Article

Determinants of Pulmonary Emphysema Severity in Taiwanese Patients with Chronic Obstructive Pulmonary Disease: An Integrated Epigenomic and Air Pollutant Analysis

Sheng-Ming Wu [1,2], Wei-Lun Sun [1,2], Kang-Yun Lee [1,2,3], Cheng-Wei Lin [4,5,6], Po-Hao Feng [1,2], Hsiao-Chi Chuang [1,2,7], Shu-Chuan Ho [1,7], Kuan-Yuan Chen [1,3,6], Tzu-Tao Chen [1,2,6], Wen-Te Liu [1,2,7], Chien-Hua Tseng [1,2,8,*] and Oluwaseun Adebayo Bamodu [9,10,11,*]

[1] Division of Pulmonary Medicine, Department of Internal Medicine, Shuang Ho Hospital, Taipei Medical University, New Taipei City 235, Taiwan; chitosan@tmu.edu.tw (S.-M.W.); weilun.sun@tmu.edu.tw (W.-L.S.); leekangyun@tmu.edu.tw (K.-Y.L.); fengpohao@tmu.edu.tw (P.-H.F.); chuanghc@tmu.edu.tw (H.-C.C.); shu-chuan@tmu.edu.tw (S.-C.H.); 14388@s.tmu.edu.tw (K.-Y.C.); 09330@s.tmu.edu.tw (T.-T.C.); b7801077@tmu.edu.tw (W.-T.L.)
[2] Division of Pulmonary Medicine, Department of Internal Medicine, School of Medicine, College of Medicine, Taipei Medical University, Taipei 110, Taiwan
[3] Graduate Institute of Clinical Medicine, College of Medicine, Taipei Medical University, Taipei 110, Taiwan
[4] Department of Biochemistry and Molecular Cell Biology, Taipei Medical University, Taipei 110, Taiwan; cwlin@tmu.edu.tw
[5] International PhD Program for Cell Therapy and Regeneration Medicine, Taipei Medical University, Taipei 110, Taiwan
[6] TMU Research Center of Thoracic Medicine, Taipei Medical University, Taipei 110, Taiwan
[7] School of Respiratory Therapy, College of Medicine, Taipei Medical University, Taipei 110, Taiwan
[8] Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei 106, Taiwan
[9] Department of Urology, Shuang Ho Hospital, Taipei Medical University, New Taipei City 235, Taiwan
[10] Division of Hematology and Oncology, Department of Internal Medicine, Shuang Ho Hospital, Taipei Medical University, New Taipei City 235, Taiwan
[11] Department of Medical Research, Shuang Ho Hospital, Taipei Medical University, New Taipei City 235, Taiwan
* Correspondence: chtseng0925@tmu.edu.tw (C.-H.T.); 16625@s.tmu.edu.tw or dr_bamodu@yahoo.com (O.A.B.); Tel.: +886-02-22490088 (C.-H.T. & O.A.B.)

Abstract: Background: Chronic obstructive pulmonary disease (COPD) continues to pose a therapeutic challenge. This may be connected with its nosological heterogeneity, broad symptomatology spectrum, varying disease course, and therapy response. The last three decades has been characterized by increased understanding of the pathobiology of COPD, with associated advances in diagnostic and therapeutic modalities; however, the identification of pathognomonic biomarkers that determine disease severity, affect disease course, predict clinical outcome, and inform therapeutic strategy remains a work in progress. Objectives: Hypothesizing that a multi-variable model rather than single variable model may be more pathognomonic of COPD emphysema (COPD-E), the present study explored for disease-associated determinants of disease severity, and treatment success in Taiwanese patients with COPD-E. Methods: The present single-center, prospective, non-randomized study enrolled 125 patients with COPD and 43 healthy subjects between March 2015 and February 2021. Adopting a multimodal approach, including bioinformatics-aided analyses and geospatial modeling, we performed an integrated analysis of selected epigenetic, clinicopathological, geospatial, and air pollutant variables, coupled with correlative analyses of time-phased changes in pulmonary function indices and COPD-E severity. Results: Our COPD cohort consisted of 10 non-, 57 current-, and 58 ex-smokers (median age = 69 ± 7.76 years). Based on the percentages of low attenuation area below -950 Hounsfield units (%LAA$_{-950insp}$), 36 had mild or no emphysema (%LAA$_{-950insp} < 6$), 22 were moderate emphysema cases ($6 \leq$ %LAA$_{-950insp} < 14$), and 9 presented with severe emphysema (%LAA-950insp ≥ 14). We found that BMI, *lnc-IL7R*, $PM_{2.5}$, PM_{10}, and SO_2 were differentially associated with disease severity, and are highly-specific predictors of COPD progression. Per geospatial levels, areas with high BMI and *lnc-IL7R* but low $PM_{2.5}$, PM_{10}, and SO_2 were associated with fewer

and ameliorated COPD cases, while high $PM_{2.5}$, PM_{10}, and SO_2 but low BMI and *lnc-IL7R* characterized places with more COPD cases and indicated exacerbation. The prediction pentad effectively differentiates patients with mild/no COPD from moderate/severe COPD cases, (mean AUC = 0.714) and exhibited very high stratification precision (mean AUC = 0.939). Conclusion: Combined BMI, *lnc-IL7R*, $PM_{2.5}$, PM_{10}, and SO_2 levels are optimal classifiers for accurate patient stratification and management triage for COPD in Taiwan. Low BMI, and *lnc-IL7R*, with concomitant high $PM_{2.5}$, PM_{10}, and SO_2 levels is pathognomonic of exacerbated/aggravated COPD in Taiwan.

1. Introduction

Chronic Obstructive Pulmonary Disease (COPD), entailing small airways inflammatory disease and parenchyma destruction (otherwise known as emphysema), is a common, avoidable, and currently incurable respiratory pathology secondary to protracted and significant exposure to noxious gases and/or particulate matter [1,2]. The pathological hallmarks of emphysema include loss of lung tissue and accelerated loss of pulmonary function [3], where the latter is associated with reduced gas exchange, altered airway dynamics, impaired expiratory airflow, and progressive air trapping [4].

Protracted exposure to air pollutants, including oxides of nitrogen (NO_x), ambient ozone (O_3), emitted hydrocarbons (HC), and fine particulate matter < 2.5 μm in aerodynamic diameter ($PM_{2.5}$), has been implicated in the progressive reduction of pulmonary function in patients with COPD/emphysema [5–7]. The significant increase in $PM_{2.5}$ exposure-specific COPD burden observed over the last decade [8], cannot be decoupled from reported reduction in lung function indices, such as the forced vital capacity (FVC), forced expiratory volume in 1s (FEV1), maximum mid-expiratory flow (MMEF), and FEV1/FVC ratio, elicited by every 5 μg/m^3 increase in $PM_{2.5}$ [6].

While our understanding of the pathobiology of COPD has greatly increased in the last two decades, its prevalence, disease burden, and mortality remain unabated, pharmacological therapies continue to exhibit limited effects on morbidity and mortality, and it remains rather unclear why the presence and severity of emphysema differs significantly between patients with COPD [9].

Although tobacco smoking is widely reported to strongly influence predisposition to COPD, other environmental factors, including noxious gases and particulate matters, and endogenous (age, genetic, and epigenetic) factors, are increasingly documented as important determinants of COPD [10,11]. With only ~14% of the overall COPD burden attributable to occupational exposures [12], the relevance of endogenous factors in the pathogenesis of COPD is accentuated by reports indicating that globally, an estimated 25–45% of COPD cases were "never smokers" [12,13].

The last decade has been characterized by accruing evidence that tobacco smoking, which is considered a principal risk factor for COPD, alongside aging, is associated with epigenetic reprogramming of the bronchial epithelium and that epigenetic signal transduction pathways regulate COPD-related airway inflammation [14]. More so, Zeng H, et al. demonstrated that cigarette smoking elicits pulmonary cell death and B-cell lymphoma/leukemia-2 (Bcl-2) promoter hypermethylation in emphysema models through induction of oxidative stress and activation of epigenetic DNA methyltransferase enzyme 1 (DNMT1) [15]. Recently, our team reported that Toll-like receptor (TLR)-related long non-coding interleukin 7 receptor (*lnc-IL7R*) levels were significantly downregulated in the peripheral blood mononuclear cells of patients with COPD, compared with those from their healthy control peers.l This suppressed *lnc-IL7R* levels was also found to be associated with impaired pulmonary function, and increased risk of COPD exacerbation [16].

Against this background, the present study probed for pathognomonic biomarkers and/or clinicopathological parameters that determine the severity of emphysema in patients with COPD (COPD-E), affect disease course, and predict clinical outcome, while informing therapeutic strategy in Taiwan. Stemming from the working hypothesis that a multi-variable model rather than single variable model may be more pathognomonic of COPD-E, we demonstrate herein that a pentad comprising of body mass index (BMI), *lnc-IL7R*, ambient $PM_{2.5}$, PM_{10}, and SO_2 concentrations were differentially associated with disease severity in Taiwanese patients with COPD-E, and are highly-specific predictors of COPD-E progression. For the first time, to the best of our knowledge, the present study demonstrated that per geospatial levels, areas with high BMI and *lnc-IL7R* but low ambient $PM_{2.5}$, PM_{10}, and SO_2 were associated with fewer and ameliorated COPD-E cases, while high $PM_{2.5}$, PM_{10}, and SO_2 but low BMI and *lnc-IL7R* characterized regions with more COPD-E cases and indicated disease exacerbation. The prediction pentad effectively differentiates patients with mild/no COPD from moderate/severe COPD cases, (mean AUC = 0.714) and exhibited very high stratification precision (mean AUC = 0.939).

2. Methods

2.1. Study Design and Patients

The study enrolled 168 subjects (patients with COPD, $n = 125$ and healthy participants, $n = 43$) who presented and underwent high resolution computed tomography (HRCT) to determine the presence and severity of COPD-E at the Department of Thoracic Medicine, Shuang Ho Hospital, Taipei Medical University (New Taipei, Taiwan) between March 2015 and February 2021. Image attenuation on the acquired HRCT scans of participants' entire lung was assessed using the APOLLO workstation version 1.2 (VIDA Diagnostics Inc., Coralville, IA, USA) at a single reading center by two highly trained experts. Based on the percentages of low attenuation area below -950 Hounsfield units ($\%LAA_{-950insp}$), patients were classified as having no or mild emphysema ($\%LAA_{-950insp} < 6$), moderate emphysema ($6 \leq \%LAA\text{-}950insp < 14$), and severe emphysema ($\%LAA_{-950insp} \geq 14$) [17]. COPD severity assessment was consistent with the Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines [5].

