with COPD over PMA.

CT-derived PMV has been previously investigated as a surrogate biomarker for muscle mass. Seo et al (24) reported that PMV can be used to quantify muscle loss up to 48 months after radiation therapy, whereas Yokosuka et al (25) reported that PMV is an independent prognostic predictor of future hospitalizations in patients with pneumonia. However, in both studies, the PMV segmentations required time-intensive user input, and automating the process could allow for large-scale analysis of data from many individuals within minutes. Yang et al (12) and Choi et al (13) published deep learning pectoralis volume segmentation models that obtained average DSC values of 0.83 and 0.98, respectively,

Table 4: Univariable and Multivariable Associations between PMA and PMV Measurements and Pulmonary Function, Exercise Capacity, and Quality-of-Life Scores in Patients with COPD

	FEV_1 (L)			VO_2 (L/min)			DLCO (mL/min/mm Hg)			COPD Assessment Test Score (N)		
Parameter	β_{stand}	*P* Value	R_{adj}^2/AIC	β_{stand}	*P* Value	R_{adj}^2/AIC	β_{stand}	*P* Value	R_{adj}^2/AIC	β_{stand}	*P* Value	R_{adj}^2/AIC
Univariable												
PMA	0.57 (0.50, 0.64)	<.001	0.29/1266	0.63 (0.56, 0.70)	<.001	0.37/767	0.56 (0.49, 0.63)	<.001	0.29/3872	−0.18 (−0.27, −0.10)	<.001	0.03/4156
PMV	0.69 (0.63, 0.75)	<.001	0.45/1104	0.72 (0.65, 0.78)	<.001	0.48/666	0.68 (0.62, 0.74)	<.001	0.43/3734	−0.19 (−0.27, −0.11)	<.001	0.03/4153
Multivariable												
BM	NA	NA	0.59/929	NA	NA	0.60/546	NA	NA	0.65/3451	NA	NA	0.13/4115
BM + PMA	0.16 (0.08, 0.24)	<.001	0.60/916	0.29 (0.20, 0.38)	<.001	0.63/507	0.20 (0.12, 0.28)	<.001	0.67/3429	−0.16 (−0.28, −0.04)	.003	0.14/4114
BM + PMV	0.28 (0.18, 0.38)	<.001	0.61/900	0.44 (0.33, 0.54)	<.001	0.65/481	0.34 (0.24, 0.43)	<.001	0.68/3403	−0.17 (−0.32, −0.03)	.006	0.14/4114

Note.—Unless otherwise noted, data are values, with 95% CIs in parentheses, or numerators and denominators. Data based on 634 patients. The data shown for the standardized beta coefficients (β_{stand}) are the standardized parameter estimates. The R^2 values reported for the model are adjusted for the number of input variables (R_{adj}^2). All multivariable models included age, sex, height, body mass index, center, smoking status, pack-years, and the presence of comorbidities (heart disease, hypertension, and diabetes mellitus). The pectoralis muscle area (PMA) and pectoralis muscle volume (PMV) measurements are extracted from all the inspiratory CT scans of the individuals with chronic obstructive pulmonary disease (COPD) in the eligible Canadian Cohort Obstructive Lung Disease (ie, CanCOLD) study. AIC = Akaike information criterion, BM = base model, DLCO = diffusing capacity of the lungs for carbon monoxide, FEV_1 = forced expiratory volume in 1 second, NA = data not available, VO_2 = peak rate of oxygen consumption at symptom-limited incremental cardiopulmonary cycle exercise testing.

on their training data. However, neither study ensured an equal balance of male and female participants, strictly included low-dose scans in their training data, nor reported performance on unseen testing groups. These factors are critical for ensuring generalizability to real-world CT imaging. In contrast, our investigation ensured an equal split of male and female participants and exclusively used low-dose CT scans in the training and validation data. Additionally, model performance in this study was evaluated on two unseen sets of data with different imaging protocols and disease severity. Therefore, the excellent generalizability of the deep learning model reported in this work lends credibility to the real-world capabilities of the developed model, particularly for use in large population-based, low-dose groups, such as the National Lung Screening Trial (26).

