

Original Research

Emphysema Detection in Smokers: Diffusing Capacity for Nitric Oxide Beats Diffusing Capacity for Carbon Monoxide-Based Models

Gerald S. Zavorsky, PhD, RRT¹ Roberto W. Dal Negro, MD² Ivo van der Lee, MD, PhD³ Alexandra M. Preisser, MD⁴

Abstract

Background: Pulmonary diffusing capacity for nitric oxide (DLNO) remains underutilized despite potential advantages over pulmonary diffusing capacity for carbon monoxide (DLCO). We evaluated whether DLNO better detects emphysema than DLCO, spirometry, or lung volumes in smokers.

Methods: We performed an individual participant data meta-analysis of adult smokers (14–43 pack years) with and without computed tomography-defined emphysema using a standardized 10±2-second breath-hold time double diffusion protocol. Variables were converted to z-scores. Prespecified models contrasted DLCO- versus DLNO-based approaches. Model selection used the Bayesian information criterion (BIC) and leave-one-out information criterion; discrimination used area under the receiver operating characteristic (AUROC) curve and Matthews correlation coefficient (MCC) with repeated cross-validation.

Results: After harmonization and quality control, 408 participants (85 emphysema, 323 controls) were analyzed. The lowest BIC (164.6) occurred for the 3-predictor model with total lung capacity (TLC), forced expiratory volume in 1 second (FEV₁), and DLNO z-scores, with an 88% probability of being superior to the next-lowest BIC model (168.5). Discrimination (AUROC 0.97, 95% confidence interval [CI] 0.95–0.98) and classification (MCC 0.80, 95% CI 0.69–0.89) were high. Hierarchical partitioning showed unique contributions from FEV₁ z-scores ($R^2=0.35$) > DLNO z-scores ($R^2=0.21$) > TLC z-scores ($R^2=0.11$), totaling McFadden's $R^2=0.663$. Adding DLCO z-scores increased the total R^2 trivially (by 0.003) and contributed largely shared information with DLNO (variance inflation factors ≤ 4.5). Category-free reclassification and Youden-threshold analyses showed small but favorable gains; the case–control risk gap improved by up to ~5% when adding DLNO to a DLCO-based model.

Interpretation: When predicting the likelihood of emphysema in smokers, a parsimonious z-score model comprising TLC, FEV₁, and DLNO z-scores provides excellent performance and stable rank superiority.

1. Department of Physiology and Membrane Biology, University of California, Davis, Davis, California, United States
2. CESFAR–National Center for Respiratory Pharmacoeconomics and Pharmacoepidemiology, Verona, Italy
3. Department of Pulmonology, Spaarne Gasthuis, Haarlem, The Netherlands
4. Institute for Occupational and Maritime Medicine, University Medical Center Hamburg–Eppendorf, Hamburg, Germany

Abbreviations:

AIC=Akaike information criterion; **AUROC**=area under the receiver operating characteristic; **BH**=Benjamini-Hochberg procedure; **BIC**=Bayesian information criterion; **BMI**=body mass index; **CI**=confidence interval; **CO**=carbon monoxide; **CT**=computed tomography; **DLCO**=diffusing capacity for carbon monoxide; **DLNO**=diffusing capacity for nitric oxide; **ELPD**=expected log pointwise predictive density; **FDA**=U.S. Food and Drug Administration; **FDR**=forced discovery rate; **FEV₁**=forced expiratory volume in 1 second; **FVC**=forced vital capacity; **GAMLSS**=generalized additive models of location, scale, and shape; **GLI**=Global Lung Function Initiative; **Hb**=hemoglobin; **IDI**=integrated discrimination index; **IPD**=individual patient data; **IQR**=interquartile range; **IV**=inspired volume; **LR**=likelihood ratio; **KCO**=carbon monoxide transfer coefficient; **KNO**=nitric oxide transfer coefficient; **LASSO**=least absolute shrinkage and selection

operator; **MCC**=Matthews correlation coefficient; **LOOIC**=leave-one-out information criteria; **NO**=nitric oxide; **NPV**=negative predictive value; **NRI**=net reclassification improvement; **OFF**=out of fold; **PC**=principal component; **PCA**=principal component analysis; **PFTs**=pulmonary function tests; **PPV**=positive predictive value; **PSis**=Pareto-smoothed importance-sampling; **RBC**=red blood cell; **RV**=residual volume; **SE**=standard error; **SLR**=segmented linear regression; **TLC**=total lung capacity; **ULN**=upper limit of normal defined as the 95th percentile; **VA**=alveolar volume; **VIF**=variance inflation factor

Funding Support:

None.

Citation:

Zavorsky GS, Dal Negro RW, van der Lee I, Preisser AM. Emphysema detection in smokers: diffusing capacity for nitric oxide beats diffusing capacity of carbon monoxide-based models. *Chronic Obstr Pulm Dis*. 2025;12(6):500-511. doi: <https://doi.org/10.15326/jcopdf.2025.0645>

Publication Dates:

Date of Acceptance: October 29, 2025

Published Online Date: November 10, 2025

For personal use only. Permission required for all other uses.

Address correspondence to:

Gerald S. Zavorsky, PhD, RRT
 Department of Physiology and Membrane Biology
 University of California, Davis
 Davis, CA
 Email: gszavorsky@health.ucdavis.edu

Keywords:

pulmonary diffusing capacity; z-scores; logistic modeling; predictive accuracy; emphysema

This article has an online supplement.**Introduction**

Emphysema remains a major health concern in the United States, affecting an estimated 3.8 million individuals¹ in 2018, with an age-adjusted mortality rate² of approximately 9.5 per 100,000 adults in 2020. This disease destroys alveoli, compromises lung elasticity, increases air trapping, and leads to dyspnea. Chronic inflammation and progressive alveolar capillary membrane damage culminate in irreversible airflow limitation. Cigarette smoking is the primary risk factor driving lung injury.^{3–6} Despite improvements in diagnostics, early-stage emphysema often goes unrecognized, delaying treatment and negatively influencing clinical and quality-of-life outcomes.

Standard pulmonary function tests—such as spirometry (measuring forced expiratory volume in 1 second [FEV₁], forced vital capacity [FVC], and their ratio) and pulmonary diffusing capacity for carbon monoxide (DLCO)—are useful for aiding diagnosis and monitoring. However, spirometry primarily assesses airway obstruction, while DLCO is affected by pulmonary capillary blood volume and hemoglobin concentration⁷ and may lack sensitivity to early alveolar damage. Pulmonary diffusing capacity for nitric oxide (DLNO), first introduced in 1983–84 as abstracts,^{8,9} offers a more direct assessment of alveolar-capillary membrane function.¹⁰

According to the classical Roughton–Forster model, the overall resistance to gas transfer is partitioned into membrane and red blood cell (RBC) components.¹¹ In this framework, resistance within the RBC interior is considered the principal limitation to carbon monoxide (CO) uptake (DLCO).^{11,12} Recent modeling work, however, indicates that DLNO is dominated by diffusion across the plasma boundary layer and reaction at the RBC surface membrane, rather than by processes within the RBC interior itself.^{12,13} Then, there is a completely different framework on NO and CO uptake, apart from the Roughton–Forster model. In this new framework, DLNO is interpreted as "surface absorption" dominated by the membrane-plasma path and accessible RBC surface, whereas DLCO reflects "volume absorption" that scales with engaged RBC volume and hematocrit.^{14,15}

Regardless of which framework is ultimately correct, the NO–CO double diffusion technique remains rarely used in clinical practice, even 42 years after its introduction.^{8,9} While

commercially available devices exist for this measurement,¹⁶ their clinical adoption is limited by a lack of clinician awareness and the absence of U.S. Food and Drug Administration (FDA) approval for any NO–CO device. As a result, use of this method has largely been confined to research settings by a small group of specialized investigators. Given the distinct technical advantages of measuring DLNO,¹⁷ it is important for manufacturers to pursue the necessary regulatory approvals so that DLNO can become as routine as DLCO in standard pulmonary function testing.