The present single-center, prospective, two-arm, non-randomized study was approved by the Joint Institutional Review Board of Taipei Medical University (TMU-JIRB Approval No.: N201803059, N201902021), and was compliant with the Declaration of Helsinki guidelines on studies involving human subjects. Written informed consent was obtained from all participants before sample collection.

2.2. Inclusion and Exclusion Criteria

Inclusion criteria: Participants were included in the COPD arm if they were aged ≥ 40 years, had established COPD diagnosis, with a FEV1/FVC < 70% after administration of a bronchodilator, according to the GOLD criteria, in stable condition, with no COPD exacerbation in the last 1 month, had no diagnosis of cardiac disease, and accepted to participate by providing written informed consent. For the healthy arm, participants had FEV1 > 80% and FEV1/FVC > 75%, and no known systemic diseases in the previous 3 months.

Exclusion criteria: Participants were excluded if they had reversible airflow obstruction greater than 12% and 200 mL after inhalation of bronchodilator (according to the American thoracic Society (ATS) guidelines), had documented clinical history of previous or present asthma episodes, or coexisting abnormalities on the CT scan.

2.3. Geographic Information System (GIS) and Ambient Air Pollutant Exposure

Ambient air pollution data, namely, concentrations of $PM_{2.5}$, PM_{10} (PM < 10 μm aerodynamic diameter), SO_2, O_3, NO_2, NO, NO_X, carbon monoxide (CO), total hydrocarbon content (THC), non-methane hydrocarbons (NMHC), and methane (CH_4), were obtained from the Air Quality Monitoring (AQM) Networks of the Taiwan Environmental Protection

Administration (https://airtw.epa.gov.tw/ENG/Sitemap.aspx; accessed 1 February 2021). PM concentration is uninterruptedly measured by all AQM stations and recorded hourly. PM exposure data per diem was assigned to participants based on their residential address. The mean daily concentrations of ambient $PM_{2.5}$ and PM_{10} in the preceding years were computed for subsequent analyses. The nearest 3 AQM stations (stations I, II and III) were identified using the ArcGIS server software version 10.8.1 (ESRI, Redlands, CA, USA), and then air pollution data was extracted. The distance range between AQM stations I, II, III, and participants' residential addresses was 0.36–8.73 (mean = 2.35), 1.28–14.33 (mean = 3.55), and 2.56–35.93 (4.73) km, respectively. The mean daily concentrations of $PM_{2.5}$ and PM_{10} were determined using the inverse distance weighting (IDW) interpolation method.

2.4. Sample Preparation and Quantitative Reverse Transcription PCR (RT-qPCR)

After the isolation of peripheral blood mononuclear cells (PBMC) from the whole blood of patients with COPD and healthy participants strictly following the Ficoll—Hypaque density gradient centrifugation protocol previously described by Chen TT, et al. [18], total RNA was isolated from the cell samples using the TRIzol™ Plus RNA Purification Kit (Cat. #12183555; Thermo Fisher Scientific, Waltham, MA, USA). Serum RNA was purified using the Plasma/Serum Circulating and Exosomal RNA Purification Kit (Slurry Format) (Cat. #42800; Norgen Biotek Corp., Thorold, ON, Canada). Following determination of total RNA concentration using the NanoDrop ND1000 spectrophotometer (Nyxor Biotech, Paris, France), PCR mix were prepared using the SYBR™ Green PCR Master Mix (Cat. #4309155; Applied Biosystems Inc., Carlsbad, CA, USA). The PCR contained the primers, the fluorogenic probe mix, and the TaqMan Universal PCR Master mix (Applied Biosystems Inc.). All amplification reactions were performed in quadruplicates from 20 ng complementary DNA (cDNA) in the Bio-Rad C1000 real-time PCR system (Bio-Rad, Cambridge, MA, USA) using the following conditions: 95 °C for 3 min, 35 cycles at 95 °C for 15 s, 60 °C for 30 s, 72 °C for 30 s, and 72 °C for 10 min. For analysis of results, all values were normalized to the levels of the housekeeping genes *18S rRNA* (cellular) or *GAPDH* (extracellular), which served as the internal control. All procedures were consistent with manufacturers' instructions, and the following PCR primer sequences were used: *lnc-IL7R* (forward): 5′-CCAGCCTTTGCCTCTTCCTTCAAT-3′, *lnc-IL7R* (reverse): 5′-CCGTA CCAAGTCTCTTAGCCCCTC-3′; *18S* rRNA (forward): 5′-TGTGCCGCTAGAGGTGAAATT-3′, *18S* rRNA (reverse): 5′-TGGCAAATGCTTTCGCTTT-3′; *GAPDH* (forward): 5′-ATGGGGAAGGTGAAGGTCG-3′, and *GAPDH* (reverse): 5′-GGGGTCATTGATGGCAACAAT-3′.

2.5. Statistical Analysis

Results are expressed as means ± standard deviations. Pearson's chi-squared (χ^2) test was used to determine the relationship or association between categorical variables. The paired *t*-test was used for comparing continuous data. The Student's *t*-test was used to assess alterations in pulmonary function based on %LAA$_{-950insp}$, concentration levels of ambient air pollutants, and *lnc-IL7R* expression levels. *p*-value ≤ 0.05 defined statistical significance. All statistical analyses were performed using IBM SPSS Statistics for Windows, Version 25.0 (IBM Corp. Released 2017, Armonk, NY, USA: IBM Corp). Geospatial visualization and analysis were performed using the ArcGIS server software version 10.8.1 (ESRI, Redlands, CA, USA).

3. Results

Emphysema (COPD-E) severity correlates with GOLD stage, and is indicative of disease progression in Taiwanese patients with COPD.

Table 1 contains the baseline characteristics of our cohort of patients with COPD (*n* = 125) and healthy controls (*n* = 43). As shown in Table 1, the median age was 69 ± 7.76 years (range: 41–87), 88.8% of all patients with COPD were male and aged 41–87 years. Of these, 6.31%

were never-smokers, while 47.75% and 45.94% were ex- and current smokers, respectively. The median BMI was of 20.0 ± 4.46 kg/m^2, and 23.66 ± 3.99 kg/m^2 for female and male COPD cases. Stratification of patients with COPD according to the GOLD-based staging showed that compared with GOLD stage I cases with median FEV$_1$ of $84.55 \pm 7.13\%$, there was 23.1-fold (median FEV$_1$ = $65 \pm 9.27\%$), 53.8-fold (median FEV$_1$ = $39.05 \pm 5.53\%$), and 70.4-fold (median FEV$_1$ = $25 \pm 3.85\%$) decrease in the pulmonary function indices of their GOLD stages II, III, and IV peers, respectively. Regardless of smoking status, compared with the healthy control (median FEV$_1$: non-smoker = 101.00 ± 5.10 vs smoker = 96.85 ± 6.68), significant decline in lung function was observed in the non-stratified COPD group (median FEV$_1$ = $57 \pm 19.16\%$) (Table 1).

Table 1. Baseline characteristics of our cohort of healthy and COPD participants (n = 168).

Variables	Healthy Controls (n = 43)		Patients with COPD (GOLD Stage, n = 125)			
	Non-Smoker (n = 21)	Smoker (n = 22)	I (n = 18)	II (n = 58)	III (n = 38)	IV (n = 11)
	Age (years)					
Mean ± SD (Min-Max)	68.33 ± 7.02 (50.00–80.00)	67.45 ± 6.75 (47.00–80.00)	69.39 ± 5.78 (61.00–80.00)	67.97 ± 8.91 (41.00–9.00)	71.24 ± 6.95 (56.00–80.00)	67.09 ± 5.38 (61.00–79.00)
Median (IQR)	69.00 (67.00–73.00)	69.00 (65.25–71.00)	68.00 (65.25–71.50)	68.50 (62.25–73.00)	70.50 (67.00–77.25)	66.00 (63.00–69.00)
	Sex, n (%)					
Male	8 (38.10)	17 (77.27)	17 (94.44)	55 (94.83)	31 (81.58)	9 (81.82)
Female	13 (61.90)	5 (22.73)	1 (5.56)	3 (5.17)	7 (18.42)	2 (18.18)
	BMI, kg·m^{-2}					
Mean ± SD (Min-Max)	22.79 ± 2.15 (20.50–28.80)	23.14 ± 2.58 (19.11–29.20)	24.05 ± 3.12 (19.10–29.36)	24.33 ± 4.41 (16.40–34.80)	22.50 ± 3.77 (15.80–36.20)	21.31 ± 3.60 (16.20–27.70)
Median (IQR)	22.00 (21.20–24.00)	22.76 (21.85–23.95)	23.90 (21.63–26.29)	24.14 (21.16–26.60)	22.30 (20.00–24.50)	20.60 (19.90–22.98)
	Tobacco Smoking, n (%)					
Current smoker	0 (0.00)	13 (59.09)	5 (27.78)	31 (53.44)	11 (28.95)	2 (18.18)
Ex-smoker	0 (0.00)	9 (40.91)	13 (72.22)	23 (39.66)	22 (57.89)	8 (72.73)
Never-smoker	100 (100)	0 (0.00)	0 (0.00)	4 (6.90)	5 (13.16)	1 (9.09)
	Smoking pack-years					
Mean ± SD (Min-Max)	0 (0.00–0.00)	65.00 ± 31.43 (30.00–145.00)	48.89 ± 35.19 (5.00–150.00)	49.02 ± 36.34 (0.00–180.00)	49.30 ± 35.66 (0.00–156.00)	56.73 ± 37.65 (0.00–123.00)
Median (IQR)	0 (0.00–0.00)	57.00 (40.00–79.50)	42.50 (20.50–60.00)	40.00 (23.00–60.00)	40.00 (25.00–75.00)	46.00 (35.00–85.00)
	Pulmonary function indices					
FEV1 (L) Mean ± SD (Min-Max)	1.98 ± 0.37 (1.26–2.76)	2.18 ± 0.42 (1.74–3.47)	1.95 ± 0.26 (1.55–2.56)	1.67 ± 0.40 (1.01–3.07)	0.98 ± 0.25 (0.61–1.51)	0.61 ± 0.13 (0.43–0.87)
Median (IQR)	1.97 (1.80–2.06)	2.09 (1.90–2.29)	1.90 (1.74–2.11)	1.61 [b'] (1.38–1.90)	0.99 [a'b'c'd'] (0.74–1.12)	0.58 [a'b'c'd'] (0.52–0.66)
FEV1 % Mean ± SD (Min-Max)	101.43 ± 5.10 (95.00–117.00)	98.60 ± 6.68 (90.00–111.00)	85.42 ± 5.24 (80.00–97.70)	63.81 ± 8.57 (50.00–79.00)	40.01 ± 5.71 (32.00–49.80)	24.85 ± 3.98 (17.50–29.90)
Median (IQR)	101.00 (98.00–103.00)	96.85 (93.00–103.00)	84.55 (81.3–86.68)	65.00 [a'b'c] (57.38–72.00)	39.05 [a'b'c'd'] (35.00–45.00)	25.00 [a'b'c'd'] (22.10–27.95)
FEV1/FVC % Mean ± SD (Min-Max)	100.76 ± 9.42 (80.00–125.00)	100.45 ± 8.01 (90.00–120.00)	63.49 ± 4.05 (54.64–68.26)	59.33 ± 6.86 (45.00–69.72)	47.93 ± 8.60 (28.00–65.00)	38.84 ± 8.41 (27.00–49.61)
Median (IQR)	100.00 (98.00–105.00)	98.00 (95.25–107.75)	63.68 (61.25–66.87)	59.25 [a'b'] (54.12–65.50)	46.50 [a'b'c'd'] (42.11–55.25)	41.41 [a'b'c'd'] (30.93–45.67)
	Emphysema severity					
Null/Mild (%)			66.67	19.05	0.00	0.00
Moderate (%)			33.33	66.67	69.23	20.00
Severe (%)			0.00	14.28	30.77	80.00