To our knowledge, this was the first study to evaluate the reproducibility of both PMA and PMV. Variations in lung volume during CT image acquisition significantly impact quantitative measurements of emphysema (27) and airways (28). However, the impact on measurements of muscle mass is unknown. The findings of this study indicated that the PMA was significantly greater at expiration than at inspiration, whereas the PMV was not. This could be explained by coupled movements of the cardiac-respiratory systems, as the aorta reportedly moves up to 8.9 mm during respiration (29). Consequently, measurements taken at a lower axial section, owing to insufficient inspiration, could lead to the underestimation of the PMA. Although breath-hold volumes are typically coached during CT imaging, inspiratory capacity is often abnormally low because of static hyperinflation (30) in people with obstructive lung diseases such as COPD, potentially explaining the sources of variability reported during evaluation (31). Therefore, this study provides important information on the greater reproducibility of 3D PMV relative to 2D PMA and suggests that the PMV can be assessed at any lung volume that is most comfortable for the patient in the CT scanner. This finding also implies that both intra- and interindividual changes in PMV are unlikely to be caused by differences in lung volume, allowing for evaluations of changes in muscle mass in response to pharmacologic treatments, pulmonary rehabilitation programs, and lifestyle modifications.

Finally, PMV performed comparably to, and in some cases exceeded, the established PMA in group comparisons and association analysis in patients with COPD. Our findings agree with those in previous work showing that pectoralis measurements are significantly lower in individuals with COPD than in those without COPD (8–11) and demonstrate strong associations with common clinical COPD markers (8–11). However, associations between PMV and clinical markers of COPD, such as VO_2 and COPD assessment scores, have never been reported. Therefore, this study presents, to our knowledge, the first evidence that PMV outperforms single-section 2D PMA in terms of these associations, laying an important foundation for the future widespread implementation of PMV as a standard index of muscle

mass on chest CT scans. It is also important to note that unlike 2D measures, volumetric indexes make it possible to study regional variations in composition (eg, fat infiltration) along the muscle, and, although this was not the focus of this study, future work should evaluate these regional differences.

Our study had important limitations. First, we acknowledge that the training and validation group sample size of 112 individuals was relatively small. However, manual segmentation of 3D volumes is time-consuming, and the final sample size was similar to that of previously published volumetric body composition deep learning training sets (32,33), with more than 16 000 individual 2D sections used during training and validation. Importantly, various countermeasures were taken during training, including cross-validation, imaging protocol variation, and internal and external testing. Furthermore, the DSC plateaued after the third batch of added data during the active learning process, indicating that adding more data to the training and validation set did not markedly alter model performance. Second, it is also acknowledged that the craniocaudal field of view (z-axis) fluctuates between chest CT scans and could act as a source of variability (34). However, the superior reproducibility and associations of PMV relative to PMA suggest that while biomarker standardization is an important area of future research, the benefits of 3D measurements outweigh these concerns. Third, the generalizability of the model on chest CT scans collected with larger section thicknesses (~2–3 mm) is unknown and should be evaluated in future work. This study also does not report attenuation or texture measurements, which could offer an enhanced understanding of the composition of the pectoralis muscle, and future work should investigate these indexes. Fourth, and finally, while identifying abnormally low respiratory muscle mass on chest CT scans for referral to pulmonary rehabilitation programs is a potential application of the developed model, healthy cutoffs stratified by age, sex, and BMI are needed to ensure accurate personalized decision-making. Determining abnormally low population cutoffs for PMV measurements should be the focus of future work.

In conclusion, this work presents a highly accurate and generalizable deep learning model to automatically extract the complete 3D PMV from low-dose chest CT scans. Compared with the established 2D single-section PMA, the 3D PMV offered greater reproducibility and showed stronger associations with clinical outcomes in individuals with COPD. These findings encourage the use of the PMV as a surrogate measure of muscle mass in chest CT scans, serving as a potential marker of comorbid sarcopenia in those with COPD. Future studies should investigate normative population reference values to determine abnormally low CT-derived PMV sarcopenic thresholds.

Author affiliations:

[1] Department of Electrical, Computer, and Biomedical Engineering, Toronto Metropolitan University, Toronto, Ontario, Canada

[2] Translational Research in Respiratory Diseases Program, Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada

[3] Center for Heart, Lung Innovation, University of British Columbia, Vancouver, British Columbia, Canada

[4] Montreal Chest Institute of the Royal Victoria Hospital, McGill University Health Centre, Montreal, Quebec, Canada

[5] Respiratory Epidemiology and Clinical Research Unit, Research Institute of McGill University Health Centre, Montreal, Quebec, Canada

[6] Department of Medicine, University of Alberta, Edmonton, Alberta, Canada

[7] Clinical Exercise and Respiratory Physiology Laboratory, Department of Kinesiology and Physical Education, McGill University, Montreal, Quebec, Canada

[8] Department of Physics, Toronto Metropolitan University, 350 Victoria St, Toronto, ON, Canada M5B 2K3

Received February 26, 2025; revision requested July 16; revision received December 29; accepted January 16, 2026.

Address correspondence to: M.K. (email: miranda.kirby@torontomu.ca).

Supplemental material: Supplemental material is available at *Radiology: Cardiothoracic Imaging* online.

Funding: D.G. acknowledges salary support from the Natural Sciences and Engineering Research Council (NSERC). D.J. holds a Canada Research Chair in Clinical Exercise & Respiratory Physiology (Tier 2) from the Canadian Institutes of Health Research. M.K. acknowledges support from NSERC and the Canada Research Chair Program (Tier 2).

Author contributions: Daniel Genkin: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Validation, Visualization, Writing - original draft, Writing - review & editing; Mustansir Verdawala: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing; Sophie Collins: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing - original draft, Writing - review & editing; Wan C Tan: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Validation, Writing - original draft, Writing - review & editing; Pei Zhi Li: Conceptualization, Data curation, Formal analysis, Methodology, Resources, Software, Visualization, Writing - original draft, Writing - review & editing; Jean Bourbeau: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Validation, Writing - original draft, Writing - review & editing; Michael K Stickland: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Validation, Writing - original draft, Writing - review & editing; and Dennis Jensen: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - original draft, Writing - review & editing; Miranda Kirby: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing.

Acknowledgments: We would like to thank the other investigators, staff, and participants of the CanCOLD and iNO studies for their essential contributions. A complete description of CanCOLD and all affiliated institutions can be found at *https://cancold.ca.*

Data sharing: Data analyzed during the study were provided by a third party. Requests for data should be directed to the provider indicated in the Acknowledgments.

Disclosures of conflicts of interest: Please see ICMJE form(s) for author conflicts of interest. These have been provided as supplemental materials.

References

1. Meng SJ, Yu LJ. Oxidative stress, molecular inflammation and sarcopenia. Int J Mol Sci 2010;11(4):1509–1526.
2. Benz E, Trajanoska K, Lahousse L, et al. Sarcopenia in COPD: a systematic review and meta-analysis. Eur Respir Rev 2019;28(154):190049.
3. Choi YJ, Kim T, Park HJ, Cho JH, Byun MK. Long-Term Clinical Outcomes of Patients with Chronic Obstructive Pulmonary Disease with Sarcopenia. Life (Basel) 2023;13(8):1628.
4. Jones SE, Maddocks M, Kon SSC, et al. Sarcopenia in COPD: prevalence, clinical correlates and response to pulmonary rehabilitation. Thorax 2015;70(3):213–218.
5. Gómez-Martínez M, Rodríguez-García W, González-Islas D, et al. Impact of Body Composition and Sarcopenia on Mortality in Chronic Obstructive Pulmonary Disease Patients. J Clin Med 2023;12(4):1321.
6. Global Initiative for Chronic Obstructive Lung Disease. 2024 GOLD Report: Global Strategy for Prevention, Diagnosis and Management of COPD: 2024 Report. GOLD Web site. https://goldcopd.org/2024-gold-report/. Published 2024. Accessed December 2025.
7. O'Brien ME, Zou RH, Hyre N, et al. CT pectoralis muscle area is associated with DXA lean mass and correlates with emphysema progression in a tobacco-exposed cohort. Thorax 2023;78(4):394–401.
8. McDonald MLN, Diaz AA, Ross JC, et al. Quantitative computed tomography measures of pectoralis muscle area and disease severity in chronic obstructive pulmonary disease. A cross-sectional study. Ann Am Thorac Soc 2014;11(3):326–334.

9. Genkin D, Jenkins AR, van Noord N, et al. A fully automated pipeline for the extraction of pectoralis muscle area from chest computed tomography scans. ERJ Open Res 2024;10(1):00485-2023.
10. Bak SH, Kwon SO, Han SS, Kim WJ. Computed tomography-derived area and density of pectoralis muscle associated disease severity and longitudinal changes in chronic obstructive pulmonary disease: a case control study. Respir Res 2019;20(1):226.
11. Wilson AC, Bon JM, Mason S, et al. Increased chest CT derived bone and muscle measures capture markers of improved morbidity and mortality in COPD. Respir Res 2022;23(1):311.
12. Yang Z, Choi I, Choi J, Jung J, Ryu M, Yong HS. Deep learning-based pectoralis muscle volume segmentation method from chest computed tomography image using sagittal range detection and axial slice-based segmentation. PLoS One 2023;18(9):e0290950.
13. Choi I, Choi J, Yong HS, Yang Z. Deep learning-based respiratory muscle segmentation as a potential imaging biomarker for respiratory function assessment. PLoS One 2024;19(7):e0306789.
14. Bourbeau J, Tan WC, Benedetti A, et al; Cancold Study Group. Canadian Cohort Obstructive Lung Disease (CanCOLD): Fulfilling the need for longitudinal observational studies in COPD. COPD 2014;11(2):125–132.
15. Rampuri ZH, Fuhr DP, Brotto AR, et al. Predicting the Clinical Phenotype of COPD Patients Who May Benefit From Selective Pulmonary Vasodilator Therapy During Exercise: A Placebo-controlled Randomized Double-blind Cross-over Trial. Am J Respir Crit Care Med 2023;207:A2535–A2535.
16. Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. Am J Respir Crit Care Med 2019;200(8):e70–e88.
17. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, November 1986. Am Rev Respir Dis 1987;136(1):225–244.
18. Lewthwaite H, Benedetti A, Stickland MK, et al; CanCOLD Collaborative Research Group and the Canadian Respiratory Research Network. Normative Peak Cardiopulmonary Exercise Test Responses in Canadian Adults Aged ≥40 Years. Chest 2020;158(6):2532–2545.
19. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. Eur Respir J 2009;34(3):648–654.
20. Çiçek Ö, Abdulkadir A, Lienkamp SS, T, Ronneberger O. Brox3D U-Net: Learning dense volumetric segmentation from sparse annotation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. Proceedings, Part II. Springer, 2016; 424–432.
21. Wasserthal J, Breit HC, Meyer MT, et al. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. Radiol Artif Intell 2023;5(5):e230024.
22. Aubrey J, Esfandiari N, Baracos VE, et al. Measurement of skeletal muscle radiation attenuation and basis of its biological variation. Acta Physiol (Oxf) 2014;210(3):489–497.
23. Fabozzi FJ, Focardi SM, Rachev ST, Arshanapalli BG. The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications. In: Appendix E: Model Selection Criterion AIC and BIC. Wiley, 2014; 400–403.
24. Seo A, Hwang JM, Lee JM, Jung TD. Changes in Pectoral Muscle Volume During Subacute Period after Radiation Therapy for Breast Cancer: A Retrospective up to 4-year Follow-up Study. Sci Rep 2019;9(1):7038.
25. Yokosuka R, Imai R, Ro S, et al. Pectoralis Muscle Mass on Chest CT at Admission Predicts Prognosis in Patients with Pneumonia. Can Respir J 2021;2021:3396950.
26. Aberle DR, Berg CD, Black WC, et al; National Lung Screening Trial Research Team. The National Lung Screening Trial: overview and study design. Radiology 2011;258(1):243–253.
27. Shaker SB, Dirksen A, Laursen LC, Skovgaard LT, Holstein-Rathlou NH. Volume adjustment of lung density by computed tomography scans in patients with emphysema. Acta Radiol 2004;45(4):417–423.
28. Chen Y, Latisenko R, Lynch DA, Ciet P, Charbonnier JP, Tiddens HAWM. Effect of inspiratory lung volume on bronchial and arterial dimensions and ratios on chest computed tomography in patients with chronic obstructive pulmonary disease. Eur Radiol 2025;35(6):2990–2998.
29. Sailer AM, Wagemans BAJM, Das M, et al. Quantification of Respiratory Movement of the Aorta and Side Branches. J Endovasc Ther 2015;22(6):905–911.
30. Deesomchok A, Webb KA, Forkert L, et al. Lung hyperinflation and its reversibility in patients with airway obstruction of varying severity. COPD 2010;7(6):428–437.
31. Motahari A, Barr RG, Han MK, et al; SPIROMICS Group. Repeatability of Pulmonary Quantitative Computed Tomography Measurements in Chronic Obstructive Pulmonary Disease. Am J Respir Crit Care Med 2023;208(6):657–665.
32. Lee YS, Hong N, Witanto JN, et al. Deep neural network for automatic volumetric segmentation of whole-body CT images for body composition assessment. Clin Nutr 2021;40(8):5038–5046.
33. Koitka S, Kroll L, Malamutmann E, Oezcelik A, Nensa F. Fully automated body composition analysis in routine CT imaging using 3D semantic segmentation convolutional neural networks. Eur Radiol 2021;31(4):1795–1804. [Published correction appears in Eur Radiol 2021;31(6):4402–4403.]
34. Cohen SL, Ward TJ, Cham MD. The relationship between CT scout landmarks and lung boundaries on chest CT: guidelines for minimizing excess z-axis scan length. Eur Radiol 2020;30(1):581–587.
35. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, Enright PL, Hankinson JL, Ip MS, Zheng J, Stocks J; ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J. 2012;40(6):1324–43.
36. Garcia-Rio F, Dorgham A, Galera R, Casitas R, Martinez E, Alvarez-Sala R, Pino JM. Prediction equations for single-breath diffusing capacity in subjects aged 65 to 85 years. Chest. 2012;142(1):175–184.
37. Lewthwaite H, Benedetti A, Stickland MK, et al. Normative Peak Cardiopulmonary Exercise Test Responses in Canadian Adults Aged ≥40 Years. Chest. 2020 Dec;158(6):2532-2545.