Studies show that the alveolar uptake for NO (KNO) has better sensitivity in detecting emphysema than DLCO,¹⁸ and that summed DLNO+DLCO z-scores outperform DLCO z-scores alone in model performance, predictive accuracy, and classification scores.^{7,19} DLNO and KNO also correlate more closely with computed tomography (CT) markers of emphysema better than DLCO or the carbon monoxide transfer coefficient (KCO).^{20,21} By enabling earlier and more accurate detection of emphysema, DLNO could facilitate timely, targeted interventions. Early identification of disease is clinically important, as patients with undiagnosed COPD face poorer outcomes and reduced quality of life²² while early diagnosis and management can reduce health care utilization and improve quality of life.²³ Detecting emphysema before substantial functional decline enables timely, evidence-based interventions—such as risk stratification, smoking-cessation support, pulmonary rehabilitation, and individualized care—that can slow disease progression and reduce emphysema progression in quitters.²⁴ Pulmonary rehabilitation, in particular, improves exercise tolerance, reduces dyspnea, and enhances quality of life.^{22,25,26}

This study examines DLNO's diagnostic performance, accuracy, and classification ability in emphysema patients compared to DLCO, spirometry, and lung volumes. Using a large cohort—predominantly smokers—from 3 hospital centers, we applied z scores derived from established reference equations.^{27–31} We hypothesized that DLNO and DLCO z scores would outperform conventional metrics in diagnosing emphysema. If DLNO proves more accurate, it could be adopted routinely alongside DLCO. This adoption could facilitate earlier diagnosis, improve patient-centered outcomes, and stimulate the development and regulatory approval of accessible DLNO measurement equipment—overcoming current technological and logistical barriers.¹⁶

Methods**Study Design and Population**

We conducted an individual participant data (IPD) meta-analysis pooling raw, participant-level data to harmonize variables, standardize analyses, and increase precision.³² The pooled dataset included 496 White participants (mostly smokers; interquartile range 6–43 pack years): 126 with computed tomography (CT)-confirmed emphysema and

370 without, from 4 European hospital centers.^{18,33-35} After harmonization, 3 centers were retained because they consistently used the simultaneous 10-second NO–CO protocol NO–CO testing^{18,33,34}; all 4 source datasets remain available in a public repository.³⁶ All original studies had ethics approval; this de-identified secondary analysis did not require additional review.

Data Collection, Conversion, and Quality Control

Pulmonary function tests (PFTs) followed American Thoracic Society/European Respiratory Society guidelines,³⁷⁻³⁹ measuring DLNO, DLCO, alveolar volume (VA), KCO, and KNO with a 10±2-second breath-hold time (denoted DLNO_{10s}, DLCO_{10s}, VA_{10s}, KCO_{10s}, and KNO_{10s}). Lung function variables were converted to z-scores using Global Lung Function Initiative (GLI) reference equations for spirometry,²⁸ lung volumes,²⁷ and DLCO_{10s},²⁹ adjusting for age, sex, and height. For the NO–CO double diffusion technique, z-scores were derived from reference equations developed with 10-second breath-hold maneuvers³⁰ and from equations that account for between device variability.^{31,40} Because available DLNO reference equations were derived in White cohorts^{30,31} and genetic ancestry influences DLNO,⁴¹⁻⁴³ analyses were restricted to White participants.

Study-level quality was graded across 9 items: (1) inclusion of COPD and non-COPD participants, (2) pack years, (3) radiologist-adjudicated CT emphysema (% volume), (4) smoking history, (5) modified Medical Research Council dyspnea scale, (6) sex, (7) height, (8) weight (coded “not provided” if imputed), and (9) technical quality control. Technical failure criteria were breath-hold outside 8–12 seconds; VA/total lung capacity (TLC) >1.0; FEV₁/FVC ≥1.0; residual volume (RV)/TLC <0.20; or inspired-volume (IV) to FVC (IV/FVC) <0.85. Studies with <5% failures met technical standards. The RV/TLC <0.20 rule excluded physiologically implausible values. A summary score (0–9) tallied the 8 availability items plus the quality control flag (Table S1 in the online supplement). R packages are listed in Tables S2–S3 in the online supplement.

Model Discovery and Comparator Definition (Post-Selection)

We assembled 34 candidate logistic models from clinically plausible and data-driven combinations of z-scores (FEV₁, FVC, FEV₁/FVC, TLC, RV/TLC, VA, DLCO_{10s}, KCO_{10s}, DLNO_{10s}, KNO_{10s}). To encourage parsimony, we screened with least absolute shrinkage and selection operator (LASSO) and ranked all candidates using Bayesian information criterion (BIC) and Pareto-smoothed importance-sampling—leave-one-out information criterion (PSIS-LOOIC) on the analysis dataset (lower values indicate superior expected out-of-sample fit and parsimony). Guided by this screening, we

defined 3 focal comparators for all downstream evaluation:

- Model A: TLC z-scores+FEV₁ z-scores+DLCO_{10s} z-scores
- Model B: Model A+DLNO_{10s} z-scores
- Model C: TLC z-scores+FEV₁ z-scores+DLNO_{10s} z-scores

TLC, FEV₁ and DLCO_{10s} z-scores were fitted using GLI equations²⁷⁻²⁹ and DLNO z-scores were fitted using generalized additive models of location, scale, and shape (GAMLSS) reference equations of Zavorsky and Cao.³¹

This workflow is data-adaptive/post-selection: information-criterion screening of individual/summed predictors informed LASSO model building, and all candidates were ultimately compared on the same information criterion scale before selecting Models A–C.

Objectives, Endpoints, and Hypotheses

Based on the results above, a primary objective was to test whether the parsimonious 3-predictor DLNO_{10s} model (Model C) is noninferior to the analogous DLCO_{10s} model (Model A) for detecting CT-defined emphysema in adult smokers, compared to smokers without emphysema, while achieving better parsimony/generalizability. A key secondary objective was to assess whether adding DLNO_{10s} to Model A (which is Model B) or expanding to higher-dimension variants yields clinically meaningful gains over Model C after accounting for complexity.

The primary endpoint was out-of-fold (OOF) Matthews correlation coefficient (MCC) on held-out folds, using thresholds learned via Youden’s J in training and applied to the paired test fold (MCC family under Benjamini–Hochberg [BH] control for prespecified contrasts). Another endpoint was the BIC and PSIS-LOOIC computed on the analysis dataset to quantify parsimony/generalization.

Secondary endpoints included the area under the receiver operating characteristic curve (AUROC) in the test folds (threshold-free discrimination), decision curve analysis net benefit across risk thresholds of 0–0.25, and a set of threshold-based performance metrics—accuracy, balanced accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), F1 score, Cohen’s kappa, false positive rate, false negative rate, false discovery rate (FDR), positive and negative likelihood ratios (LR+ and LR–), and the diagnostic odds ratio—each calculated out-of-fold at the Youden-derived threshold. Exploratory endpoints included category-free net reclassification improvement (NRI), integrated discrimination improvement (IDI), calibration intercept and slope with bootstrap confidence intervals (CIs), principal component analysis (PCA) loadings and variance explained, and hierarchical partitioning of McFadden’s R².

The primary hypothesis was that Model C is noninferior to Model A on OOF, MCC, and AUROC and shows lower BIC and/or PSIS-LOOIC (i.e., superior parsimony/generalization). The secondary hypotheses were that Model B and higher-numbered variable models do not provide clinically meaningful gains in OOF performance or decision-curve net benefit over Model C once complexity penalties are considered.

Decision Rules, Inference, and Multiplicity

For each contrast we estimated Δ MCC and Δ AUROC using paired, fold-level bootstrap (10,000 resamples); 2-sided bootstrap p -values were computed, with BH control applied only within the MCC family across contrasts. Parsimony/generalization superiority was judged by BIC/PSIS-LOOIC (with $|\Delta| \geq 2$ typically indicating a small but nontrivial improvement). Secondary and exploratory endpoints were interpreted descriptively without alpha allocation.