COPD, chronic obstructive pulmonary disease; GOLD, Global Initiative for Chronic Obstructive Lung Disease; M, male; F, female; FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity; BMI, body mass index; IQR, interquartile range. The values of FEV1/FVC % and FEV1 % were analyzed by Kruskal–Wallis tests and Dunn's multiple comparisons ([a'] $p < 0.01$, compared with non-smoker; [b'] $p < 0.01$, compared with smoker; [c] $p < 0.05$, [c'] $p < 0.01$, compared with COPD patients with GOLD stage I; [d'] $p < 0.01$, compared with COPD patients with GOLD stage II.

Furthermore, based on the working hypothesis that a phenotype-based categorization of COPD which takes into account presence/severity of COPD-E and exacerbation history, rather than solely on level of airflow limitation (FEV_1) better reflects the nosological complexity and constitutive heterogeneity of COPD, based on $\%LAA_{-950insp}$, we found 66.67% and 33.33% of the GOLD 1 cases had mild and moderate COPD-E, respectively. 72.73% of all patients with severe COPD-E were GOLD 3–4, 27.27% were GOLD 2, and 0.00% were GOLD 1 (Table 1). Conversely, all mild COPD-E cases were GOLD 2 (100%), while 17.24% of all moderate COPD-E was found to be GOLD 1 (Table 1). These data do indicate that COPD-E is not only diagnostic of COPD, but is also indicative of disease progression or exacerbation.

3.1. Emphysema Risk Modulators in Taiwanese Patients with COPD

Observing that COPD-E reflects the nosological complexity and disease progression in patients with COPD, we explored for COPD-E-specific determinants of disease severity, and treatment success in Taiwanese patients with COPD. Establishing a significantly strong correlation between the imaging emphysema severity indices, total LAA% and $\%LAA_{-950insp}$ (r = 0.82, $p < 0.001$) (Figure 1A), we next probed for probable correlation between emphysema severity and ambient air pollutants, epigenetic, spirometric, anthropometric, and clinical variables. The $\%LAA_{-950insp}$ was inversely correlated with epigenetic *lnc-IL7R* (r = −0.30, $p = 0.002$), post-bronchodilator FEV_1/FVC (r = −0.41, $p < 0.001$), and BMI (r = −0.47, $p < 0.001$) (Figure 1B–D). Ambient $PM_{2.5}$ (r = 0.31, $p < 0.001$), PM_{10} (r = 0.30, $p = 0.001$), NO_2 (r = 0.18, $p = 0.043$), SO_2 (r = 0.25, $p = 0.004$), and THC (r = 0.20, $p = 0.024$) concentrations were positively correlated with $\%LAA_{-950insp}$, while O_3 (r = −0.09, $p = 0.341$) was inversely correlated (Figure 1E–K). In addition, and equivocal association was found between $\%LAA_{-950insp}$ and pack-year (r = 0.17, $p = 0.057$) or age (r = 0.11, $p = 0.240$) (Figure 1L,M). These data are suggestive of a multifactorial risk modulatory cluster for COPD-E in Taiwanese patients with COPD.

3.2. Delineating Predictors of Disease Severity in Taiwanese Patients with COPD-E

Consistent with earlier data, using a supervised machine learning algorithm, artificial neural network (ANN) modelling based on the hyperbolic tangent activation (TanH) model with random holdback validation showed that ambient air pollutants ($PM_{2.5}$, PM_{10}, NO_2, SO_2, THC, O_3), epigenetic (*lnc-IL7R*), anthropometric (Age, BMI), lifestyle (smoking history, pack-year), and geospatial components (longitude and latitude of participants' residential addresses) all contribute differentially to emphysema (COPD-E) severity ($\%LAA_{-950insp}$) (Figure 2A,B). The initial fitting of the ANN model on training dataset (80% of our COPD cohort) for variable selection and parameter estimation to predict $\%LAA_{-950insp}$ indicated strong association between listed variables and COPD-E severity ($R^2 = 0.49$; root mean square error, RMSE = 0.64; mean absolute deviation, MAD = 0.460) (Figure 2A). This was validated by the unbiased evaluation and hyper-parameter fine-tuning of the training dataset-fitted ANN model using the validation dataset (20% of our COPD cohort) ($R^2 = 0.65$; root mean square error, RMSE = 0.60; MAD = 0.37) (Figure 2A). Interestingly, our prediction profile analysis indicate that BMI < 23.62 kgm^{-2}, blood *lnc-IL7R* level < 0.53, smoking history (ex or current) with pack-year > 50.58, SO_2 > 3.08 parts per billion (ppb), O_3 < 26.27 ppb, NO_2 > 20.08 ppb, THC < 2.18 parts per million (ppm), PM_{10} < 40.10 mg m^{-3}, and $PM_{2.5}$ > 22.18 mg m^{-3} define an exacerbated COPD-E phenotype (Figure 2B). Though not optimal, goodness of fit analysis showed that the ANN model can relatively predict COPD-E ($\%LAA_{-950insp}$) severity fairly accurately (RMSE = 0.58, $R^2 = 0.46$, $p < 0.0001$) (Figure 2C). In parallel unsupervised machine learning analysis using the hierarchical clustering model, the COPD cohort was pooled into three clusters reflecting COPD-E severity, with cluster 0, 1, and 2 representing no emphysema, mild-moderate emphysema, and moderate-severe emphysema ($R^2 = 0.39$; Pearson's g = 0.47; Dunn index = 0.17; Silhoutte = 0.250; Entropy = 0.41, Calinski-Harabasz index = 20.84) (Figure 2D). The relatively low R^2, Dunn index, Pearson's g, and silhouette index, coupled

with high entropy score necessitated model optimization. For optimal variable selection, the effect of BMI (t-ratio = -5.75, F-ratio = 33.01, $p < 0.0001$), *lnc-IL7R* (t-ratio = -2.07, F-ratio = 4.28, $p = 0.042$), and SO_2 (t-ratio = 2.00, F-ratio = 3.99, $p = 0.049$) on COPD-E severity were statistically significant, while smoking history (t-ratio = 1.94, F-ratio = 3.75, $p = 0.056$), PM_{10} (t-ratio = -1.51, F-ratio = 2.30, $p = 0.133$), and $PM_{2.5}$ though exhibiting good effect (F-ratio ≥ 2), were statistically non-significant (Figure 2E,F).
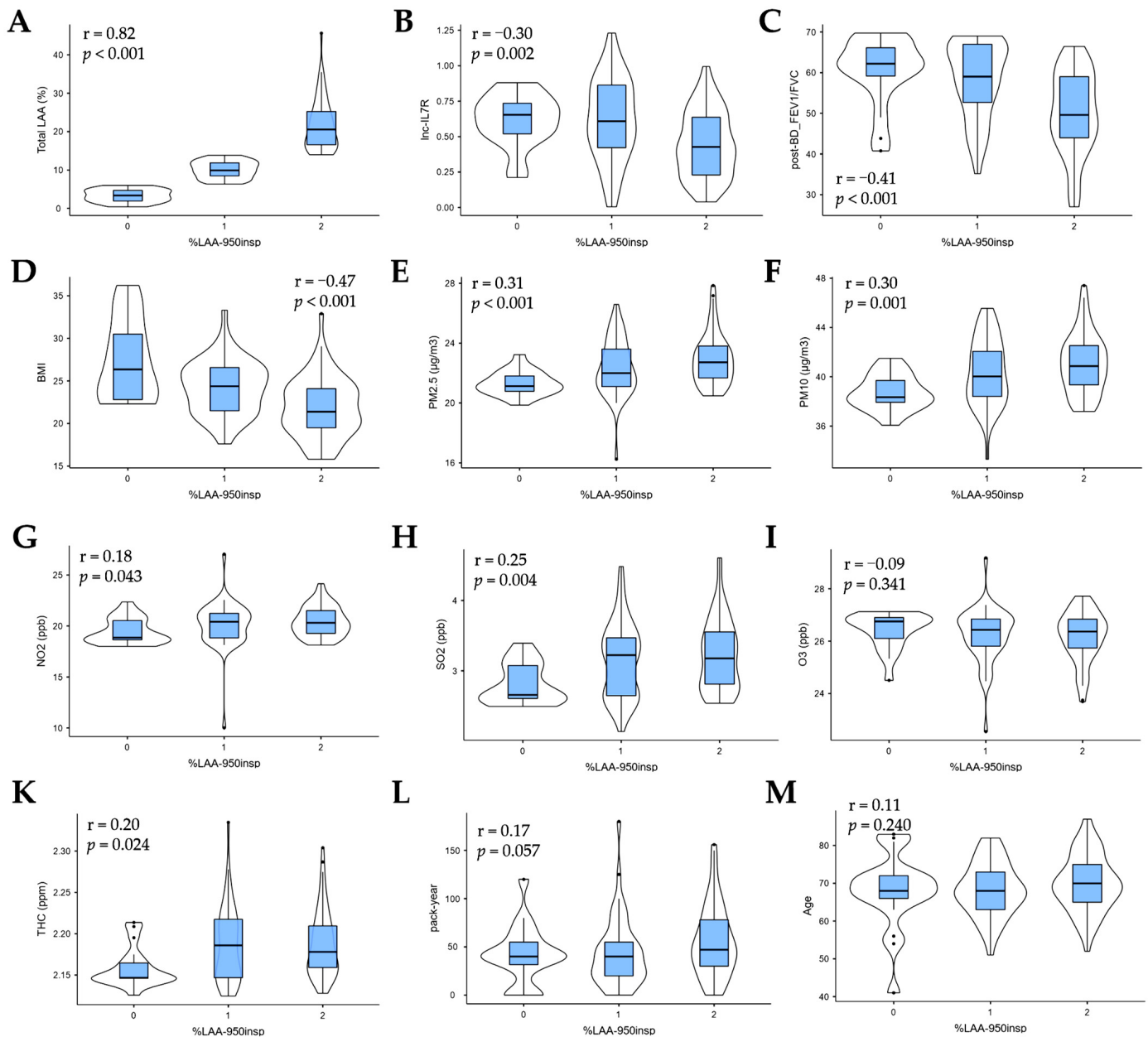
Figure 1. Disease risk modulators in Taiwanese patients with COPD-E. Violin plots with included box plots showing the correlation between (**A**) Total LAA, (**B**) *lnc-IL7R*, (**C**) post-bronchodilator FEV1/FVC ratio, (**D**) BMI, (**E**) $PM_{2.5}$, (**F**) PM_{10}, (**G**) NO_2, (**H**) SO_2, (**I**) O_3, (**K**) THC, (**L**) pack-year, or (**M**) age, with %LAA$_{-950insp}$-based COPD-E severity. COPD-E, emphysema in patients with COPD; r, Pearson correlation coefficient; %LAA$_{-950insp}$, percentage of lung attenuation area with values less than -950 Hounsfield units on inspiratory CT scan. For %LAA$_{-950insp}$: 1, no or mild emphysema (%LAA$_{-950insp} < 6$); 2, moderate emphysema ($6 \leq$ %LAA$_{-950insp} < 14$); and 3, severe emphysema (%LAA$_{-950insp} \geq 14$).
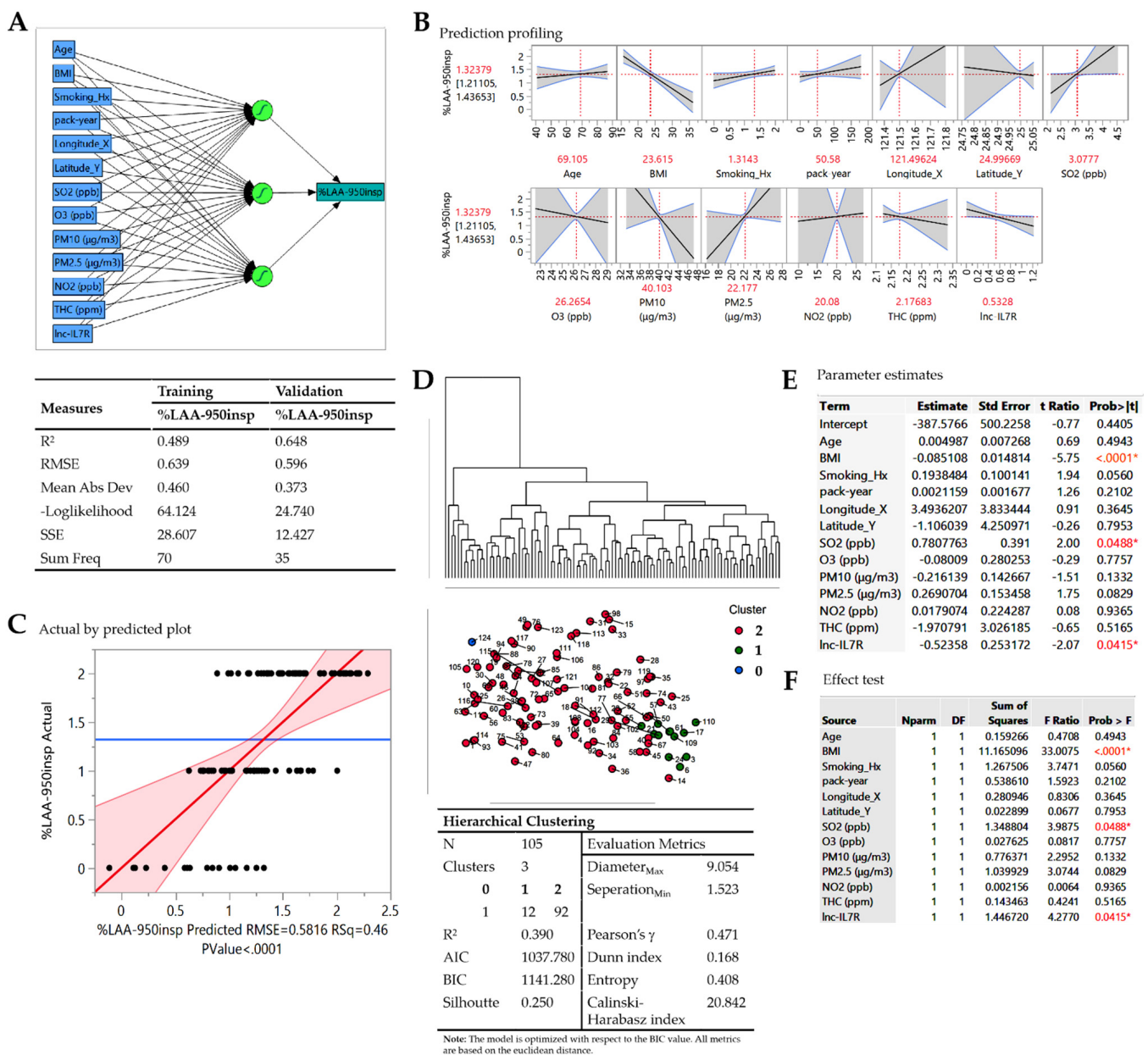
Figure 2. Delineating predictors of disease severity in Taiwanese patients with COPD-E. (**A**) Artificial neural network (ANN) model schema (upper panel) and statistics chart (lower panel) showing the 13 disease-related variables of interest input, 3 auto-determined hidden, and predicted %LAA$_{-950insp}$-based COPD-E severity output layers. (**B**) Prediction profiler showing the effect of age, BMI, smoking history, pack-year, longitude, latitude, SO$_2$, O$_3$, PM$_{2.5}$, PM$_{10}$, NO$_2$, THC, and *lnc-IL7R* on predicted %LAA$_{-950insp}$-based COPD-E severity. Cut-off values are indicated in red. (**C**) Actual vs. predicted %LAA$_{-950insp}$-based COPD-E severity plot. (**D**) Hierarchical clustering dendrogram and statistics of severity-stratified COPD-E cases. (**E**) Parameter estimates, and (**F**) effect test charts of the panel of variables. R^2, coefficient of determination; RMSE, root mean square error; SSE, sum of squared estimate of errors; Hx, history; ppb, part per billion; ppm, part per million; N, number of cases; AIC, Akaike's Information Criteria; BIC, Bayesian Information Criteria; Values in red, statistically significant.

3.3. Severity-Stratified Spatiofunctional Interaction between Individual Predictors of COPD-E

To determine if and to what extent these COPD-E-associated factors interact and/or form a pathogenic cascade, we performed a bootstrapped network analysis (bootstrap *n* = 1000). Of pathophysiological relevance, we found no interaction between any of the COPD-E pathognomonic factors in patients with no or mild emphysema (%LAA$_{-950insp}$ < 6) (Figure 3A, left). For the moderate emphysema cases (6 ≤ %LAA$_{-950insp}$ < 14), we found

that ambient SO_2, NO_2, THC, PM_{10}, $PM_{2.5}$, and O_3 form a loose cascade with endogenous factors BMI, *lnc-IL7R*, and Age in the context of patients' geolocation, while smoking history and pack-year were non-contributors to the cascade (Figure 3A, middle). Conversely, a tight-knit cluster was observed between all variables in patients with severe COPD-E (%LAA-950insp $\geq$ 14), with relatively high interaction density (Figure 3A, right). This partial deciphering of the structural and functional networks of COPD-E on a spatiotemporal scale, was corroborated by the Barrat, Onnela, Watts and Strogatz (WS), and Zhang clustering coefficients [19,20], which measure the propensity to which the nodes/factors tend to cluster together, and quantify the abundance of connected triangles in the weighted severity-based pathogenic networks (Figure 3B). Next, exploiting the ability of the centrality plot to identify important nodes/variables that determine disease severity, as well as the nature of their influence, we identified SO_2, PM_{10}, $PM_{2.5}$, O_3, and *lnc-IL7R tetrad* as an important predictor of disease progression and severity, and that while SO_2, PM_{10}, and $PM_{2.5}$, drove disease progression and were positively correlated with COPD-E severity, O_3 and *lnc-IL7R* attenuated progression and were negatively associated with disease severity (Figure 3C).
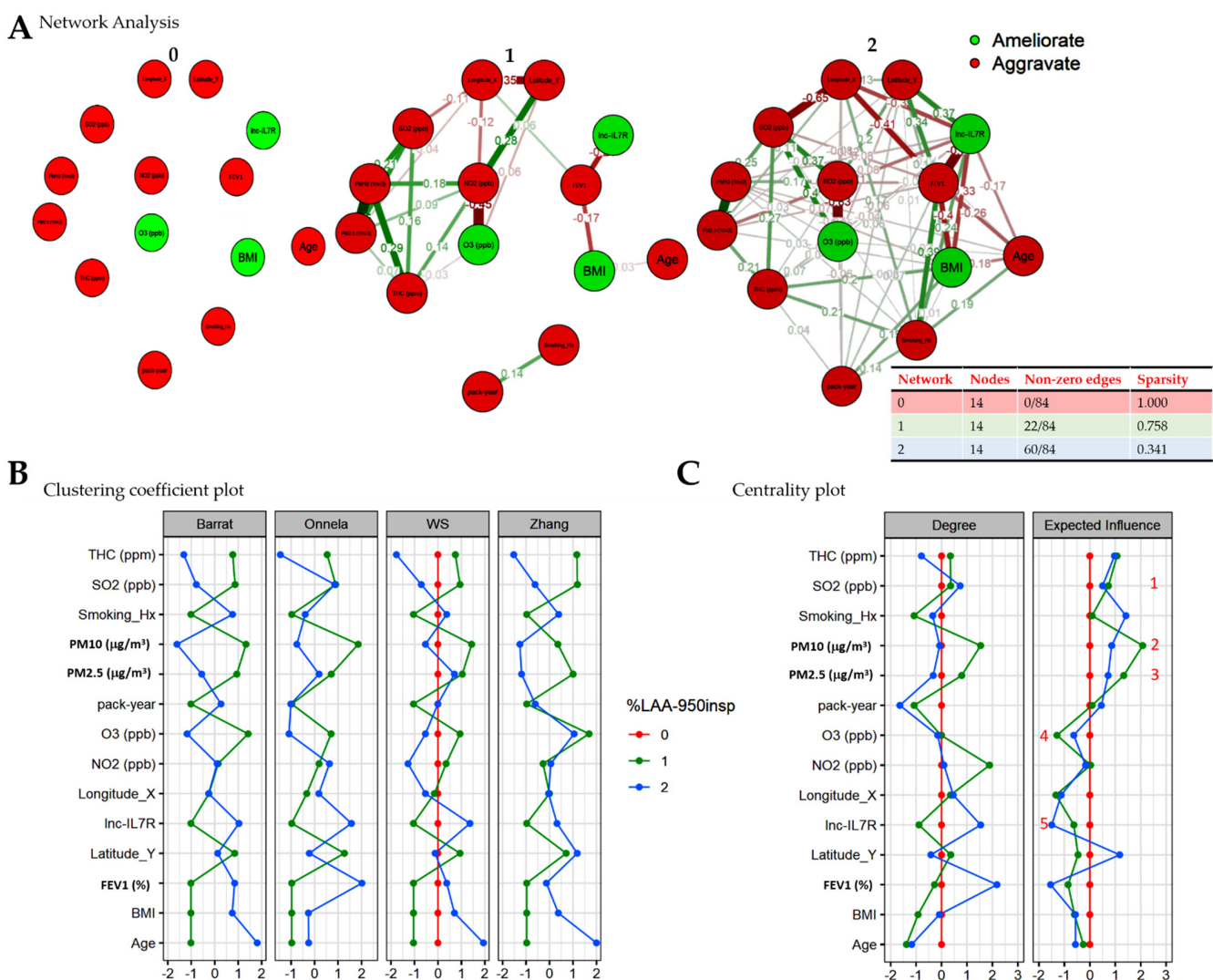

Network	Nodes	Non-zero edges	Sparsity
0	14	0/84	1.000
1	14	22/84	0.758
2	14	60/84	0.341

Figure 3. Severity-stratified spatiofunctional interaction between individual predictors of COPD-E progression in Taiwan. (**A**) Visualization, (**B**) clustering coefficient plot, and (**C**) centrality plot of the interaction between indicated variables in the determination of COPD-E severity as shown by bootstrapped network analysis. Number of bootstraps = 1000. red circle, aggravating factor; green circle, ameliorating factor. WS, Watts and Strogatz.

3.4. BMI, lnc-IL7R, PM$_{2.5}$, PM$_{10}$, and SO$_2$ Levels Are Excellent Classifiers for Accurate Patient Stratification and COPD-E Management Triage in Taiwan

Finally, receiver operating characteristic (ROC) curve analysis was used to evaluate the discriminatory power of all COPD-E-associated factors, namely geospatial variables (longitude, latitude), ambient particulate matter and pollutants (PM$_{2.5}$, PM$_{10}$, SO$_2$, O$_3$, THC, NO$_2$), lifestyle (pack-year), anthropometric (BMI), and endogenous (age, epigenetic *lnc-IL7R*) factors. All factors, except latitude (area under the ROC curve, AUC = 0.55), age (AUC = 0.51), O$_3$ (AUC = 0.58), and pack-year (AUC = 0.57), exhibited acceptable capability to discriminate between patients with COPD-E and those with non-emphysematous COPD phenotype, regardless of severity status (Figure 4A,B). In parallel analysis of our COPD cohort using precision–recall curves, we found the latitude (AUC = 0.89, associated criterion $\leq$ 25.06), longitude (AUC = 0.93, associated criterion $\leq$ 121.80), SO$_2$ (AUC = 0.95, associated criterion $\geq$ 2.14), THC (AUC = 0.94, associated criterion $\geq$ 2.14), PM$_{2.5}$ (AUC = 0.95, associated criterion $\geq$ 16.24), PM$_{10}$ (AUC = 0.95, associated criterion $\geq$ 16.24), BMI (AUC = 0.95, associated criterion $\leq$ 29.36), and *lnc-IL7R* (AUC = 0.91, associated criterion $\leq$ 1.23) excellently stratified patients into COPD-E and non-COPD-E groups (Figure 4C–G). Since a high AUC represents both high recall (related to low false negative rate) and high precision (associated with low false positive rate), where 0.5 denotes a bad classifier and 1, an excellent classifier, these results indicate that BMI, *lnc-IL7R*, PM$_{2.5}$, PM$_{10}$, and SO$_2$ levels are excellent classifiers for accurate patient stratification and management triage for COPD-E in Taiwan.

3.5. BMI, lnc-IL7R, PM$_{2.5}$, PM$_{10}$, and SO$_2$ Are Highly Specific Predictors of COPD-E Severity and Disease Progression in New Taipei City

Furthermore, predictor screening and effect ranking for predictive factor optimization showed that the highest ranked contributors to or determinants of emphysema (COPD-E) severity were BMI (contribution: 2.31, portion: 0.21, rank: 1), *lnc-IL7R* (contribution: 2.04, portion: 0.19, rank: 2), PM$_{2.5}$ (contribution: 1.33, portion: 0.12, rank: 3), PM$_{10}$ (contribution: 0.98, portion: 0.09, rank: 4), and SO$_2$ (contribution: 0.81, portion: 0.07, rank: 5) (Figure 5A). Next, using the Gaussian Processes (GP) generic supervised learning algorithm which is designed to solve regression and probabilistic classification problems, we generated contour and surface profile plots for the GP prediction model of COPD-E (%LAA$_{-950insp}$) severity. First, for geospatial relevance, we found that most non-COPD-E or mild/ameliorated COPD-E cases, as defined by %LAA$_{-950insp}$, were resident above longitude 121.496° along latitude 24.996°, while the largest proportion of exacerbated or severe COPD-E cases were resident beyond this geographical level (Figure 5B). These exacerbated or severe COPD-E cases were defined by serum *lnc-IL7R* levels $\leq$ 0.54, BMI $\geq$ 23.54 kgm^{-2}, and resident under ambient PM$_{2.5}$ $\geq$ 22.48 mg/m^3, PM$_{10}$ $\geq$ 40.46 mg/m^3, and SO$_2$ $\geq$ 3.13 ppb (Figure 5C,D). More so, understanding that sensitivity implies the ability of any test to designate a subject with COPD-E as positive, our GP model of COPD-E severity showed that PM$_{2.5}$ (total sensitivity = 0.833, theta = 1.70 $\times$ 10^{-9}, nugget = 0.001), longitude (total sensitivity = 0.997, theta = 0.005, nugget = 0.001), BMI (total sensitivity = 0.633, theta = 8.21 $\times$ 10^{-6}, nugget = 0.001), and SO$_2$ (total sensitivity = 0.317, theta = 8.75 $\times$ 10^{-5}, nugget = 0.001) are independent determinants or drivers of COPD-E in New Taipei City, Taiwan (Tables 2–4). Since the lower the sensitivity, the greater the ability of the GP model to designate a case without COPD-E as negative, we confirmed that *lnc-IL7R* (total sensitivity = 0.167, theta = 0.000, nugget = 0.001), and latitude (total sensitivity = 0.001, theta = 0.000) are factors of amelioration in the specified region (Tables 2–4). The very low theta values and nugget parameter are of statistical and predictive relevance, as they rule out measurement error and short scale variability [21].
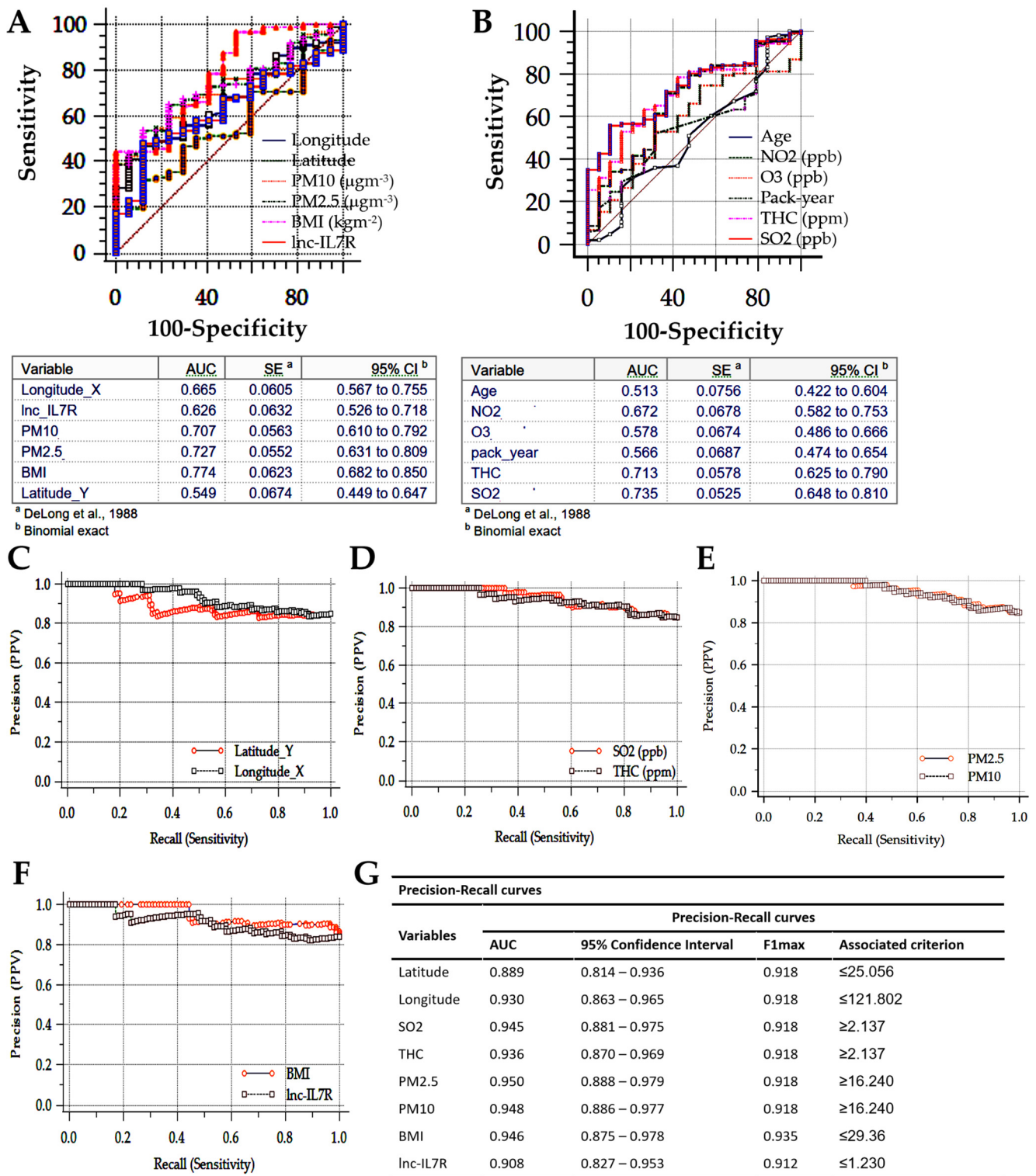
Figure 4. Combined BMI, *lnc-IL7R*, PM$_{2.5}$, PM$_{10}$, and SO$_2$ levels are optimal classifiers for accurate patient stratification and COPD-E management triage in Taiwan. Visual representation (upper panel) and statistics chart (lower panel) of the area under the receiver operating characteristics curves for (**A**) longitude, latitude, PM$_{10}$, PM$_{2.5}$, BMI, *lnc-IL7R*, (**B**) age, NO$_2$, O$_3$, pack-year, THC, and SO$_2$. Precision–recall curves showing the tradeoff between precision and recall for (**C**) longitude, latitude, (**D**) THC, SO$_2$, (**E**) PM$_{2.5}$, PM$_{10}$, (**F**) BMI, and *lnc-IL7R*. (**G**) Statistics chart of the precision–recall curves for (**C–F**).
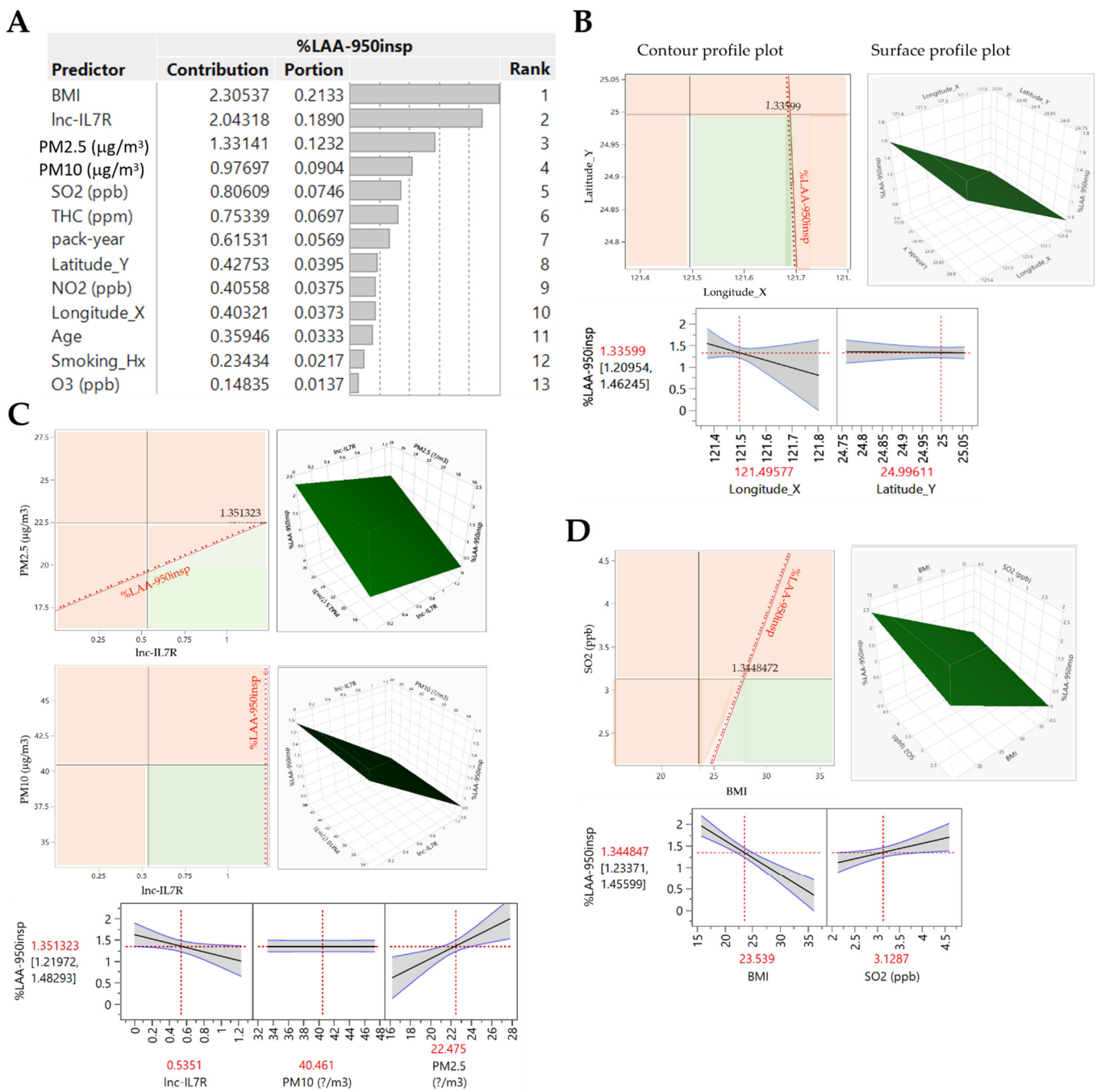
Figure 5. BMI, lnc-IL7R, PM$_{2.5}$, PM$_{10}$, and SO$_2$ are highly specific predictors of COPD-E severity and disease progression in New Taipei City. (**A**) Predictor screening chart showing the contribution and ranking of each variable in predicting COPD—E severity. Contour profile (upper left), surface profile (upper right), and predictor profiler (lower panel) plots for (**B**) longitude, latitude, (**C**) PM$_{2.5}$, PM$_{10}$, lnc-IL7R, (**D**) SO$_2$, and BMI based on the Gaussian process model of %LAA$_{-950insp}$-based COPD-E severity. Fit used Gaussian correlation function. Nugget parameter was set to avoid singular variance matrix.

Table 2. Gaussian process model of %LAA-950insp-based COPD-E severity and progression: *lnc-IL7R*, PM$_{2.5}$, and PM$_{10}$.

Column	Theta	Total Sensitivity	Main Effect	lnc-IL7R Interaction	PM10 Interaction	PM2.5 Interaction
lnc-IL7R	0.0003	0.1667	0.1667	-	0	5.35×10^{-9}
PM10	4.82×10^{-9}	0	0	0	-	0
PM2.5	1.70×10^{-9}	0.8333	0.8333	5.35×10^{-9}	0	-
μ	σ^2	Nugget				
1.3155	470.0283	0.001				
−2*Loglikelihood						
240.6087						

Table 3. Gaussian process model of %LAA-950insp-based COPD-E severity and progression: longitude, and latitude.

Column	Theta	Total Sensitivity	Main Effect	Longitude Interaction	Latitude Interaction
Longitude	0.0046	0.9986	0.9986	-	0
Latitude	0.0003	0.0014	0.0014	0	-
μ	σ^2	Nugget			
1.4838	520.33	0.001			
−2*Loglikelihood					
285.8087					

Table 4. Gaussian process model of %LAA-950insp-based COPD-E severity and progression: BMI, and SO$_2$.

| Column	Theta	Total Sensitivity	Main Effect	BMI Interaction	SO$_2$ Interaction
BMI	8.21×10^{-6}	0.6833	0.6833	-	2.74×10^{-7}
SO2	8.75×10^{-5}	0.3167	0.3167	2.74×10^{-7}	-
μ	σ^2	Nugget			
1.1118	390.41	0.001			
−2*Loglikelihood					
248.4064					

3.6. Low BMI, and lnc-IL7R, with Concomitant High PM$_{2.5}$, and SO$_2$ Levels Is Pathognomonic of Exacerbated/Severe COPD-E in New Taipei City, Taiwan

For geospatial contextualization and visualization of the effect of the delineated determinants of COPD-E severity in New Taipei City, after digitizing our study areas at landscape, regional, and national scales on a global map as polygons, site and cases were marked as point features. The generated site-of-interest map is shown in Figure 6A, with COPD-E cases were concentrated around the Wanhua, Banqiao, Tucheng, Xinzhuang, Hsintien, Zhonghe, and Yonghe Districts. Consistent with Figure 5, the most severe COPD-E cases were in the Wanhua, Banqiao, Zhonghe, and Yonghe Districts, located below longitude 121.496° latitude 24.996°, with %LAA$_{-950insp}$ ≥ 10 (i.e., ≥ 1.3 on a statistical scale of 1–3) (Figure 6B), ambient PM$_{2.5}$ > 24.0 mg/m^3 (Figure 6C), and SO$_2$ > 3.5 ppb (Figure 6D), coupled with endogenous *lnc-IL7R* levels < 0.8 (Figure 6E), and BMI < 25.0 kgm^{-2} (Figure 6F). These data indicated that low BMI, and lnc-IL7R, with concomitant high PM$_{2.5}$, and SO$_2$ levels is pathognomonic of exacerbated/severe COPD-E in New Taipei City, Taiwan.
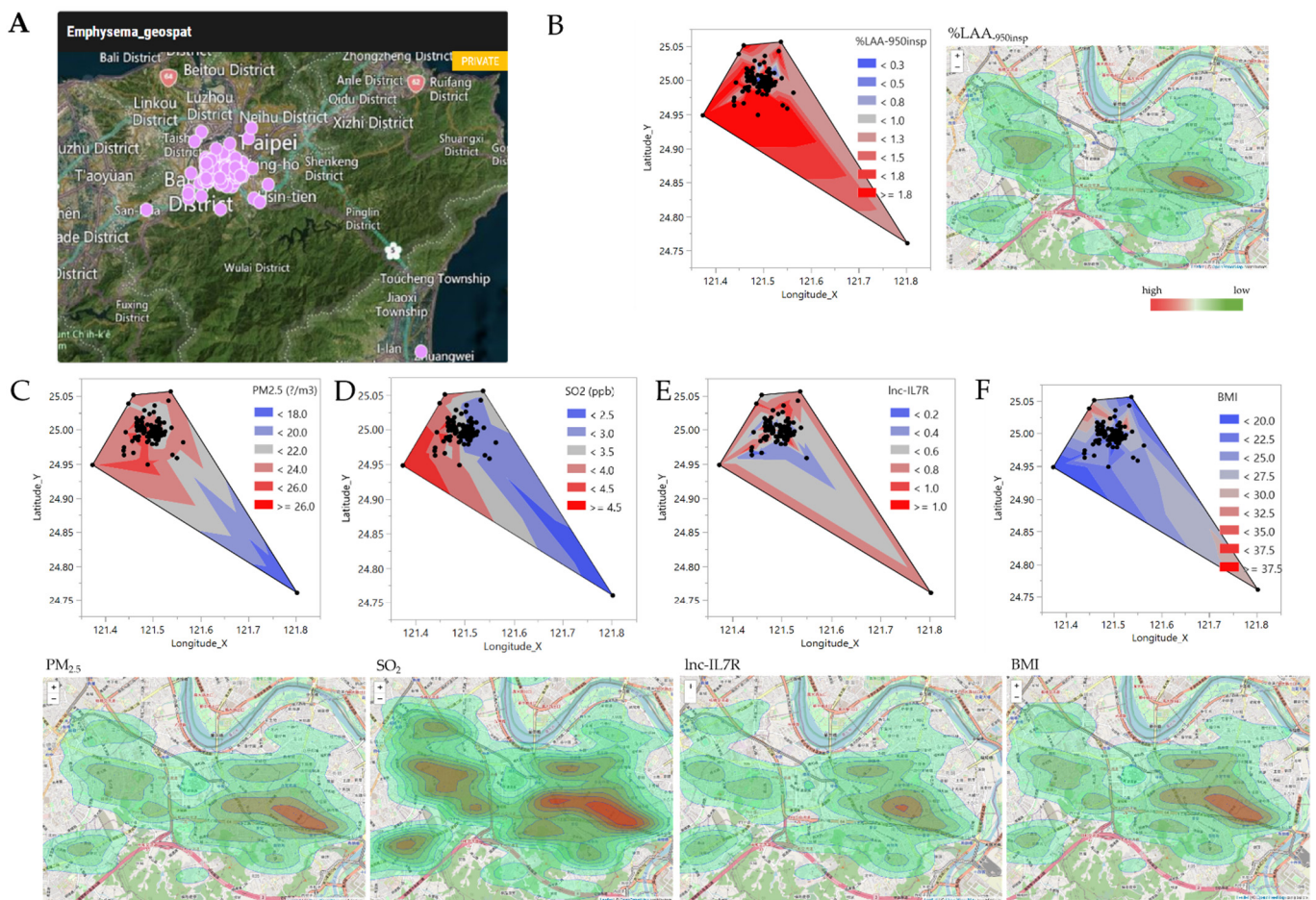
Figure 6. Low BMI, and lnc-IL7R, with concomitant high $PM_{2.5}$, and SO_2 levels is pathognomonic of exacerbated/severe COPD-E in New Taipei City, Taiwan. (**A**) Image of the geospatial mapping patients with COPD-E in New Taipei City. Contour-based polygon plots and heatmap aerial view maps of case distribution according to (**B**) disease severity, and levels of (**C**) $PM_{2.5}$, (**D**) SO_2, (**E**) lnc-IL7R, and (**F**) BMI in patients with COPD-E in New Taipei City.

4. Discussion

Severe emphysema (COPD-E) remains a therapeutic challenge, especially in the light of the limited efficacy of contemporary anti-COPD therapeutic strategy. The present study accentuates the role of a multifactorial risk modulatory cluster for development and/or progression of COPD-E in Taiwanese patients with COPD. We showed that ambient air pollutants ($PM_{2.5}$, PM_{10}, NO_2, SO_2, THC, O_3), epigenetic (*lnc-IL7R*), anthropometric (Age, BMI), lifestyle (smoking history, pack-year), and geospatial components (longitude and latitude of participants' residential addresses) all contribute differentially to emphysema (COPD-E) severity ($\%LAA_{-950insp}$). This is in part consistent with reports by Wang M, et al. suggesting a significant association between observed increase in COPD-E severity over time and baseline concentration of ambient $PM_{2.5}$, NO_X, THC, or O_3 [7].

While we cannot fully explain the inverse correlation between COPD-E status/severity and ambient O_3 concentration in our study, we posit that this may be corollary to the significantly enhanced concentration of $PM_{2.5}$ in the study sites, and this rationalization would be consistent with results of a recent study showing that a 40% reduction in $PM_{2.5}$ over a period of 5 years in the North China Plains was in part responsible for a 1–3 ppb annual increase in O_3 observed in megacity clusters of eastern China [22]. It is also possible that the geospatial localization of the study sites allows for enhanced natural influx of "good O_3" from the stratosphere into the troposphere, due to heightened vertical air movements, causing this "good O_3" to contribute immensely to the background concentration of ground-level O_3 in the districts of New Taipei City specified in the present study [23]. Thus,

rather than the increased ambient O_3 concentrations exerting detrimental effects such as "breathing problems, triggering asthma attacks, reducing lung function, and increasing incidence of respiratory diseases" [23], which are characteristic of troposphere-O_3-associated COPD-E, an inversely correlation was found between the O_3 and COPD-E, suggesting a protective effect consistent with those reported by Alberto Hernández's team in the context of COVID-19 [24].

The BMI, indirectly representing an individual's degree of obesity, is a vital indicator or determinant of the phenotypic expression of COPD, its course, and prognosis [25–28]. Our study found that high BMI was associated with less incidence and severity of COPD-E. While this contradicts the prevalent forgone conclusion that obesity or high BMI is associated with disease exacerbation or progression in an array of theme-relevant publications [25–28], our finding is consistent with recent reports that patients with COPD with high BMI exhibit reduced dyspnea symptoms, relatively better lung function, and quality of life, health-wise [28]. More so, our findings corroborate those from a nationwide analysis of the Taiwan Obstructive Lung Disease study data from 12 hospitals in Taiwan, which showed that high BMI (BMI ≥ 24 kgm^{-2}) is associated with a lower frequency of COPD exacerbation in Taiwan [29]. Furthermore, a systematic review of available literature supporting the evolving obesity paradox in COPD, concluded that compared with normal BMI (18.5–24.9 kgm^{-2}), "low BMI is a risk factor for accelerated lung function decline, whilst high BMI has a protective effect" [30].

The present study provides some evidence that BMI, *lnc-IL7R*, $PM_{2.5}$, PM_{10}, and SO_2 levels are excellent classifiers for accurate patient stratification and management triage for COPD-E in Taiwan. Our results revealed that emphysema in patients with COPD is positively correlated with particulate matter and noxious gases exposure, and that high-level exposure to $PM_{2.5}$, PM_{10}, and SO_2 causes greater decline in pulmonary function. Interestingly, we also showed that alongside the ambient pollutants, endogenous *lnc-IL7R* is a highly specific predictors of COPD-E severity and disease progression in New Taipei City. This is corollary to our previously published work indicating that downregulated expression of plasma or tissue *lnc-IL7R* in patients with COPD enhances inflammation and is associated with acute exacerbation, and more so frequently [16]. Moreover, *lnc-IL7R* RNA expression in the serum and lung tissues of patients with COPD-E was positively correlated with BMI, but negatively correlated with $PM_{2.5}$, PM_{10}, and SO_2 exposure. This aligns with the assumption that the large variability in COPD-E onset and progression is driven principally by a compound gene–environment cascade. As succinctly put by Devadoss et al., "the transcriptomic and epigenetic memory potential of lung epithelial and innate immune cells drive responses, such as mucus hyperreactivity and airway remodeling, that are tightly regulated by various molecular mechanisms, for which several candidate susceptibility genes have been described" [31]. We posit that by interacting with and suppressing *lnc-IL7R* expression, the ambient $PM_{2.5}$, PM_{10}, SO_2, facilitates aryl hydrocarbon receptor (AHR)-mediated CYP1A1 activation, enhances generation of reactive oxygen species (ROS), induces oxidative stress with associated inflammatory responses, and consequently elicits chronic inflammatory diseases, including COPD-E [32]. This is corroborated by our data suggesting that *lnc-IL7R* induction may have a potential regulatory role in normal bronchial cells exposed to $PM_{2.5}$ because endogenous *lnc-IL7R* expression was upregulated in normal but not COPD lung epithelial cells or PBMC, and lower *lnc-IL7R* expression was associated with emphysema in COPD (unpublished data).

Furthermore, our finding indicating that Low BMI, and *lnc-IL7R*, with concomitant high $PM_{2.5}$, and SO_2 levels is pathognomonic of exacerbated/severe COPD-E in New Taipei City, Taiwan, corroborates the conclusions of a recent comprehensive analysis of two large cohorts that a panel of disease-related variables improve predictive value for disease outcomes, compared with lone clinical variables and individual biomarkers [33].

Against the background of the findings documented herein, we are cognizant of and do point out that there are suggestions that chronic or longterm exposure, rather than momentary exposure to ambient air pollution significantly affects the incidence and

prevalence of both emphysematous and non-emphysematous COPD, however, such reports remain inconclusive, howbeit with plausible biological mechanisms [34]. Consistent with conclusions drawn by Tamara Schikowski and her team [34], there is probable substantial evidence for a causal link between ambient air pollution and development of COPD, to the extent that diminished pulmonary function earlier in life translates into COPD with age. While we concur that a case for causality may be substantiated should the critical role of ambient air pollutants in repeated or cumulative exacerbation be considered in the development of COPD-E [3,5,10,34,35], the extent to which such cumulative or long-term exposure to ambient air pollution outweighs current or short-term exposure, in the development of COPD-E remains largely inconclusive.

Limitations

As with studies of this nature, the present study has some limitations. First, the use of a relatively small sample size (n = 168) and single-center nature of the study may harbor suggestions of high variability and low reliability of the findings reported herein. A larger cohort from a multi-center setting with varied characteristics may be warranted for more accurate representation of the disease population, and for deriving robustly generalizable inferences. This will help to establish more accurate prediction tools and clinical decision support systems for COPD-E management. Secondly, considering the heterogeneous nature of the classifiers, namely ambient air pollutants ($PM_{2.5}$, PM_{10}, NO_2, SO_2, THC, O_3), epigenetic (*lnc-IL7R*), anthropometric (Age, BMI), lifestyle (smoking history, pack-year), and geospatial components (longitude and latitude of participants' residential addresses), the inclusion of a bridging factor such as blood oxidative stress marker may have been appropriate for mechanistic insight; however, this was omitted. Studies on this bridging factor is currently ongoing.

5. Conclusions

In conclusion, combined BMI, *lnc-IL7R*, $PM_{2.5}$, PM_{10}, and SO_2 levels are optimal classifiers for accurate patient stratification and management triage for COPD-E in Taiwan. Low BMI, and *lnc-IL7R*, with concomitant high $PM_{2.5}$, PM_{10}, and SO_2 levels is pathognomonic of exacerbated/aggravated COPD-E in Taiwan. These findings may help inform management efforts and environmental health policy formulation for lowering disease risk and severity in Taiwan.

Abbreviations

%LAA$_{-950insp}$: percentage of lung attenuation area with values less than$-$ 950 Hounsfield Units on inspiratory CT scan; AQM: air quality monitoring; AUC: area under the curve; BMI: body mass index; COPD: chronic obstructive pulmonary disease; COPD-E: emphysema in patients with COPD; CT: computed tomography; FEV1%: percent predicted forced expiratory volume in 1 s; FVC: predicted forced vital capacity; GOLD: Global Initiative for Chronic Obstructive Lung Disease; lncRNA: long non-coding RNA; lnc-IL7R: long non-coding interleukin-7 receptor α-subunit gene; MMEF: maximum mid-expiratory flow; PBMC: peripheral blood mononuclear cells; PM: particulate matter; NMHC: non-methane hydrocarbons; THC: total hydrocarbon content.

References

1. Muneswarao, J.; Verma, A.K.; Hassali, M.A.A. Global initiative for chronic obstructive lung disease (GOLD) 2018 report: Highlighting an incorrect information. *Pulm. Pharmacol. Ther.* **2018**, *49*, 10. [CrossRef]
2. Halpin, D.M.G.; Criner, G.J.; Papi, A.; Singh, D.; Anzueto, A.; Martinez, F.J.; Agusti, A.A.; Vogelmeier, C.F. Global Initiative for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease. The 2020 GOLD Science Committee Report on COVID-19 and Chronic Obstructive Pulmonary Disease. *Am. J. Respir. Crit. Care Med.* **2021**, *203*, 24–36. [CrossRef]
3. Celli, B.R.; Locantore, N.; Tal-Singer, R.; Riley, J.; Miller, B.; Vestbo, J.; Yates, J.C.; Silverman, E.K.; Owen, C.A.; Divo, M.; et al. ECLIPSE Study Investigators. Emphysema and extrapulmonary tissue loss in COPD: A multi-organ loss of tissue phenotype. *Eur. Respir. J.* **2018**, *51*, 1702146. [CrossRef] [PubMed]
4. Shah, P.L.; Herth, F.J.; van Geffen, W.H.; Deslee, G.; Slebos, D.J. Lung volume reduction for emphysema. *Lancet Respir. Med.* **2017**, *5*, 147–156. [CrossRef]
5. Global Initiative for Chronic Obstructive Lung Disease Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease Global Initiative for Chronic Obstructive Lung Disease; 2020 [Updated 4 November 2020]. Available online: http://www.goldcopd.org/ (accessed on 12 December 2020).
6. Guo, C.; Zhang, Z.; Lau, A.K.H.; Lin, C.Q.; Chuang, Y.C.; Chan, J.; Jiang, W.K.; Tam, T.; Yeoh, E.K.; Chan, T.C.; et al. Effect of long-term exposure to fine particulate matter on lung function decline and risk of chronic obstructive pulmonary disease in Taiwan: A longitudinal, cohort study. *Lancet Planet. Health* **2018**, *2*, e114–e125. [CrossRef]
7. Wang, M.; Aaron, C.P.; Madrigano, J.; Hoffman, E.A.; Angelini, E.; Yang, J.; Laine, A.; Vetterli, T.M.; Kinney, P.L.; Sampson, P.D.; et al. Association between Long-term Exposure to Ambient Air Pollution and Change in Quantitatively Assessed Emphysema and Lung Function. *JAMA* **2019**, *322*, 546–556. [CrossRef]
8. Yang, X.; Zhang, T.; Zhang, Y.; Chen, H.; Sang, S. Global burden of COPD attributable to ambient PM2.5 in 204 countries and territories, 1990 to 2019: A systematic analysis for the Global Burden of Disease Study 2019. *Sci. Total Environ.* **2021**, *796*, 148819. [CrossRef]
9. Agusti, A.; Calverley, P.M.; Celli, B.; Coxson, H.O.; Edwards, L.D.; Lomas, D.A.; MacNee, W.; Miller, B.E.; Rennard, S.; Silverman, E.K.; et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints (ECLIPSE) investigators. Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respir. Res.* **2010**, *11*, 122. [CrossRef]
10. Barnes, P.J. Chapter 47—Genomics in the Evaluation and Management of Chronic Obstructive Pulmonary Disease. In *Essentials of Genomic and Personalized Medicine*; Ginsburg, G.S., Willard, H.F., Eds.; Academic Press: Cambridge, MA, USA, 2010; pp. 603–615. [CrossRef]
11. Wu, D.D.; Song, J.; Bartel, S.; Krauss-Etschmann, S.; Rots, M.G.; Hylkema, M.N. The potential for targeted rewriting of epigenetic marks in COPD as a new therapeutic approach. *Pharmacol. Ther.* **2018**, *182*, 1–14. [CrossRef]
12. Ruvuna, L.; Sood, A. Epidemiology of Chronic Obstructive Pulmonary Disease. *Clin. Chest Med.* **2020**, *41*, 315–327. [CrossRef] [PubMed]
13. Tan, W.C.; Sin, D.D.; Bourbeau, J.; Hernandez, P.; Chapman, K.R.; Cowie, R.; FitzGerald, J.M.; Marciniuk, D.D.; Maltais, F.; Buist, A.S.; et al. Characteristics of COPD in never-smokers and ever-smokers in the general population: Results from the CanCOLD study. *Thorax* **2015**, *70*, 822–829. [CrossRef] [PubMed]
14. Schamberger, A.C.; Mise, N.; Meiners, S.; Eickelberg, O. Epigenetic mechanisms in COPD: Implications for pathogenesis and drug discovery. *Expert Opin. Drug Discov.* **2014**, *9*, 609–628. [CrossRef] [PubMed]

15. Zeng, H.; Li, T.; He, X.; Cai, S.; Luo, H.; Chen, P.; Chen, Y. Oxidative stress mediates the apoptosis and epigenetic modification of the Bcl-2 promoter via DNMT1 in a cigarette smoke-induced emphysema model. *Respir. Res.* **2020**, *21*, 229. [CrossRef]

16. Wu, S.M.; Feng, P.H.; Chuang, H.C.; Ho, S.C.; Fan, C.K.; Chen, K.Y.; Wu, G.S.; Chen, T.T.; Tseng, C.H.; Liu, W.T.; et al. Impaired lnc-IL7R modulatory mechanism of Toll-like receptors is associated with an exacerbator phenotype of chronic obstructive pulmonary disease. *FASEB J.* **2020**, *34*, 13317–13332. [CrossRef]

17. Occhipinti, M.; Paoletti, M.; Bartholmai, B.J.; Rajagopalan, S.; Karwoski, R.A.; Nardi, C.; Inchingolo, R.; Larici, A.R.; Camiciottoli, G.; Lavorini, F.; et al. Spirometric assessment of emphysema presence and severity as measured by quantitative CT and CT-based radiomics in COPD. *Respir. Res.* **2019**, *20*, 101. [CrossRef] [PubMed]

18. Chen, T.T.; Wu, S.M.; Ho, S.C.; Chuang, H.C.; Liu, C.Y.; Chan, Y.F.; Kuo, L.W.; Feng, P.H.; Liu, W.T.; Chen, K.Y.; et al. SUV39H1 Reduction Is Implicated in Abnormal Inflammation in COPD. *Sci. Rep.* **2017**, *7*, 46667. [CrossRef] [PubMed]

19. Masuda, N.; Sakaki, M.; Ezaki, T.; Watanabe, T. Clustering Coefficients for Correlation Networks. *Front. Neuroinform.* **2018**, *12*, 7. [CrossRef]

20. Clemente, G.P.; Grassi, R. Directed clustering in weighted networks: A new perspective. *Chaos Soliton Fract.* **2018**, *107*, 26–38. [CrossRef]

21. Pepelyshev, A. The Role of the Nugget Term in the Gaussian Process Method. In *mODa 9—Advances in Model-Oriented Design and Analysis. Contributions to Statistics*; Giovagnoli, A., Atkinson, A., Torsney, B., May, C., Eds.; Physica HD: Geneva, Switzerland, 2010. [CrossRef]

22. Li, K.; Jacob, D.J.; Liao, H.; Shen, L.; Zhang, Q.; Bates, K.H. Anthropogenic drivers of 2013–2017 trends in summer surface ozone in China. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 422–427. [CrossRef]

23. Zhang, J.J.; Wei, Y.; Fang, Z. Ozone Pollution: A Major Health Hazard Worldwide. *Front. Immunol.* **2019**, *10*, 2518. [CrossRef]

24. Hernández, A.; Viñals, M.; Isidoro, T.; Vilás, F. Potential Role of Oxygen-Ozone Therapy in Treatment of COVID-19 Pneumonia. *Am. J. Case Rep.* **2020**, *21*, e925849. [CrossRef]

25. Celli, B.R.; Cote, C.G.; Marin, J.M.; Casanova, C.; Montes de Oca, M.; Mendez, R.A.; Pinto Plata, V.; Cabral, H.J. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N. Engl. J. Med.* **2004**, *350*, 1005–1012. [CrossRef]

26. Vestbo, J.; Hurd, S.S.; Agustí, A.G.; Jones, P.W.; Vogelmeier, C.; