



Enhancing COPD classification using combined quantitative computed tomography and texture-based radiomics: a CanCOLD cohort study

Kalysta Makimoto ¹, James C. Hogg², Jean Bourbeau ^{3,4}, Wan C. Tan², Miranda Kirby ^{1,2}
and the CanCOLD Collaborative Research Group

¹Toronto Metropolitan University, Toronto, ON, Canada. ²Center for Heart, Lung Innovation, University of British Columbia, Vancouver, BC, Canada. ³Montreal Chest Institute of the Royal Victoria Hospital, McGill University Health Centre, Montreal, QC, Canada. ⁴Respiratory Epidemiology and Clinical Research Unit, Research Institute of McGill University Health Centre, Montreal, QC, Canada.

Corresponding author: Miranda Kirby (miranda.kirby@torontomu.ca)



Shareable abstract (@ERSpublications)

Texture-based CT radiomics provides additional and complementary information to conventional qCT features on lung structural changes, which improves COPD classification <https://bit.ly/3x4Yrdm>

Cite this article as: Makimoto K, Hogg JC, Bourbeau J, *et al.* Enhancing COPD classification using combined quantitative computed tomography and texture-based radiomics: a CanCOLD cohort study. *ERJ Open Res* 2024; 10: 00968-2023 [DOI: 10.1183/23120541.00968-2023].

Copyright ©The authors 2024

This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

Received: 2 Dec 2023
Accepted: 13 March 2024

Abstract

Background Recent advances in texture-based computed tomography (CT) radiomics have demonstrated its potential for classifying COPD.

Methods Participants from the Canadian Cohort Obstructive Lung Disease (CanCOLD) study were evaluated. A total of 108 features were included: eight quantitative CT (qCT), 95 texture-based radiomic and five demographic features. Machine-learning models included demographics along with texture-based radiomics and/or qCT. Combinations of five feature selection and five classification methods were evaluated; a training dataset was used for feature selection and to train the models, and a testing dataset was used for model evaluation. Models for classifying COPD status and severity were evaluated using the area under the receiver operating characteristic curve (AUC) with DeLong's test for comparison. SHapely Additive exPlanations (SHAP) analysis was used to investigate the features selected.

Results A total of 1204 participants were evaluated (n=602 no COPD; n=602 COPD). There were no differences between the groups for sex (p=0.77) or body mass index (p=0.21). For classifying COPD status, the combination of demographics, texture-based radiomics and qCT performed better (AUC=0.87) than the combination of demographics and texture-based radiomics (AUC=0.81, p<0.05) or qCT alone (AUC=0.84, p<0.05). Similarly, for classifying COPD severity, the combination of demographics, texture-based radiomics and qCT performed better (AUC=0.81) than demographics and texture-based radiomics (AUC=0.72, p<0.05) or qCT alone (AUC=0.79, p<0.05). Texture-based radiomics and qCT features were among the top five features selected (15th percentile of the CT density histogram, CT total airway count, pack-years, CT grey-level distance zone matrix zone distance entropy, CT low-attenuation clusters) for classifying COPD status.

Conclusion Texture-based radiomics and conventional qCT features in combination improve machine-learning models for classification of COPD status and severity.

Introduction

COPD is a common, progressive and incurable lung disease often associated with tobacco smoking [1]. Individuals with COPD have increased symptom burden [2], reduced exercise capacity [3], risk of exacerbations [2] and a doubled risk of lung cancer [4–6]. While spirometry is the standard technique used to detect airflow obstruction to confirm the presence of COPD [1], it is often not performed in primary care or in lung cancer screening trials, where COPD is largely underdiagnosed [7]. However, low-dose computed tomography (CT) scans are used for lung cancer screening and offer an opportunity to identify undiagnosed COPD. Estimates suggest a significant portion of screening trial participants have COPD [6]; therefore, there is potential to use this information in risk models to determine screening eligibility and frequency.



CT imaging is used to quantify structural lung changes in COPD, including the extent of emphysema [8] and airway remodelling [9]. In lung cancer screening trials, CT emphysema has been shown to classify individuals with COPD [10]; however, many individuals with COPD have no or minimal emphysema [11]. Another pathological feature of COPD is airway remodelling, and the addition of CT airway measurements to emphysema improves COPD classification accuracy in lung cancer screening trials [12]. However, these conventional quantitative CT (qCT) measurements provide a single global measurement and may miss subtle changes in the lung tissue characteristics.

Emerging texture-based radiomics can provide additional information about the spatial patterns in an image [13]. CT texture-based radiomics have been shown in previous studies to classify COPD status [14–16], COPD severity [15] and progression to COPD [17]. A previous study demonstrated that radiomic features, including texture-based radiomics and shape features, extracted from low-dose CT images improved accuracy for COPD detection compared to CT emphysema alone in a cohort with moderate-severe COPD [14]. However, there are several CT emphysema measures that reflect both emphysema extent and clustering [18, 19]. Further, CT airway remodelling measurements have been shown to provide complementary information that improves COPD classification [20] and numerous features derived from the CT airway tree that have prognostic relevance in COPD [9, 20, 21]. Therefore, these CT measurements should also be considered in COPD classification models.

We hypothesised that the addition of texture-based radiomics to conventional qCT, including both CT emphysema and airway measures, would provide additional information about disease heterogeneity among individuals with COPD, which will improve machine-learning (ML) model performance for COPD classification. To investigate our hypothesis, we aimed to compare the performance of ML models with conventional CT measurements, texture-based CT radiomic measurements and their combination for classifying COPD status and severity. We also aimed to investigate the CT features being selected in our models to determine which features are most predictive.

Materials and methods

Study participants

The participants evaluated in this study were from Canadian Cohort Obstructive Lung Disease (CanCOLD), a longitudinal population-based study [22]. Participants were excluded if they were missing CT images at the baseline visit (V1) or V1 demographics or spirometry (figure 1). The remaining