Model Fitting, Selection, and Internal Validation

Logistic models were initially fit as generalized linear mixed models with a study-level random intercept; model fit and diagnostics favored generalized linear models without the random intercept for 8/34 models, which were used for primary analyses (Table S10 in the online supplement). Primary selection used BIC; expected out-of-sample performance used PSIS-LOOIC. Discrimination used AUROC with 95% CIs (BH-controlled where applicable). We applied stratified 10-fold cross-validation $\times 1000$ repeats with within-fold standardization; thresholds were learned in training and applied to held-out folds. OOF probabilities were averaged across repeats; each model's global threshold was the mean of fold-level Youden-J thresholds. Uncertainty used paired, fold-level bootstraps; decision-curve analysis assessed standardized net benefit.

Ancillary and Rank-Based Analyses

PCA assessed latent structure/collinearity. Hierarchical partitioning decomposed McFadden's R^2 into unique and joint components for FEV₁, TLC, DLCO_{10s}, and DLNO_{10s} z-scores. Reclassification (category-free and threshold-based NRI, IDI) compared Models B–A and Models C–A. Calibration used logistic recalibration (intercept α , slope β) with bootstrap CIs and calibration plots. Sensitivity analyses included alternative operating points and leave-one-center-out checks.

To synthesize signals across metrics, we performed rank-based comparisons of MCC, AUROC, BIC, and PSIS-LOOIC across the top 10 models.

To assess whether model comparisons were robust to the weighting of evaluation metrics, we prespecified 6 ranking schemes (equal weight, weighted average, generalization

emphasis, discrimination emphasis, BIC omitted, and AUROC emphasis). For each comparator, we computed the rank difference (comparator – the best ranked model) within each scheme and then the mean Δ rank across schemes. To reflect sensitivity to weighting choice, we obtained 95% bootstrap intervals by resampling schemes ($n=6$) with replacement ($B=10,000$) and recomputing the across scheme mean. We did not resample the derived “average rank” column. Instead, intervals were based on the 6 scheme specific differences. Intervals entirely >0 indicate the comparator is consistently ranked worse than Model C across the prespecified schemes.

Statistical Software

Analyses were performed using RStudio (2025.09.0), Build 387, with R (version 4.4.2). Two-sided $p < 0.05$ was considered statistically significant. Additional information on the statistical analyses can be found in the online supplementary material in the Supplementary Methods section, and Tables S1 to S9.

Results

Participant Characteristics

Of 496 eligible individuals, 408 (85 emphysema, 323 controls) met harmonization and quality control criteria (See Table S1, Figure S1 in the online supplement). Participant characteristics are presented in Table 1. The median breath-hold time for diffusing capacity was 10 seconds (interquartile range [IQR], 9.2–10). The median DLNO_{10s}/DLCO_{10s} ratio was 4.43 (IQR 4.09–4.88) for emphysema and 4.50 (IQR 4.22–4.84) for non-emphysema ($p=0.226$). Nearly all variables in Table 1 were statistically different between the 2 groups except for pack years of smoking and the proportion of participants with pulmonary restriction. The pooled data demonstrate that 71% of the variance in DLNO_{10s} z-scores is shared with DLCO_{10s} z-scores (Figure S2 in the online supplement).

Information Criteria and Generalization

The top 17 models are ranked by their difference in BIC and LOOIC relative to the top-performing model (Figure 1). The bottom 18 models are presented in Figure S3 in the online supplement. Absolute values for BIC and LOOIC for all 34 models are presented in Table S10 in the online supplement. A compact cluster—including the 3-predictor DLNO model (Model C)—lies near $\Delta 0$ for both BIC and LOOIC. Single-index models (e.g., VA alone, DLCO alone) are markedly inferior to the multiple variable predictor models. As such, the model with the lowest BIC is presented in Table 2 with the AUROC, and its precision-recall curves are presented in Figure S4 in the online supplement.

Table 1. Participant Characteristics

Variable	Smokers With No Emphysema (n=323)	Smokers With Emphysema (n=85)
Number of Males and Females	306M, 17 F (95% M)	78M, 7 F (92% M)
Pack Years of Smoking (Interquartile Range)	8 to 44	22 to 42
Age (years)	58 (11)	61 (10) ^a
Height (cm)	177 (8)	174 (9) ^a
Weight (kg)	80.5 (9.2)	81.9 (14.3)
BMI (kg/m ²)	25.6 (2.5)	27.0 (3.9) ^a
Hb (g/dL)	14.6 (0.7)	15.1 (1.3) ^a
FEV₁ (L)	3.54 (0.54)	2.15 (0.73) ^a
FVC (L)	4.29 (0.66)	3.38 (0.73) ^a
TLC (L)	7.44 (0.95)	7.66 (1.38)
FEV₁/FVC	0.83 (0.05)	0.63 (0.13) ^a
FEV₁ z-scores (GLI equations)	−0.15 (0.69)	−2.26 (1.29) ^a
FVC z-scores (GLI equations)	−0.53 (0.64)	−1.46 (1.11) ^a
FEV₁/FVC Ratio z-scores (GLI equations)	0.69 (0.79)	−1.79 (1.45) ^a
TLC z-scores (GLI equations)	0.35 (0.68)	0.97 (1.17) ^a
RV/TLC ratio z-scores (GLI equations)	1.16 (0.59)	2.09 (0.89) ^a
DLCO_{10s} z-scores (GLI equations)	−0.24 (1.13)	−2.23 (1.54) ^a
VA_{10s} z-scores (GLI equations)	−0.13 (1.09)	−0.66 (1.68) ^a
DLNO_{10s} z-scores (SLR, Zavorsky and Cao ³¹)	−0.75 (1.05)	−2.34 (1.27) ^a
Combined DLNO_{10s}+ DLCO_{10s} z-scores (SLR, Zavorsky and Cao ³¹)	−1.52 (1.92)	−4.71 (2.38) ^a
Number and Percentage of Individuals With Obstruction (FEV ₁ /FVC < LLN ₅ and FVC > LLN ₅)	1 (0.3%)	15 (18%) ^a
Number and Percentage of Individuals With Mixed Disorders (FEV ₁ /FVC < LLN ₅ and FVC < LLN ₅)	1 (0.3%)	20 (2%) ^a
Number and Percentage of Individuals With Confirmed Restrictions as Assessed by Total Lung Capacity (TLC < LLN ₅)	0 (0%)	1(1%)
Number and Percentage of Individuals With Hyperinflation (RV/TLC z-score > ULN ₉₅)	66 (20%)	56 (66%) ^a
Number and Percentage of Individuals With a DLCO_{10s} Abnormality	46 (14%)	54 (64%) ^a
Number and Percentage of Individuals With a DLNO_{10s} < LLN₅	61 (19%)	59 (69%) ^a
Number and Percentage of Individuals With Either Obstruction, Restriction, Hyperinflation, DLCO, or DLNO < LLN₅	105 (33%)	77 (91%) ^a

Mean (SD)

^aFor all variables marked with a footnote a, $p < 0.05$. The p -values were determined after controlling for a false discovery rate of 1% (38). The combined Z-scores for diffusing capacity were calculated using segmented (piecewise) linear regression (SLR) of Zavorsky and Cao.³¹

BMI=body mass index; Hb=hemoglobin; FEV₁=forced expiratory volume in 1 second; FVC=forced vital capacity; TLC=total lung capacity; RV=residual volume; DLCO=diffusing capacity for carbon monoxide; VA=alveolar volume; DLNO=diffusing capacity for nitric oxide; SLR=segmented linear regression; LLN₅=lower limit of normal defined as the 5th percentile; ULN₉₅=upper limit of normal defined as the 95th percentile

Discrimination Area Under the Receiving Operating Characteristic and Matthews Correlation Coefficient at the Youden Cut-Point

Figure 2 displays MCC (bars) and AUROC (points with 95% CIs) ranked by MCC at the Youden threshold for the top 17 models. The bottom 17 models are presented in Figure S5 in the online supplement. Three- to 6-predictor DLNO models achieve AUROC ≈ 0.96 – 0.97 with MCC ≈ 0.8 . Model A (DLCO-based) is modestly worse, and Model B (A+DLNO) trades sensitivity and specificity without clear net gain (for fold-averaged metrics and CIs).

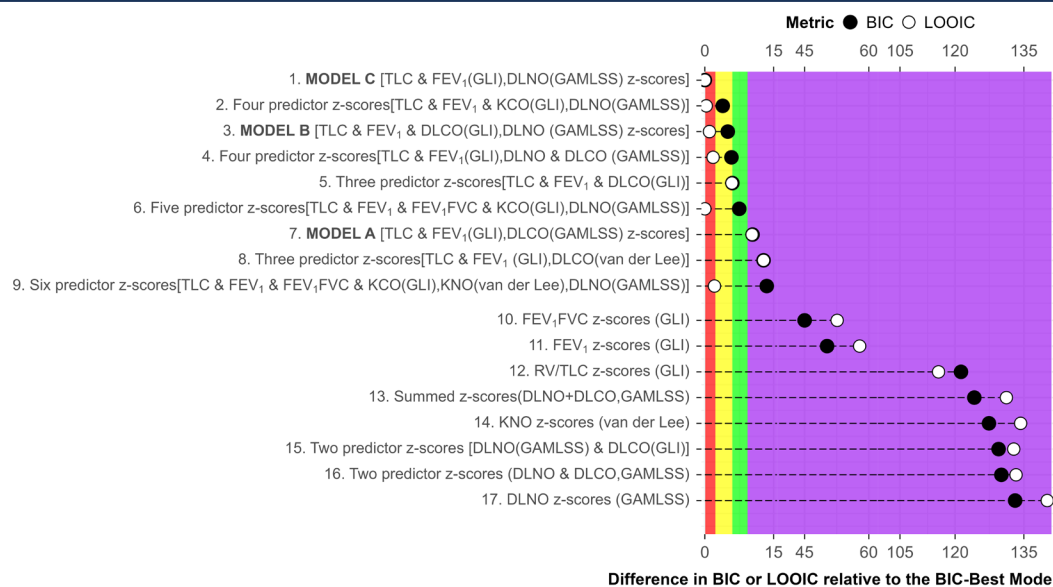
Principal Component Analysis and Hierarchical Partitioning

PCA revealed principal component (PC) 1 (gas transfer: DLNO/DLCO), PC2 (hyperinflation: TLC/VA), and

PC3 (obstruction/air-trapping: FEV₁/FVC, RV/TLC) (Tables S11–S13 in the online supplement). PC4 was not found to add any benefit (Table S14, Figure S6 in the online supplement). Replacing z-score predictors with PC1–PC3 did not improve discrimination or generalization (Δ AUROC ≈ 0 – 0.01 ; Δ LOOIC/ Δ BIC < 2); hence, we favor Model C (TLC+FEV₁+DLNO z-scores) specification for interpretability. Hierarchical partitioning ranked unique contributions as the FEV₁ z-scores > DLNO_{10s} z-scores > TLC z-scores > DLCO_{10s} z-scores (Tables S15–S16 in the online supplement).

Classification, Reclassification and Decision Analysis

Model B z-scores (FEV₁, TLC, DLNO_{10s}, DLCO_{10s} z-scores) or Model C (z-scores FEV₁, TLC, DLNO_{10s}, z-scores) compared

Figure 1. Bayesian Information Criterion and Leave-One-Out Information Criterion of Models for Emphysema Prediction and Generalizability

Models are ranked by their difference in BIC (black circles) and LOOIC (white circles) relative to the top-performing model (top=best). BIC penalizes model complexity; LOOIC evaluates predictive performance via cross-validation. Colored zones indicate performance tiers:

Red zone (BIC or LOOIC difference ≤ 2.2): Models nearly as good as the best model.

Yellow zone (BIC or LOOIC difference 2.3–5.9): Models with substantial but acceptable performance differences compared to the best model.

Green zone (BIC or LOOIC difference 6.0–9.2): Models with considerably weaker performance compared to the best model.

Purple zone (BIC or LOOIC difference ≥ 9.3): Models with significantly poorer fit compared to the best model.

The x-axis shows the difference from the best model—smaller is better. The best-performing model is the 3-predictor z-score model of TLC + FEV₁ + DLNO_{10s} (GAMLSS) derived from the GLI equations^{27,28} and DLNO z-scores from the GAMLSS equations³¹ (n=323 smokers with and without emphysema; n=85 smokers with emphysema).

BIC=Bayesian information criterion FEV₁=forced expiratory volume in 1 second; GLI=Global Lung Function Initiative; DLNO=diffusing capacity for nitric oxide; GAMLSS=generalized additive models of location, scale, and shape; TLC=total lung capacity; KCO=carbon monoxide transfer coefficient; DLCO=diffusing capacity for nitric oxide RV=residual volume; KNO=nitric oxide transfer coefficient; LOOIC=leave-one-out information criteria

to Model A (FEV₁, TLC, DLCO_{10s} z-scores) demonstrated no real difference in 17 metrics when considering the 95% CI (Tables S17–S18 in the online supplement). At Youden-optimized thresholds, the overall net improvement in reclassification when DLNO_{10s} is added to the Model A was not significant. However, the average predicted risk-gap between predicting smokers with and without emphysema improved by as much as 5% when DLNO_{10s} is added to Model A (Table S19 in the online supplement).

At category-free reclassification, there was a 34% overall net improvement in reclassification (95% CI=–12 to 96%) when DLNO_{10s} was added to Model A. Simply, this means that, compared with the old model, the new model moved people in the right direction (up for true cases, down for true noncases) 34 percentage points more often than it moved them in the wrong direction. Moreover, the average predicted risk-gap between predicting smokers with and without emphysema improved by as much as 5% when DLNO_{10s} was added to Model A (Table S20 in the online supplement).

Decision-curve analysis using out-of-fold predictions (Figure S7 in the online supplement) showed all 3 models (Models A, B, C) delivered positive net benefit across threshold probabilities 0–0.25, exceeding Treat None and—apart from the very lowest thresholds—exceeding

Treat All. The Treat-All curve crossed zero at ~0.21 (cohort prevalence), while all model curves remained positive. The e curves overlapped closely; the 4-predictor model (Model B) offered no discernible advantage over either Model A or C. Absolute net benefit was ~0.16–0.20, i.e., ~16–20 more correctly flagged smokers with emphysema per 100 smokers than doing nothing, at thresholds 0–0.25

Model Rankings

Across schemes, 8 out of 9 comparators ranked significantly worse than the best ranked model (Model C, Δ rank >0 ; 95% CI is >0). The 5-predictor DLNO model was the only top 10 model that had a ranking that was not different to Model C. (Δ =+0.22 [–0.09, 0.57]). As such, Model C (TLC z-scores, FEV₁ z-scores, and DLNO_{10s} z-scores) offered near-top performance yet delivered comparable discrimination/generalization with fewer predictors with greater parsimony than any other model. Thus, Model C is the best choice (Figure 3).

Table 2. Binary Logistic Regression Results of the Model With the Best Performance as Determined by the Lowest Bayesian Information Criterion

Predictors Fixed Effects	Estimate [95% CI]	SE	z-value	VIF	McFadden's Pseudo R ² [Overall=0.66]	Nagelkerke R ² [Overall=0.77]	Odds Ratios [95%CI]
Intercept	−5.27 [−6.33, −4.33]	0.51	−10.30	–	–	–	0.01 [0.00, 0.02]
FEV ₁ z-scores	−1.43 [−1.85, −1.03]	0.21	−6.87	1.15	0.52	0.64	0.24 [0.16, 0.35]
TLC z- scores	1.55 [1.05, 2.08]	0.26	5.92	1.40	0.08	0.13	4.73 [2.87, 8.07]
DLNO _{10s} z-scores (GAMLSS)	−0.73 [−1.02, −0.46]	0.14	−5.10	1.45	0.32	0.43	0.49 [0.36, 0.64]

The z-scores of TLC, FEV₁, and DLNO_{10s} (GAMLSS) best predict emphysema.

We compared a 3-predictor logistic model with and without a study-level random intercept. The model without the study-level random intercept had slightly higher expected out-of-sample accuracy [the change in expected log predictive density (ΔELPD)=0.22±0.27 is within 1 SE, stacking weight=1.00]. Therefore, fixed-effects model without a study random intercept was used. Null deviance: 417.6 on 407 degrees of freedom. Residual deviance: 140.6 on 404 degrees of freedom. AIC (null)=419.6; BIC (null)=423.6. AIC for the 3-predictor model =148.6; BIC for the 3-predictor model=164.6. McFadden's pseudo-R² is a measure that assesses how well the model fits the data relative to a baseline (null) model of no predictors. It is not the same as the R² in linear regression. McFadden's pseudo-R² is a scaled measure of improvement in log-likelihood by comparing the model's log-likelihood with the 3 predictors to the log-likelihood of a null model with no predictors, just the intercept. Overall, McFadden's pseudo R²=0.66 suggests that TLC, FEV₁, and DLNO_{10s} z-scores explain a significant portion of the log-likelihood compared to the null model, suggesting a strong fit. However, McFadden's R² does not directly quantify the percentage of variance in emphysema explained by those specific predictors. However, the Craig-Uhler R² (Nagelkerke R²) was 0.77, adjusted for the binary outcome, indicating that the 3-predictor model explained 77% of the explainable variance.

Note: The overlapping McFadden pseudo-R² occurs because the predictors are correlated. For example, TLC z-scores and FEV₁ z-scores have related aspects of lung function sharing variance explaining disease. The TLC z-scores and FEV₁ z-scores are obtained from GLI equations,^{27,28} while the DLNO_{10s} z-scores were obtained using GAMLSS equations of Zavorsky and Cao.³¹ These specific reference equations presented here were a better fit than DLNO reference equations, be they segmented linear regression reference equations³¹ or otherwise.³⁰ Furthermore, DLCO_{10s} z-scores obtained from any reference equation, including GLI equations,^{29,30} displayed inferior fitting compared to DLNO_{10s}, and removing it resulted in a negligible loss of 0.2-0.3% in pseudo-R² and a slight increase in AIC (from 148.6 to 149.5) and a BIC (from 164.6 to 169.6). This implies that DLCO_{10s} add little unique predictive value for emphysema in this dataset. Thus, the final 3-predictor models are included without DLCO_{10s}. In this analysis, there were 85 smokers with emphysema and 323 smokers without emphysema.

SE=standard error; VIF=variance inflation factor; CI=confidence interval; FEV₁=forced expiratory volume in 1 second; TLC=total lung capacity; DLNO=diffusing capacity for nitric oxide; GAMLSS=generalized additive models of location, scale, and shape; ELPD=expected log pointwise predictive density; AIC=Akaike information criterion; BIC=Bayesian information criterion; DLCO=diffusing capacity for carbon monoxide

Discussion

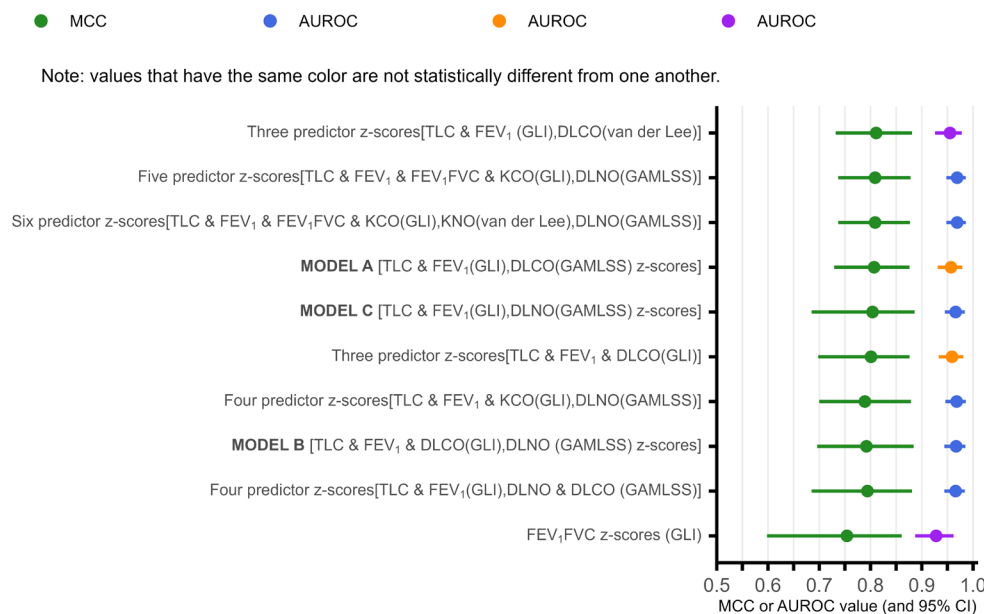
Our multicenter IPD meta-analysis demonstrates that compact models incorporating DLNO offer robust, high-quality classification of emphysema compared to smokers without emphysema. In the 3-predictor DLNO z-score model (Model C: TLC z-scores, FEV₁ z-scores, and DLNO_{10s} z-scores), the TLC z-scores and FEV₁ z-scores were obtained using GLI equation for White participants,^{27,28} while the best fitting DLNO z-scores were obtained by using the DLNO GAMLSS equation from Zavorsky and Cao.³¹ The 3-predictor DLNO z-score model (Model C) occupied a consistently superior or cosuperior position across BIC/LOOIC (Figure 1) and AUROC/MCC (Figure 2) and maintaining top ranks across weighting scenarios (Figure 3).

The incremental advantage in Model C (TLC z-scores, FEV₁ z-scores, and DLNO_{10s} z-scores) versus a DLCO-based analogue (Model A) is modest, but consistency across metrics and resampling supports a genuine performance edge for Model C. Reclassification indices were small (Tables S19–S20 in the online supplement), suggesting DLNO's benefits manifest more as improved overall performance and calibration (Figure S7 in the online supplement) than as wholesale shifts in categorical assignment at a single threshold.

The pooled data demonstrate that 71% of the variance in DLNO_{10s} z-scores is shared with DLCO_{10s} z-scores (Figure S2 in the online supplement) displaying substantial collinearity; yet the residual ~29% “unique” variance is not necessarily predictive for emphysema. Nevertheless, adding DLCO z-scores to Model C, model barely changes performance (McFadden's R²=0.663→0.666; and BIC worsens), and AUROC gains were negligible. That is strong evidence that the “unique” DLCO portion does not add a meaningful predictive signal for emphysema beyond the z-score model of TLC+FEV₁+DLNO_{10s} z-scores.

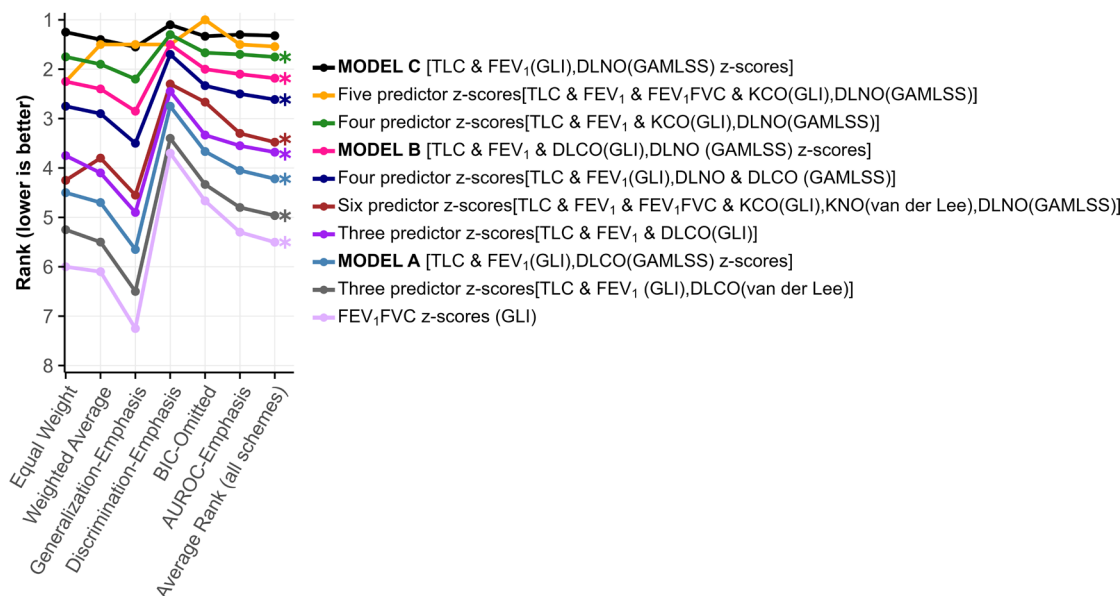
Physiologic plausibility of the results is strong. One view is that NO uptake primarily reflects membrane resistance, whereas CO uptake reflects the resistance that occurs within the red cell membrane (Zavorsky et al Figure 1).⁷ The finding that adding DLCO z-scores to a DLNO z-score-based model contributes little, coheres with early membrane-dominant injury in emphysema. PCA structure (Tables S11–S14 in the online supplement) and hierarchical partitioning (Table S15–S16 in the online supplement) further support construct validity by aligning dominant components with expected physiologic domains.

In our IPD meta-analysis, Model C had similar discrimination and classification of the 3-predictor model using TLC, FEV₁, and DLCO_{10s} z-scores, while achieving

Figure 2. The Discriminatory Classification Performance of the Top 10 Predictive Models

Models are ordered by Matthews correlation coefficient (MCC; green points with horizontal 95% CIs). Corresponding AUROC values are shown to the right as colored points with 95% CIs (blue, orange, and purple). Colors denote groups of models whose AUROC values are not statistically different from one another (same color = no significant difference, different colors = $p < 0.05$, Benjamini-Hochberg FDR 1%), based on 100,000 bootstrap samples. The top model is Model C, which achieves the highest MCC and AUROC; lower-ranked models show progressively weaker discrimination.

MCC=Matthews correlation coefficient; AUROC=area under the receiver operating characteristic; TLC=total lung capacity; FEV₁=forced expiratory volume in 1 second; GLI=Global Lung Function Initiative; DLCO=diffusing capacity for carbon monoxide; FVC=forced vital capacity; KCO=carbon monoxide transfer coefficient; DLNO=diffusing capacity for nitric oxide; GAMLSS=generalized additive models of location, scale, and shape; KNO=nitric oxide transfer coefficient; CI=confidence interval; FDR=false discovery rate

Figure 3. Weight Sensitivity: Model Rankings Across Weighted Schemes

For each of the 6 ranking schemes (equal weight, weighted average, generalization-emphasis, discrimination-emphasis, BIC-omitted, AUROC-emphasis), we computed a per-scheme rank difference (Δ rank) between each comparator and the baseline MODEL C [TLC & FEV₁(GLI), DLNO(GAMLSS) z-scores] (Δ rank = rank_{comparator} - rank_{MODEL C}; positive=worse). We then took the mean Δ rank across the 6 schemes for each comparator and obtained 95% CIs via nonparametric bootstrap by resampling schemes with replacement ($B=10,000$). We computed per-scheme rank differences vs MODEL C for the 6 schemes, then took the mean of those 6 differences for each comparator, and then bootstrapped CIs by resampling schemes ($n=6$) with replacement across-scheme mean Δ rank vs MODEL C (95% CI). Eight comparators have significantly higher ranking numbers (higher=worse) than MODEL C (CIs entirely > 0): FEV₁/FVC z-scores (GLI): +4.18 [3.36, 4.98]; Three predictor z-scores [TLC & FEV₁ (GLI), DLCO (van der Lee)]: +3.64 [2.98, 4.33]; MODEL A [TLC & FEV₁ (GLI), DLCO (GAMLSS)]: +2.90 [2.28, 3.53]; Three predictor z-scores [TLC & FEV₁ & DLCO (GLI)]: +2.36 [1.88, 2.85]; Six predictor z-scores [TLC & FEV₁ & FEV₁/FVC & KCO (GLI), KNO (van der Lee), DLNO (GAMLSS)]: +2.16 [1.57, 2.72]; Four predictor z-scores [TLC & FEV₁ (GLI), DLNO & DLCO (GAMLSS)]: +1.29 [0.95, 1.64]; MODEL B [TLC & FEV₁ & DLCO (GLI), DLNO (GAMLSS)]: +0.86 [0.63, 1.08]; Four predictor z-scores [TLC & FEV₁ & KCO (GLI), DLNO (GAMLSS)]: +0.43 [0.32, 0.55]. The 5-predictor DLNO model is not significantly different from MODEL C (CI includes 0); However, due to the desire of having a parsimonious model, MODEL C wins.

TLC=total lung capacity; FEV₁=forced expiratory volume in 1 second; GLI=Global Lung Function Initiative; DLNO=diffusing capacity for nitric oxide; GAMLSS=generalized additive models of location, scale, and shape; FVC=forced vital capacity; KCO=carbon monoxide transfer coefficient; DLCO=diffusing capacity for carbon monoxide; BIC=Bayesian information criterion; AUROC=area under the receiving operating characteristic

lower BIC and LOOIC—indicating a more parsimonious specification with better expected out-of-sample fit. Practically, this suggests DLNO_{10s} can substitute for DLCO_{10s} as the gas-transfer input in emphysema when simultaneous NO–CO testing is available, and quality control is assured. We are not advocating DLNO_{10s} in isolation from spirometry and lung volumes across all pathologies – summed DLNO+DLCO z-score models can outperform either measure alone in other conditions.^{19,7} Rather, among transfer measures, DLNO_{10s} alone (without DLCO_{10s}) appears sufficient in this emphysema-screening context. DLCO_{10s} still has broader clinical roles (e.g., interstitial lung^{44,45} and pulmonary vascular disease⁴⁶) so laboratories without NO–CO capability can continue to rely on DLCO_{10s}, whereas centers with NO–CO may reasonably prioritize DLNO_{10s} or DLNO_{5s}, where validated, in parsimonious models. Decision-curve analyses indicate potential utility at low thresholds common to screening/case-finding (Figure S5 in the online supplement).

Despite promising diagnostic performance, DLNO remains underutilized since its introduction^{8,9} in 1983–84, largely due to limited awareness, regulatory hurdles, and the high cost of sensitive NO analyzers. DLNO testing is available as an add-on to DLCO on at least one commercial platform (e.g., MGC Diagnostics; St. Paul, Minnesota), requiring only an NO sensor, cylinder, and minimal training. These systems typically use low-cost electrochemical sensors (e.g., 7NT CiTiceL[®]), but their limited range (0–100 ppm) and slow response (~15 second) make them unsuitable for longer breath-holds where exhaled NO concentrations drop below detectable thresholds. Chemiluminescence remains the gold standard for NO detection due to its rapid response (<1 second) and wide dynamic range (1 ppb–100 ppm), but devices like the CLD 855 Yh (Eco Physics; Ann Arbor, Michigan) cost ~\$35,000 USD. Mass production could lower this to \$10,000–\$21,000 per unit if adopted across the estimated 2000–5000 DLCO-equipped sites in the United States.⁴⁷

Moreover, the similar median DLNO_{10s}/DLCO_{10s} ratios between emphysema and non-emphysema groups (4.43 versus 4.50) suggest that raw ratios lack discriminatory power—reinforcing the need for z-score standardization in diagnostic models. Future work should prioritize affordable, regulatory-approved DLNO_{10s} systems and longitudinal studies to assess their clinical utility.

Limitations

Our pooled analytic sample comprised primarily White adults who were current or former smokers from European centers using a 10±2-second simultaneous NO–CO protocol; only ~14% were never-smokers. Accordingly, generalizability to never-smokers, other ancestral groups, pediatric or very elderly populations, and to laboratories

using different devices or protocols may be limited. CT-confirmed emphysema improves case specificity but may miss early/subclinical or airway-predominant disease. Although we harmonized variables across sites, applied uniform quality control, and used repeated cross-validation with leave-one-center-out checks and calibration assessment, we lacked an independent external cohort; thus, performance estimates and optimal thresholds may shift in other settings. Furthermore, it is known that genetic ancestry affects DLNO,^{41–43} so standardization relying on DLNO reference equations and GLI equations for TLC and DLCO derived largely from White cohorts, can constrain calibration and promote bias. Reference equations developed for specific genetic ancestries (versus only White) are needed to maintain precision. We also lacked uniform data on comorbidities, medications, and socioeconomic context, limiting adjustment for potential confounding and spectrum effects. Finally, the dataset did not include longitudinal outcomes (e.g., exacerbations, CT progression, mortality), so we evaluated discrimination and reclassification rather than long-term clinical impact; strict and complete-case analysis may also introduce selection bias, and residual site/device effects may persist despite adjustment.

Broader Implications for Diffusing Capacity for Nitric Oxide Adoption in Clinical Practice

The demonstrated superiority of DLNO_{10s} z-scores in predicting and classifying smokers with emphysema compared to smokers without emphysema offers a compelling case for DLNO inclusion in routine pulmonary function testing. Despite the current barriers to widespread adoption—such as the high cost of NO analyzers and regulatory hurdles—this study provides strong evidence for the clinical value of DLNO. Its ability to detect subtle changes in alveolar-capillary membrane functionality positions it as a critical tool for early COPD diagnosis, particularly in high-risk populations such as smokers.

Future efforts should focus on developing cost-effective, FDA-approved devices for DLNO measurement to facilitate its integration into pulmonary function laboratories. Additionally, prospective studies validating these findings in diverse populations are necessary to further establish DLNO's role as a diagnostic benchmark.

Conclusion

The 3-predictor model incorporating DLNO_{10s} z-scores with TLC z-scores and FEV₁ z-scores offers superior model performance, predictive accuracy, and classification for emphysema detection compared to DLCO_{10s}-based models. These findings advocate integrating DLNO_{10s} into routine clinical practice, potentially improving early diagnosis and patient outcomes in emphysema management.

Acknowledgements

Author contributions: GSZ was responsible for the conceptual design of the work, data curation, formal analysis, methodology, software, validation, visualization, and original and final draft preparation. RWD-N, IvdL, and AMP provided the study participants and data collection and reviewed and edited the manuscript. All authors reviewed and approved the final manuscript submitted for publication.

Data sharing statement: The data supporting this study's findings are accessible on Mendeley Data,³⁶ a cloud-based repository for research data.

Declaration of Interest

GSZ is a GLI Network member. The GLI Network has published reference equations for spirometry, DLCO, and static lung volumes using GAMLSS models. GSZ is the current cochair of the European Respiratory Society Task Force on the interpretation of DLNO. The remaining authors declare that no conflicts of interest exist.

References

1. American Lung Association. Trends in emphysema by race/ethnicity, sex, and age. American Lung Association website. Accessed February 25, 2025. <https://www.lung.org/research/trends-in-lung-disease/copd-trends-brief/data-tables/trends-in-emphysema-by-race-sex-age>
2. Saeed H, Arshad MK, Shahnoor S, et al. Temporal trends, gender, and ethnoracial disparities in mortality from pulmonary emphysema: a retrospective nationwide analysis. *Medicine (Baltimore)*. 2024;103(52):e41032. <https://doi.org/10.1097/MD.00000000000041032>
3. Himes BE, Dai Y, Kohane IS, Weiss ST, Ramoni MF. Prediction of chronic obstructive pulmonary disease (COPD) in asthma patients using electronic medical records. *J Am Med Inform Assoc*. 2009;16(3):371-379. <https://doi.org/10.1197/jamia.M2846>
4. Adatia A, Wahab M, Shahid I, Moinuddin A, Killian KJ, Satia I. Effects of cigarette smoke exposure on pulmonary physiology, muscle strength and exercise capacity in a retrospective cohort with 30,000 subjects. *PLoS One*. 2021;16(6):e0250957. <https://doi.org/10.1371/journal.pone.0250957>
5. Hoffmann RF, Zarrintan S, Brandenburg SM, et al. Prolonged cigarette smoke exposure alters mitochondrial structure and function in airway epithelial cells. *Respir Res*. 2013;14(1):97. <https://respiratory-research.biomedcentral.com/articles/10.1186/1465-9921-14-97>
6. Calverley PMA, Walker P. Chronic obstructive pulmonary disease. *Lancet*. 2003;362(9389):1053-1061. [https://doi.org/10.1016/S0140-6736\(03\)14416-9](https://doi.org/10.1016/S0140-6736(03)14416-9)
7. Zavorsky GS, Barisione G, Gille T, et al. Enhanced detection of patients with previous COVID-19: Superiority of the double diffusion technique. *BMJ Open Respir Res*. 2025;12(1):e002561. <https://doi.org/10.1136/bmjresp-2024-002561>
8. Borland C, Chamberlain A, Higenbottam T. The fate of inhaled nitric oxide. *Clin Sci (Lond)*. 1983;65(3):37P. <https://doi.org/10.1042/cs065037P>
9. Borland C, Cracknell N, Higenbottam T. Is the measurement of "DLNO" a true measure of membrane diffusing capacity? *Clin Sci (Lond)*. 1984;67(S9):41P. <https://doi.org/10.1042/cs067041Pa>
10. Zavorsky GS, Hsia CCW, Hughes JMB, et al. Standardisation and application of the single-breath determination of nitric oxide uptake in the lung. *Eur Respir J*. 2017;49(2):1600962. <https://doi.org/10.1183/13993003.00962-2016>
11. Roughton FJW, Forster RE. Relative importance of diffusion and chemical reaction rates in determining rate of exchange of gases in the human lung, with special reference to true diffusing capacity of pulmonary membrane and volume of blood in the lung capillaries. *J Appl Physiol*. 1957;11(2):290-302. <https://doi.org/10.1152/jappl.1957.11.2.290>
12. Borland CDR, Hughes JMB. Lung diffusing capacities (D(L)) for nitric oxide (NO) and carbon monoxide (CO): the evolving story. *Compr Physiol*. 2020;10(1):73-97. <https://doi.org/10.1002/cphy.c190001>
13. Borland C, Patel R. Comparing in vitro nitric oxide blood uptake to its pulmonary diffusing capacity. *Nitric Oxide*. 2024;143:29-43. <https://doi.org/10.1016/j.niox.2023.11.006>
14. Kang MY, Sapoval B. Time-based understanding of DLCO and DLNO. *Respir Physiol Neurobiol*. 2016;225:48-59. <https://doi.org/10.1016/j.resp.2016.01.008>
15. Kang MY, Grebenkov D, Guenard H, Katz I, Sapoval B. The Roughton-Forster equation for DLCO and DLNO re-examined. *Respir Physiol Neurobiol*. 2017;241:62-71. <https://doi.org/10.1016/j.resp.2016.12.014>
16. Giraud X, Le-Dong NN, Hogben K, Martinot JB. The measurement of DLNO and DLCO: a manufacturer's perspective. *Respir Physiol Neurobiol*. 2017;241:36-44. <https://doi.org/10.1016/j.resp.2017.02.006>
17. Zavorsky GS, van der Lee I. Can the measurement of pulmonary diffusing capacity for nitric oxide replace the measurement of pulmonary diffusing capacity for carbon monoxide? *Respir Physiol Neurobiol*. 2017;241:9-16. <https://doi.org/10.1016/j.resp.2016.11.008>
18. van der Lee I, Gietema HA, Zanen P, et al. Nitric oxide diffusing capacity versus spirometry in the early diagnosis of emphysema in smokers. *Respir Med*. 2009;103(12):1892-1897. <https://doi.org/10.1016/j.rmed.2009.06.005>
19. Zavorsky GS, Agostoni P. Two is better than one: the double diffusion technique in classifying heart failure. *ERJ Open Res*. 2024;10(1):00644-2023. <https://doi.org/10.1183/23120541.00644-2023>
20. Dressel H, Filser L, Fischer R, et al. Lung diffusing capacity for nitric oxide and carbon monoxide in relation to morphological changes as assessed by computed tomography in patients with cystic fibrosis. *BMC Pulm Med*. 2009;9:30. <https://doi.org/10.1186/1471-2466-9-30>
21. Barisione G, Brusasco C, Garlaschi A, Baroffio M, Brusasco V. Lung diffusing capacity for nitric oxide as a marker of fibrotic changes in idiopathic interstitial pneumonias. *J Appl Physiol*. 2016;120(9):1029-1038. <https://doi.org/10.1152/jappphysiol.00964.2015>
22. Lin CH, Cheng SL, Chen CZ, Chen CH, Lin SH, Wang HC. Current progress of COPD early detection: key points and novel strategies. *Int J Chron Obstruct Pulmon Dis*. 2023;18:1511-1524. <https://doi.org/10.2147/COPD.S413969>
23. Aaron SD, Montes de Oca M, Celli B, et al. Early diagnosis and treatment of chronic obstructive pulmonary disease: the costs and benefits of case finding. *Am J Respir Crit Care Med*. 2024;209(8):928-937. <https://doi.org/10.1164/rccm.202311-2120PP>
24. Postolache P, Nemes RM, Petrescu O, Merisanu IO. Smoking cessation, pulmonary rehabilitation and quality of life at smokers with COPD. *Rev Med Chir Soc Med Nat Iasi*. 2015 Jan-Mar;119(1):77-80. <https://www.revmedchir.ro/index.php/revmedchir/article/view/494/435>

25. Rochester CL, Alison JA, Carlin B, et al. Pulmonary rehabilitation for adults with chronic respiratory disease: an official American Thoracic Society clinical practice guideline. *Am J Respir Crit Care Med*. 2023;208(4):e7-e26. <https://doi.org/10.1164/rccm.202306-1066ST>
26. Mendes Xavier D, Lanza Galvao E, Aliane Fonseca A, de Souza GM, Pereira Lima V. Effects of home-based pulmonary rehabilitation on dyspnea, exercise capacity, quality of life and impact of the disease in COPD patients: a systematic review. *COPD*. 2022;19(1):18-46. <https://doi.org/10.1080/15412555.2021.2020234>
27. Hall GL, Filipow N, Ruppel G, et al. Official ERS technical standard: Global Lung Function Initiative reference values for static lung volumes in individuals of European ancestry. *Eur Respir J*. 2021;57(3):2000289. <https://doi.org/10.1183/13993003.00289-2020>
28. Quanjer PH, Stanojevic S, Cole TJ, et al. 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):1324-1343. <https://doi.org/10.1183/09031936.00080312>
29. Stanojevic S, Graham BL, Cooper BG, et al. Official ERS technical standards: Global Lung Function Initiative reference values for the carbon monoxide transfer factor for Caucasians [published correction appears in *Eur Respir J*. 2020;56:1700010. <https://doi.org/10.1183/13993003.00010-2017>] *Eur Respir J*. 2017;50(3):1700010. <https://doi.org/10.1183/13993003.00010-2017>
30. van der Lee I, Zanen P, Stigter N, van den Bosch JM, Lammers JW. Diffusing capacity for nitric oxide: reference values and dependence on alveolar volume. *Respir Med*. 2007;101(7):1579-1584. <https://doi.org/10.1016/j.rmed.2006.12.001>
31. Zavorsky GS, Cao J. Reference equations for pulmonary diffusing capacity using segmented regression show similar predictive accuracy as GAMLSS models. *BMJ Open Respir Res*. 2022;9(1):e001087. <https://doi.org/10.1136/bmjresp-2021-001087>
32. Stewart LA, Cochrane Working Group On Meta-Analysis Using Individual Patient Data. Practical methodology of meta-analyses (overviews) using updated individual patient data. *Stat Med*. 1995;14(19):2057-2079. <https://doi.org/10.1002/sim.4780141902>
33. Moinard J, Guénard H. Determination of lung capillary blood volume and membrane diffusing capacity in patients with COLD using the NO-CO method. *Eur Respir J*. 1990;3(3):318-322. <https://doi.org/10.1183/09031936.93.03030318>
34. Diener L, Herold R, Harth V, Preisser A. DLNO/DLCO in different patient collectives; influencing factors have different effects on DLNO and DLCO [in German] *Pneumologie*. 2021;75(S01):S3. <https://doi.org/10.1055/s-0041-1723272>
35. Dal Negro RW, Turco P, Povero M. Single-breath simultaneous measurement of DLNO and DLCO as predictor of the emphysema component in COPD - a retrospective observational study. *Int J Chron Obstruct Pulmon Dis*. 2024;19:2123-2133. <https://doi.org/10.2147/COPD.S467138>
36. Zavorsky GS, Dal-Negro R, van der Lee I, Preisser A. Data from: Nitric oxide uptake in the lungs in smokers with emphysema. Mendeley Data, V2, 2025. Deposited February 26, 2025. <https://data.mendeley.com/datasets/g4jp8wd6f6/2>
37. 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. <https://doi.org/10.1164/rccm.201908-1590ST>
38. Wanger J, Clausen JL, Coates A, et al. Standardisation of the measurement of lung volumes. *Eur Respir J*. 2005;26(3):511-522. <https://doi.org/10.1183/09031936.05.00035005>
39. Macintyre N, Crapo RO, Viegi G, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J*. 2005;26(4):720-735. <https://doi.org/10.1183/09031936.05.00034905>
40. Radtke T, de Groot Q, Haile SR, Maggi M, Hsia CCW, Dressel H. Lung diffusing capacity for nitric oxide measured by two commercial devices: a randomised crossover comparison in healthy adults. *ERJ Open Res*. 2021;7(3):00193-2021. <https://doi.org/10.1183/23120541.00193-2021>
41. Zavorsky GS, Almamary AS, Alqahtani MK, Shan SHS, Gardenhire DS. The need for race-specific reference equations for pulmonary diffusing capacity for nitric oxide. *BMC Pulm Med*. 2021;21:232. <https://doi.org/10.1186/s12890-021-01591-7>
42. Gochicoa-Rangel L, De-Los-Santos-Martinez A, Reyes-Garcia A, et al. Reference equations for DLNO and DLCO in Mexican Hispanics: influence of altitude and race. *BMJ Open Respir Res*. 2024;11(1):2002341. <https://doi.org/10.1136/bmjresp-2024-002341>
43. Simaga B, Forton K, Motoji Y, Naeije R, Faoro V. Lung diffusing capacity in sub-Saharan Africans versus European Caucasians. *Respir Physiol Neurobiol*. 2017;241:23-27. <https://doi.org/10.1016/j.resp.2017.01.003>
44. Alhamad EH, Cal JG. Predictors of mortality in interstitial lung disease patients without pulmonary hypertension. *Ann Thorac Med*. 2020;15(4):238-243. https://doi.org/10.4103/atm.ATM_438_20
45. Wong AW, Ryerson CJ, Guler SA. Progression of fibrosing interstitial lung disease. *Respir Res*. 2020;21:32. <https://doi.org/10.1186/s12931-020-1296-3>
46. Stadler S, Mergenthaler N, Lange TJ. The prognostic value of DLCO and pulmonary blood flow in patients with pulmonary hypertension. *Pulm Circ*. 2019;9(4):1-9. <https://doi.org/10.1177/2045894019894531>
47. Kearsley P. Understanding medical device manufacturing costs: key factors influencing medical device manufacturing costs. Meridian Medical website. Published February 2024. Accessed September 2025. <https://www.meridian-medical.com/key-factors-influencing-medical-device-manufacturing-costs/>