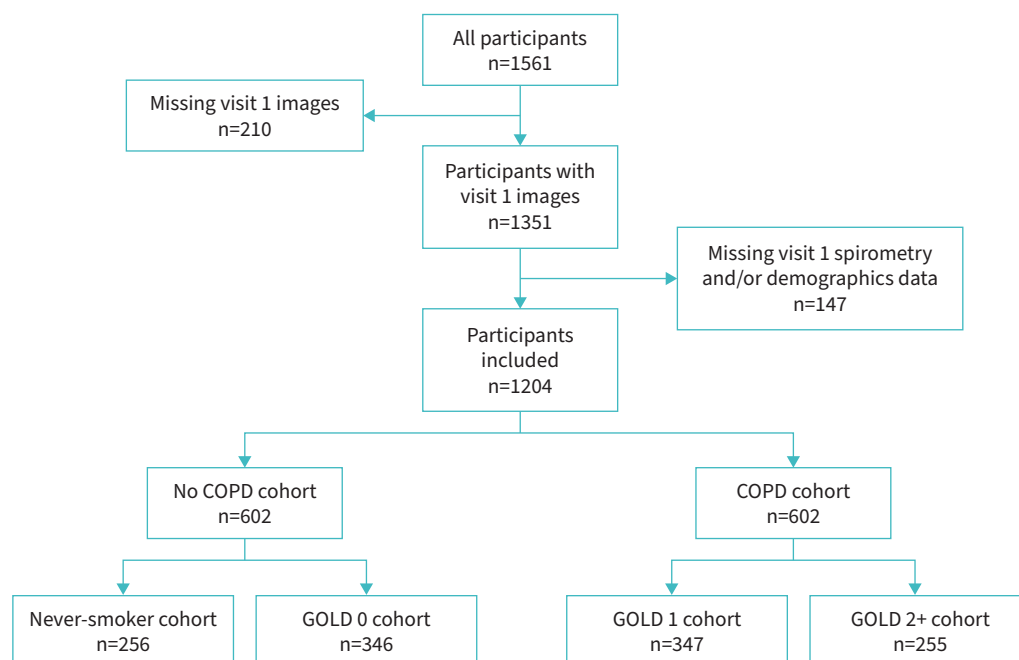


FIGURE 1 CONSORT diagram illustrating the participants included. Any participant visit 1 CT images, spirometry and/or demographics data were excluded. Participants were then classified as “No COPD” or “COPD” based on their baseline forced expiratory volume in 1 s/forced vital capacity measurements.

participants were then grouped into “No COPD” and “COPD” based on their baseline spirometry (forced expiratory volume in 1 s (FEV₁)/forced vital capacity (FVC)). The groups were further classified into never-smokers and Global Initiative for Chronic Obstructive Lung Disease (GOLD) 0, GOLD 1 and GOLD 2+ based on the GOLD guidelines [1]. Written informed consent was obtained from all participants and the study received institutional review board approval at each study site (REB 2019-244). The same CanCOLD participant population has been previously reported [16, 23].

Spirometry and questionnaires

The American Thoracic Society guidelines [24] were used to obtain spirometry measurements at V1. Questionnaires were collected at V1, including COPD Assessment Test (CAT) scores [25] and St George’s Respiratory Questionnaire (SGRQ) scores [26].

CT image acquisition and CT features

The CT images were acquired using a multi-slice CT scanner in the supine position from the apex to base of the lung at full inspiration, based on CanCOLD imaging guidelines [22]. CT scanner acquisition details are shown in supplementary table S1.

Conventional qCT features were extracted using VIDA Diagnostics Inc. clinical image analysis service (Apollo 2.0 software package, Coralville, IA, USA), except for normalised join count (NJC) and total airway count (TAC). The emphysema qCT features extracted include low-attenuation areas below -950 HU (LAA₉₅₀) [8], 15th percentile of the density histogram (HU₁₅) [27], low-attenuation clusters (LAC) [19] and NJC [18]. The airway qCT features extracted include TAC [21], average wall thickness for a hypothetical airway of 10 mm lumen perimeter [9], the average airway wall area % (WA%) for five segmental airways (RB1, RB4, RB10, LB1 and LB10) [20] and the average airway lumen area (LA) for five segmental airways [20].

To extract the texture-based CT radiomic features, an in-house-developed pipeline that uses the Standardized Environment for Radiomics Analysis [28] (MATLAB-based framework) was constructed to calculate the features in compliance with the Image Biomarker Standardisation Initiative (IBSI) [13]. Prior to radiomic feature extraction, CT images were pre-processed by resampling the voxels to a 1 mm³ isotropic resolution, segmentation masks were used to remove the airways and extract the lungs, and a threshold of -1000 HU to 0 HU was applied [23]. The 95 texture-based radiomics [13] that were extracted included 25 grey-level co-occurrence matrix (GLCM) features, 16 grey-level run-length matrix (GLRLM) features, 16 grey-level size zone matrix (GLSZM) features, 16 grey-level distance zone matrix (GLDZM) features, 17 neighbouring grey-level dependence matrix (NGLDM) features and five neighbourhood grey-tone difference matrix (NGTDM) features. To account for the various feature scales, a z-normalisation (mean of 0 and sd of 1) was applied to all features.

Machine-learning pipeline

Figure 2 illustrates the proposed methods. Demographics (age, sex, race, body mass index (BMI), height, smoking status and pack-years) and CT features, including eight qCT and 95 texture-based radiomics, were obtained for all participants. For each combination of features that were investigated, the dataset was split into a training dataset (75%) and a testing dataset (25%), such that the class labels were balanced. Outlier removal was performed on the training dataset by removing outliers outside 2 sd for a Gaussian distribution [16]. Pearson’s correlation coefficients were used to determine highly correlated features in the training dataset ($|r| \geq 0.90$), which were removed from further analyses. Following outlier removal and the removal of highly correlated features, the training dataset was then input into various feature selection methods to select five features [16] to train the models with a 5-fold cross-validation to tune the hyper-parameters for classifying binary COPD status and multiclass COPD severity. The hold-out testing dataset was used to evaluate the models. Combinations of five feature selection methods and five classification methods (supplementary table S2), available from the scikit-learn package (version 0.24.2; Python version 3.9.6), were evaluated. The feature selection and classification combination that obtained the highest area under the receiver operating characteristic (ROC) curve (AUC) value in the test dataset was determined to be the optimal ML model for each outcome being investigated and the performance metrics for the optimal model are reported.

Statistical analysis

To evaluate the ML model performance, various performance metrics were used in the testing dataset, including AUC, accuracy and F1 scores. DeLong’s test was used to compare AUC values for significant differences [29]. To investigate the impact of the features selected on the model, SHapely Additive exPlanations (SHAP) analysis [30] summary plots were investigated. Pearson’s correlation test was used to

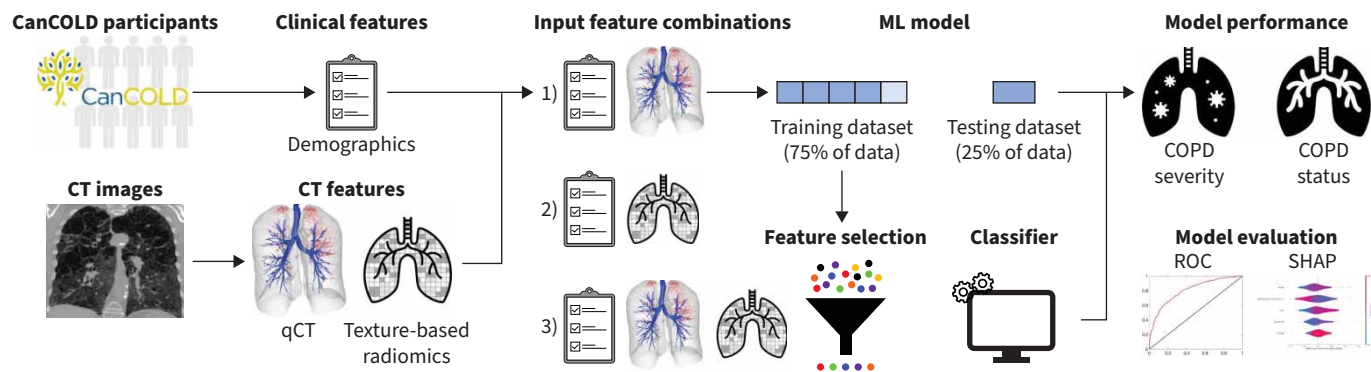


FIGURE 2 Proposed methods. Machine-learning (ML) models were constructed using different combinations of five demographic, eight quantitative computed tomography (qCT) and 95 texture-based CT radiomics measurements. The dataset was split into a 5-fold cross-validation training dataset (75% of the data) and testing dataset (25% of the data). The training dataset was used with feature selection methods to select five features, which were then input into a ML classifier to be trained. The ML models were then tested with the testing dataset for COPD status and COPD severity classification. ROC: receiver operating characteristic; SHAP: SHapely Additive exPlanations.

evaluate correlations between CT features and baseline spirometry measurements and symptoms scores (CAT and SGRQ). An ANOVA was performed for texture-based radiomics between the groups (never-smokers, GOLD 0, GOLD 1 and GOLD 2+).

Results

Study participant demographics

Table 1 shows the demographics and lung function measurements for the participants included in this study. There was no significant difference between the “No COPD” and “COPD” participants for sex (p=0.77) or BMI (p=0.21) (table 1). The COPD participants were slightly but significantly older, with greater pack-years of smoking and lower lung function than the participants without COPD (p<0.05). There were no significant differences between the participants included (1204 participants) and the participants excluded (357 participants) from the study for age (p=0.40), sex (p=0.86) or BMI (p=0.10)

TABLE 1 Subject demographics for participants included

	No COPD	COPD	p-value
Subjects (n)	602	602	
Age (years)	66.16±9.64	67.38±9.94	0.03
Female sex	279 (46)	232 (39)	0.77
BMI (kg·m⁻²)	27.74±5.14	27.37±4.94	0.21
Pack-years	10.9±18.2	22.7±25.1	<0.001
Smoking status			
Never	308 (51)	196 (33)	0.03
Former	238 (40)	302 (50)	0.19
Current	56 (9)	104 (17)	0.08
FEV₁ (L)	2.75±0.77	2.36±0.76	<0.001
FEV₁ (% pred)	100.2±17.4	83.3±18.9	<0.001
FVC (L)	3.58±1.01	3.83±1.10	<0.001
FVC (% pred)	97.7±16.5	101.0±17.5	<0.001
FEV₁/FVC	77.1±4.7	61.4±7.9	<0.001
GOLD grade			
GOLD 0	346 (57)	–	–
GOLD 1	–	347 (58)	–
GOLD 2	–	224 (37)	–
GOLD 3	–	31 (5)	–
GOLD 4	–	0 (0)	–

Data are presented as mean±SD or n (%), unless otherwise indicated. Bold text indicates a statistically significant difference at p<0.05. BMI: body mass index; FEV₁: forced expiratory volume in 1 s; FVC: forced vital capacity; GOLD: Global Initiative for Chronic Obstructive Lung Disease.

(supplementary table S3). The outlier cohort had slightly but significantly lower BMI ($p=0.01$), FEV₁ ($p=0.003$) and FVC ($p=0.002$) measurements than the training dataset (supplementary table S4). There were no significant differences between age ($p=0.80$), sex ($p=0.42$), BMI ($p=0.41$), pack-years ($p=0.97$) or COPD labels ($p=0.91$) between the training and testing cohorts (supplementary table S5).

Machine-learning model performance

COPD status

Table 2 shows the performance metrics for the models with demographics, qCT and/or texture-based radiomics for classifying binary COPD status in the testing dataset. Figure 3 shows the ROC curves and SHAP analyses for the different feature set combinations for binary COPD status classification. The combination of demographics, qCT and texture-based radiomics features (AUC=0.87) significantly increased the AUC value compared to demographics with qCT features alone (AUC=0.84, $p<0.05$), and compared to demographics with texture-based radiomics (AUC=0.81, $p<0.05$). However, there was no significant difference between demographics with qCT (AUC=0.84) and demographics with texture-based radiomics (AUC=0.81, $p>0.05$).

The SHAP analyses (figure 3) show that for the model with demographics and qCT, as well as demographics, qCT and texture-based radiomics, lower HU₁₅, lower TAC, higher pack-years and higher LAC values had high positive contributions to the models, which demonstrates that these features are likely to predict a positive class for COPD status. When texture-based radiomics was included in the model with demographics and qCT, a texture-based radiomic feature (GLDZM_{zone distance entropy} (GLDZM_{zdentr})) was selected, instead of age, as a top feature.

A sensitivity analysis was performed to evaluate the addition of texture-based radiomics to qCT features without demographics (supplementary table S6 and supplementary figure S1). It also showed that two texture-based radiomics (GLDZM_{large distance low grey-level emphasis} and NGTDM_{coarseness}) were selected and replaced two qCT features (LAA₉₅₀ and WA%) in the models with qCT and texture-based radiomics (supplementary figure S1). In all the model combinations that included qCT features, TAC and HU₁₅ were consistently selected as important features for COPD status.

COPD severity

Table 2 shows the performance metrics for the models with demographics, qCT and/or texture-based radiomics for classifying multiclass COPD severity in the testing dataset. Figure 4 shows the ROC curves and SHAP analyses for the different feature set combinations for COPD severity classification. The combination of demographics, texture-based radiomics and qCT features (AUC=0.81) significantly increased the AUC value compared to demographics with qCT features (AUC=0.79, $p<0.05$) and to demographics with texture-based radiomics (AUC=0.72, $p<0.05$). However, the AUC for the model with demographics and qCT (AUC=0.79) was significantly higher than for the model with demographics and texture-based radiomics (AUC=0.72, $p<0.05$).

The SHAP analysis (figure 4) shows that for the model with demographics, qCT and texture-based radiomics, as well as the model with demographics and qCT, higher NJC, lower TAC, higher pack-years and average higher smoking status had high positive contributions to the model, demonstrating the model is likely to predict higher severity groups. In the model with qCT and texture-based radiomics, a texture-based radiomics feature (GLDZM_{zone distance non-uniformity} (GLDZM_{zdnv})) was selected, instead of a qCT feature (WA%), as a top feature.

A sensitivity analysis was performed to evaluate the addition of texture-based radiomics to qCT features alone (supplementary table S6 and supplementary figure S2). It also showed that a texture-based radiomic feature (GLDZM_{zdnv}) was selected and replaced a qCT feature (LA) in the model including texture-based radiomics with qCT features (supplementary figure S2). Also, in all the model combinations including qCT features, TAC and NJC were consistently selected as important features for COPD severity.

Figure 5 shows a coronal CT image slice, a pre-processed CT image slice used to extract the texture-based radiomics and three-dimensional images for never-smoker, GOLD 0, GOLD 1 and GOLD 2+ participants. The three-dimensional images show that participants with greater disease severity had more emphysema, shown in red, indicated by the increased LAA₉₅₀ values. The pre-processed image shows that with increasing disease severity there was an increased pattern of parenchymal heterogeneity, reflected by the increased GLDZM_{zdentr} and GLDZM_{zdnv} values.

TABLE 2 Models comparing the impact of the addition of texture-based radiomics to conventional measurements (demographics and qCT features) for classifying COPD status and COPD severity in the testing dataset

Feature combination	Outcome					
	COPD status [#]			COPD severity [¶]		
	AUC (95% CI)	Accuracy (95% CI)	F1 score (95% CI)	AUC (95% CI)	Accuracy (95% CI)	F1 score (95% CI)
Demographics+qCT	0.84 (0.80–0.88)	0.77 (0.72–0.82)	0.77 (0.72–0.82)	0.79 (0.74–0.83)	0.56 (0.50–0.62)	0.56 (0.52–0.63)
Demographics+texture-based radiomics	0.81 (0.76–0.85)	0.74 (0.69–0.79)	0.74 (0.69–0.79)	0.72 (0.67–0.77) ⁺	0.46 (0.41–0.52)	0.43 (0.38–0.49)
Demographics+qCT+texture-based radiomics	0.87 (0.83–0.91) ^{+,§}	0.79 (0.74–0.84)	0.79 (0.74–0.84)	0.81 (0.74–0.84) ^{+,§}	0.57 (0.51–0.62)	0.56 (0.50–0.61)

qCT: quantitative computed tomography; AUC: area under the receiver operating characteristic curve. [#]: machine-learning model: Lasso feature selection method+support vector machine classification method; [¶]: machine-learning model: gradient boosting decision tree feature selection method+multilayer perceptron classification method; ⁺: significantly different AUC from demographics and qCT model (p<0.05); [§]: significantly different AUC from demographics+texture-based radiomics model (p<0.05).

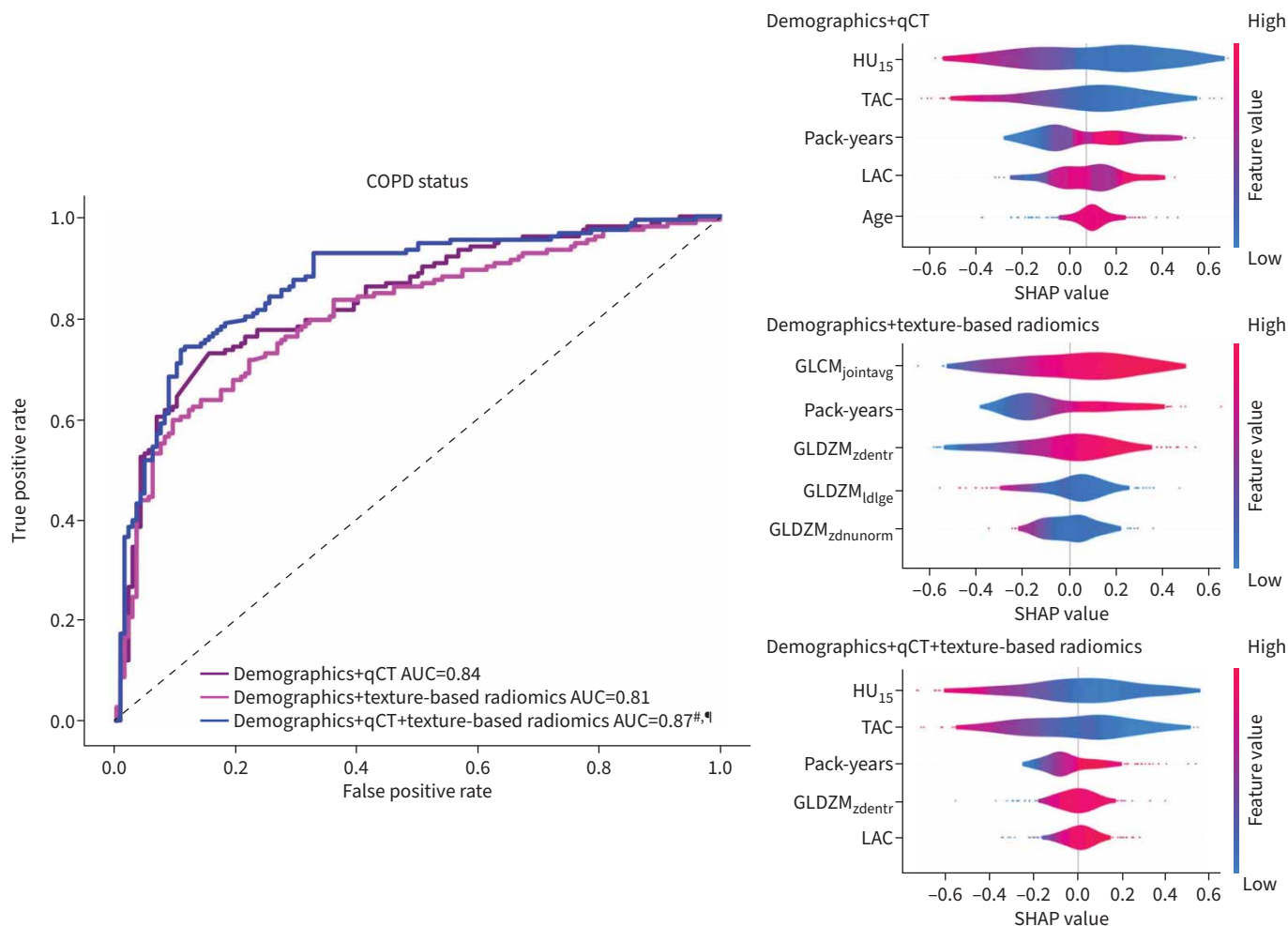


FIGURE 3 Receiver operating characteristic curves and SHapely Additive exPlanations (SHAP) analysis for COPD status with different input feature set combinations. qCT: quantitative computed tomography; AUC: area under the receiver operating characteristic curve; HU₁₅: 15th percentile of the density histogram; TAC: total airway count; LAC: low-attenuation clusters; GLCM_{jointavg}: grey-level co-occurrence matrix (GLCM) joint average; GLDZM_{zdentr}: grey-level distance zone matrix (GLDZM) zone distance entropy; GLDZM_{ldlge}: GLDZM large distance low grey-level emphasis; GLDZM_{zdnunorm}: GLDZM zone distance non-uniformity normalised. #: significantly different AUC from demographics and qCT model; *: significantly different AUC from demographics and texture-based radiomics model.

Pearson correlations for top CT features with lung function and symptom scores

The Pearson's correlation coefficients for lung function and symptoms with the top CT features selected by the all-features ML model (demographics, qCT and texture-based radiomics) for classifying COPD status and COPD severity are shown in table 3. All texture-based radiomic features and qCT features were significantly correlated with FEV₁ ($p < 0.05$), FVC ($p < 0.05$) and FEV₁/FVC ($p < 0.05$). For CAT and SGRQ scores, only CT TAC, NJC and GLDZM_{zdentr} features had significant correlations ($p < 0.05$).

Supplementary table S7 shows the Pearson's correlation coefficients for all the qCT and texture-based radiomics features selected by the ML models without demographics (supplementary figures S1 and S2). The ANOVA between groups shows that qCT and texture-based radiomic features (GLDZM_{small distance high grey-level emphasis}, GLDZM_{zdnunorm}, GLDZM_{grey-level non-uniformity}, GLCM_{joint maximum} and NGTDM_{coarseness}) were able to differentiate the severity groups (supplementary table S8).

Discussion

COPD remains largely undiagnosed, yet CT imaging is acquired for clinical investigation and in lung cancer screening trials, which presents the opportunity to identify COPD in populations at risk. In a multicentre, population-based study consisting of individuals with and without smoking histories, and with

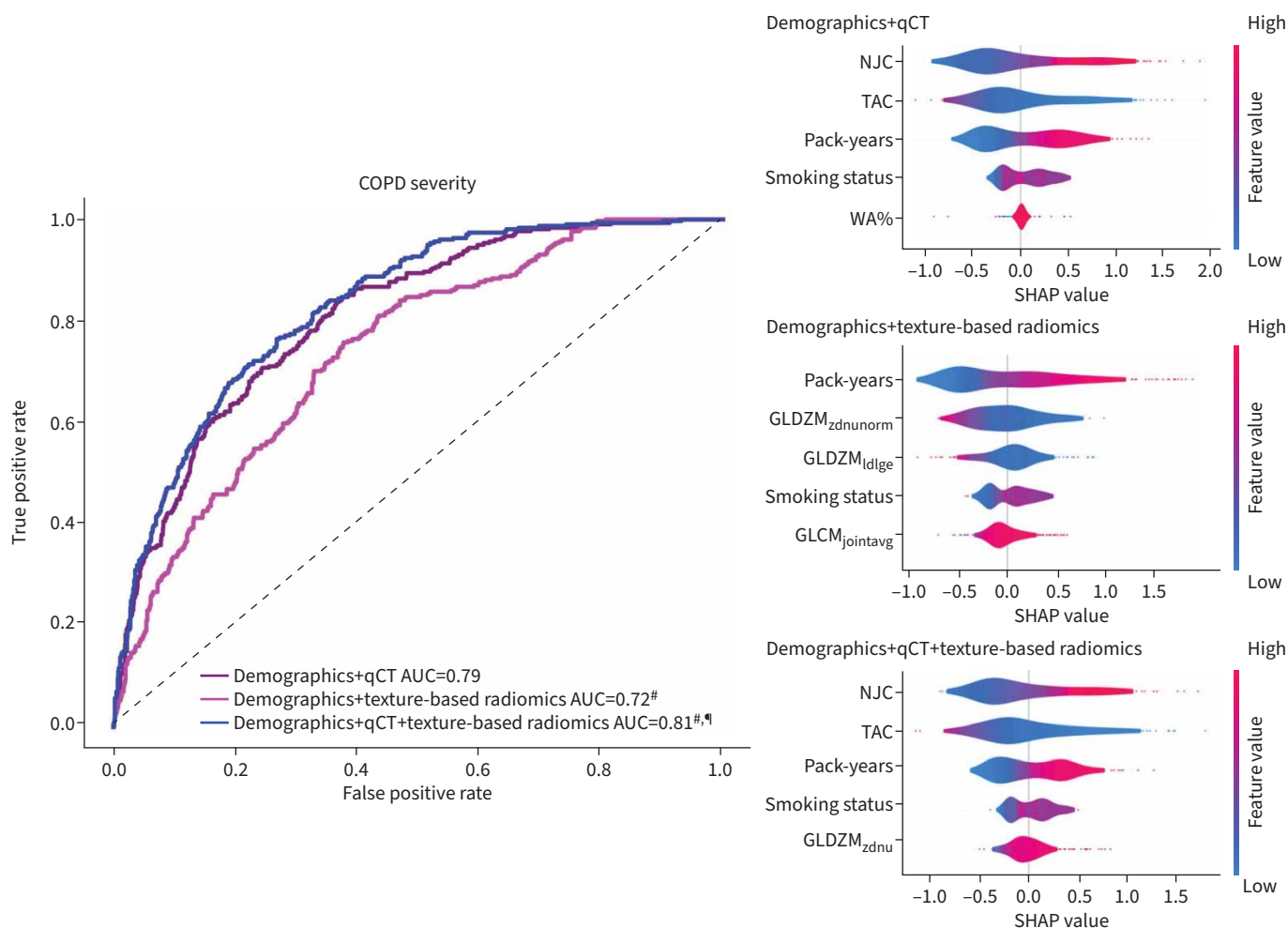


FIGURE 4 Receiver operating characteristic curves and SHapely Additive exPlanations (SHAP) analysis for COPD severity with different input feature set combinations. qCT: quantitative computed tomography; AUC: area under the receiver operating characteristic curve; NJC: normalised joint count; TAC: total airway count; WA%: wall area %; GLDZM_{zdnunorm}: grey-level distance zone matrix (GLDZM) zone distance non-uniformity normalised; GLDZM_{ldlge}: GLDZM large distance low grey-level emphasis; GLCM_{jointavg}: grey-level co-occurrence matrix joint average; GLDZM_{zdnunorm}: GLDZM zone distance non-uniformity. #: significantly different AUC from demographics and qCT model; #,: significantly different AUC from demographics and texture-based radiomics model.

mainly mild COPD, we aimed to investigate the performance of texture-based CT radiomic features, conventional qCT features and their combination for predicting COPD status and severity with ML. Our results show that the combination of texture-based CT radiomics and conventional qCT features resulted in significantly improved model performance for classifying COPD status and COPD severity than either feature set alone. Our results also demonstrate that a variety of texture-based CT radiomics and qCT features reflecting emphysema and airway remodelling were among the most important features affecting the ML model performance.

Our results show that a combination of CT imaging features, both texture-based radiomics and more conventional CT emphysema and airway disease measurements, classified COPD status and severity. Similar results were found in our sensitivity analysis (supplementary table S5), where we demonstrated the benefit of using only CT imaging to classify COPD. In agreement with these findings, several studies show that conventional qCT features classify COPD status and severity in moderate-severe COPD cohorts [31, 32]. There are also studies that show texture-based radiomics can classify COPD status and severity with ML [14, 15]. However, few studies have investigated the combination of texture-based radiomics and qCT features for COPD classification. One recent study by PUCHAKAYALA *et al.* [14] demonstrated that the addition of a single CT emphysema measurement to CT radiomic features did not significantly improve the performance of a COPD classification model in comparison to a model that included only CT radiomics

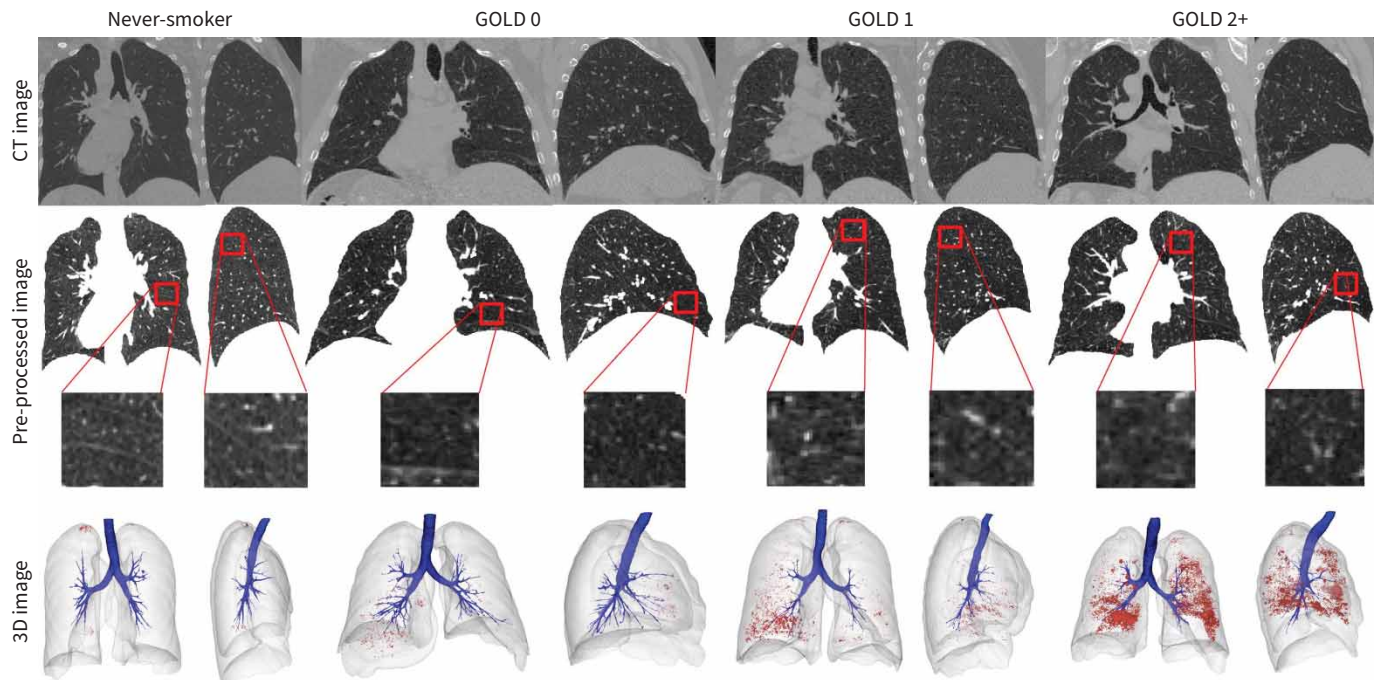


FIGURE 5 Original computed tomography (CT) image, pre-processed CT image and three-dimensional image for never-smoker, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 0, GOLD 1 and GOLD 2+ participants. Never-smoker: 57-year-old female, $FEV_1/FVC=70.46$, $FEV_1=2.29$ L, $GLDZM_{z\text{dnu}}=1.14e^{+5}$, $GLDZM_{z\text{dentr}}=13.3$, $LAA_{950}=0.4$. GOLD 0: 65-year-old male, $FEV_1/FVC=80.76$, $FEV_1=2.56$ L, $GLDZM_{z\text{dnu}}=1.17e^{+5}$, $GLDZM_{z\text{dentr}}=13.3$, $LAA_{950}=4.3$. GOLD 1: 63-year-old female, $FEV_1/FVC=60.45$, $FEV_1=2.16$ L, $GLDZM_{z\text{dnu}}=1.19e^{+5}$, $GLDZM_{z\text{dentr}}=13.4$, $LAA_{950}=4.6$. GOLD 2+: 65-year-old male, $FEV_1/FVC=52.76$, $FEV_1=1.72$ L, $GLDZM_{z\text{dnu}}=1.11e^{+5}$, $GLDZM_{z\text{dentr}}=13.4$, $LAA_{950}=7.0$. FEV_1 : forced expiratory volume in 1 s; FVC: forced vital capacity; $GLDZM_{z\text{dnu}}$: grey-level distance zone matrix (GLDZM) zone distance non-uniformity; $GLDZM_{z\text{dentr}}$: GLDZM zone distance entropy; LAA_{950} : low attenuation areas below -950 HU.

features. However, in contrast with our study, their study only included one conventional qCT feature and not all the IBSI-defined texture-based radiomic features. Our study included a variety of conventional qCT features that capture the various structural changes that occur in COPD, with all six IBSI-defined texture-based radiomic features. Our cohort also consisted of patients with mainly mild COPD, which may be why our model performance was poorer than that of PUCHAKAYALA *et al.* [14]; however, our results demonstrate the potential of these features for identifying COPD in the early/mild disease stages. Taken together, these findings further illustrate the benefit of both texture-based radiomics and conventional qCT measurements in ML models for classifying COPD status and severity, especially in the early and mild disease stages.

TABLE 3 Pearson's correlation coefficients (r) for CT features (all qCT and texture-based radiomics selected in the machine-learning models) with baseline spirometry measurements for the whole cohort

Feature	FEV_1		FVC		FEV_1/FVC		CAT score [#]		SGRQ total score [¶]	
	r	p-value	r	p-value	r	p-value	r	p-value	r	p-value
HU ₁₅	-0.11	<0.001	-0.31	<0.001	0.40	<0.001	0.03	0.33	-0.01	0.97
TAC	0.29	<0.001	0.14	<0.001	0.34	<0.001	-0.14	<0.001	-0.18	<0.001
NJC	-0.08	0.005	0.13	<0.001	-0.43	<0.001	0.08	0.004	0.10	0.001
$GLDZM_{z\text{dnu}}$	0.41	<0.001	0.57	<0.001	-0.26	<0.001	-0.02	0.40	-0.01	0.79
$GLDZM_{z\text{dentr}}$	0.12	<0.001	0.16	<0.001	-0.06	0.04	0.10	<0.001	0.07	0.04

Total cohort n=1204. Bold text indicates a statistically significant difference at $p<0.05$. qCT: quantitative computed tomography; FEV_1 : forced expiratory volume in 1 s; FVC: forced vital capacity; CAT: COPD Assessment Test; SGRQ: St George's Respiratory Questionnaire; HU₁₅: 15th percentile of the density histogram; TAC: total airway count; NJC: normalised join count; $GLDZM_{z\text{dnu}}$: grey-level distance zone matrix (GLDZM) zone distance non-uniformity; $GLDZM_{z\text{dentr}}$: GLDZM zone distance entropy. #: n=1180 owing to missing data; ¶: n=957 owing to missing data.

We found that both texture-based CT radiomic features and conventional qCT features were selected in our ML models as important predictive features for COPD status and severity. In the models for COPD status that included all feature sets (demographics, qCT and texture-based radiomics), the CT features selected included three qCT features (HU₁₅, LAC and TAC) and one texture-based radiomics features (GLDZM_{zdenr}). In the model with all feature sets for COPD severity, the CT features selected included two qCT features (NJC and TAC) and one texture-based radiomic feature (GLDZM_{zdnv}). In both the COPD status and severity models, the CT TAC measurement, which can differentiate COPD severity and predict lung function decline [21, 33], was selected as a top feature and was the only qCT airway feature selected in both models. A previous study that included airway shape radiomic features with texture-based radiomics also demonstrated that the airway shape features were selected as important features [14]. In our models for classifying COPD status and severity, the texture-based radiomic features that were selected as important were from the GLDZM feature set (GLDZM_{zdenr} and GLDZM_{zdnv}), which has not been investigated in other studies using texture-based radiomics for COPD classification [14]. Compared to the other texture-based radiomic feature sets that only consider the grey levels, the GLDZM feature set considers both the grey levels and the distance to the border in the image, indicating that the GLDZM feature set is a hybrid of texture and morphological features [13]. We also demonstrated that texture-based radiomics are significantly correlated with lung function and symptoms (supplementary table S7) and can be used to differentiate COPD severity groups (supplementary table S8), demonstrating their usefulness in a mild disease population where there is minimal disease present that may not be detected visually or by CT emphysema measurements. Overall, these results indicate that texture-based radiomic features and conventional qCT features are able to quantify the heterogeneous patterns within the lung and provide complementary information.

A main strength of our study is that we used participants from a population-based mild COPD cohort that may better reflect very subtle structural changes that occur in early/mild disease. However, there are limitations of this study to be addressed. First, CanCOLD is a multicentre study with various CT scanners, which introduce variability into texture-based radiomic measurements [34], and therefore the results may not be generalisable to other CT systems. The number of COPD participants included was limited and may have contributed to nonsignificant results in the DeLong's test because it is dependent on sample size [29]. However, using a larger dataset may allow for significant differences to be determined. Additionally, although we used a 5-fold cross-validation to train the model and a hold-out testing dataset for model evaluation, this study lacked an external dataset for validation purposes to test the models for generalisability. Future studies should investigate the combination of texture-based radiomic features compared to conventional features in other external cohorts to validate the models. One advantage of texture-based radiomics is that it does not require airway or vessel segmentation, which can be a complex and time-consuming process. However, with recent advances in deep learning, it is now possible to automatically segment numerous anatomical structures [35]. This development significantly enhances the accessibility of both texture-based radiomics and qCT features for ML models used in risk prediction.

Interpretation

In our cohort of individuals with mainly mild COPD, we demonstrated the benefit of combining texture-based radiomics and conventional CT measurements for classifying COPD status and severity with ML. These findings indicate that models investigating COPD outcomes should include a variety of CT features that can capture various disease characteristics.

Provenance: Submitted article, peer reviewed.

Acknowledgments: The authors would also like to thank the men and women who participated in the study. This work in this article was previously presented at the American Thoracic Society Conference in 2023.

CanCOLD Collaborative Research Group: Jonathon Samet (Keck School of Medicine of USC, Los Angeles, CA, USA); Milo Puhan (John Hopkins School of Public Health, Baltimore, MD, USA); Qutayba Hamid, Carolyn Baglole, Palmira Mancino, Pei-Zhi Li, Zhi Song, Dennis Jensen and Benjamin McDonald Smith (McGill University, Montreal, QC, Canada); Yvan Fortier and Mina Dligui (Sherbrooke University, Sherbrooke, QC, Canada); Kenneth Chapman, Jane Duke, Andrea S. Gershon and Teresa To (University of Toronto, Toronto, ON, Canada); J. Mark Fitzgerald and Mohsen Sadatsafavi (University of British Columbia, Vancouver, BC, Canada); Christine Lo, Sarah Cheng, Elena Un, Michael Cheng, Cynthia Fung, Nancy Haynes, Liyun Zheng, LingXiang Zou, Joe Comeau, Jonathon Leipsic and Cameron Hague (UBC James Hogg Research Center, Vancouver, BC, Canada); Brandie L. Walker and Curtis Dumonceaux (University of Calgary, Calgary, AB, Canada); Paul Hernandez and Scott Fulton (University of Dalhousie, Halifax, NS, Canada); Shawn Aaron and Kathy Vandemheen (University of Ottawa, Ottawa, ON, Canada); Denis O'Donnell, Matthew McNeil and Kate Whelan (Queen's University, Kingston, ON, Canada); François Maltais

and Cynthia Brouillard (University of Laval, Quebec City, QC, Canada); and Darcy Marciniuk, Ron Clemens and Janet Baran (University of Saskatchewan, Saskatoon, SK, Canada).

Ethics statement: Written informed consent was obtained from all participants and the study received institutional review board approval at each study site (REB 2019-244).

Author contributions: K. Makimoto contributed substantially to the study design, data analysis and interpretation, and the writing of the manuscript. J.C. Hogg, J. Bourbeau, W.C. Tan and M. Kirby had full access to all of the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis. M. Kirby had final approval of the version to be published.

Conflict of interest: We would like to note that there is no overlap with our study and other previously published CanCOLD studies. Further, there are no conflicts of interest or industry support in relation to this project for any of the authors. M. Kirby is a consultant for VIDA Diagnostics Inc. (Coralville, IA, USA).

Support statement: M. Kirby acknowledges support from the Natural Sciences and Engineering Research Council Discovery Grant, the Early Researchers Award Program and the Canada Research Chair Program (Tier II). Funding information for this article has been deposited with the Crossref Funder Registry.

References

- 1 Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global Strategy for the Prevention, Diagnosis and Management of COPD. 2023. Available from: <https://goldcopd.org/>
- 2 Ding B, Small M, Bergström G, et al. COPD symptom burden: impact on health care resource utilization, and work and activity impairment. *Int J Chron Obstruct Pulmon Dis* 2017; 12: 677–689.
- 3 Liesker JJW, Wijkstra PJ, Ten Hacken NHT, et al. A systematic review of the effects of bronchodilators on exercise capacity in patients with COPD. *Chest* 2002; 121: 597–608.
- 4 Gao YH, Guan WJ, Liu Q, et al. Impact of COPD and emphysema on survival of patients with lung cancer: a meta-analysis of observational studies. *Respirology* 2016; 21: 269–279.
- 5 Young RP, Hopkins RJ, Christmas T, et al. COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. *Eur Respir J* 2009; 34: 380–386.
- 6 Amirahmadi R, Kumar AJ, Cowan M, et al. Lung cancer screening in patients with COPD – a case report. *Medicina (Kaunas)* 2019; 55: 364.
- 7 Lamprecht B, Soriano JB, Studnicka M, et al. Determinants of underdiagnosis of COPD in national and international surveys. *Chest* 2015; 148: 971–985.
- 8 Müller NL, Staples CA, Miller RR, et al. “Density mask”: an objective method to quantitate emphysema using. *Chest* 1988; 94: 782–787.
- 9 Charbonnier JP, Pompe E, Moore C, et al. Airway wall thickening on CT: relation to smoking status and severity of COPD. *Respir Med* 2019; 146: 36–41.
- 10 Mets OM, Buckens CFM, Zanen P, et al. Identification of chronic obstructive pulmonary disease in lung cancer screening computed tomographic scans. *JAMA* 2011; 306: 1775–1781.
- 11 Lynch DA, Austin JHM, Hogg JC, et al. CT-definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner Society. *Radiology* 2015; 277: 192–205.
- 12 Mets OM, Schmidt M, Buckens CF, et al. Diagnosis of chronic obstructive pulmonary disease in lung cancer screening computed tomography scans: independent contribution of emphysema, air trapping and bronchial wall thickening. *Respir Res* 2013; 14: 59.
- 13 Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* 2020; 295: 328–338.
- 14 Amudala Puchakayala PR, Sthanam VL, Nakhmani A, et al. Radiomics for improved detection of chronic obstructive pulmonary disease in low-dose and standard-dose chest CT scans. *Radiology* 2023; 307: e222998.
- 15 Li Z, Liu L, Zhang Z, et al. A novel CT-based radiomics features analysis for identification and severity staging of COPD. *Acad Radiol* 2022; 29: 663–673.
- 16 Makimoto K, Au R, Moslemi A, et al. Comparison of feature selection methods and machine learning classifiers for predicting chronic obstructive pulmonary disease using texture-based CT lung radiomic features. *Acad Radiol* 2023; 30: 900–910.
- 17 Makimoto K, Hogg JC, Bourbeau J, et al. CT imaging with machine learning for predicting progression to COPD in individuals at risk. *Chest* 2023; 164: 1139–1149.
- 18 Virdee S, Tan WC, Hogg JC, et al. Spatial dependence of CT emphysema in chronic obstructive pulmonary disease quantified by using join-count statistics. *Radiology* 2021; 301: 702–709.
- 19 Gietema HA, Müller NL, Nasute Fauerbach PV, et al. Quantifying the extent of emphysema: factors associated with radiologists’ estimations and quantitative indices of emphysema severity using the ECLIPSE cohort. *Acad Radiol* 2011; 18: 661–671.

- 20 Smith BM, Hoffman EA, Rabinowitz D, *et al.* Comparison of spatially matched airways reveals thinner airway walls in COPD. The Multi-Ethnic Study of Atherosclerosis (MESA) COPD study. *Thorax* 2014; 69: 987–996.
- 21 Kirby M, Tanabe N, Tan WC, *et al.* Total airway count on computed tomography and the risk of chronic obstructive pulmonary disease progression. Findings from a population-based study. *Am J Respir Crit Care Med* 2018; 197: 56–65.
- 22 Bourbeau J, Tan WC, Benedetti A, *et al.* Canadian Cohort Obstructive Lung Disease (CanCOLD): fulfilling the need for longitudinal observational studies in COPD. *COPD* 2014; 11: 125–132.
- 23 Au RC, Tan WC, Bourbeau J, *et al.* Impact of image pre-processing methods on computed tomography radiomics features in chronic obstructive pulmonary disease. *Phys Med Biol* 2021; 66: 245015.
- 24 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
- 25 Jones P, Harding G, Wiklund I, *et al.* Improving the process and outcome of care in COPD: development of a standardised assessment tool. *Prim Care Respir J* 2009; 18: 208–215.
- 26 Jones PW, Quirk FH, Baveystock CM. The St George’s Respiratory Questionnaire. *Respir Med* 1991; 85: Suppl. B, 25–37.
- 27 Shaker SB, Stavngaard T, Laursen LC, *et al.* Rapid fall in lung density following smoking cessation in COPD. *COPD* 2011; 8: 2–7.
- 28 Ashrafinia S. Quantitative Nuclear Medicine Imaging Using Advanced Image Reconstruction and Radiomics. Baltimore, MD, Johns Hopkins University, 2019.
- 29 DeLong E, DeLong D, Clarke-Pearson D. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44: 837–845.
- 30 Lipovetsky S, Conklin M. Analysis of regression in game theory approach. *Appl Stoch Model Bus Ind* 2001; 17: 319–330.
- 31 Yang K, Yang Y, Kang Y, *et al.* The value of radiomic features in chronic obstructive pulmonary disease assessment: a prospective study. *Clin Radiol* 2022; 77: e466–e472.
- 32 Castaldi PJ, Estépar RSJ, Mendoza CS, *et al.* Distinct quantitative computed tomography emphysema patterns are associated with physiology and function in smokers. *Am J Respir Crit Care Med* 2013; 188: 1083–1090.
- 33 Kirby M, Smith BM, Tanabe N, *et al.* Computed tomography total airway count predicts progression to COPD in at-risk smokers. *ERJ Open Res* 2021; 7: 00307-2021.
- 34 Yasaka K, Akai H, Mackin D, *et al.* Precision of quantitative computed tomography texture analysis using image filtering: a phantom study for scanner variability. *Medicine (Baltimore)* 2017; 96: e6993.
- 35 Wasserthal J, Breit HC, Meyer MT, *et al.* TotalSegmentator: robust segmentation of 104 anatomic structures in CT images. *Radiol Artif Intell* 2023; 5: e230024.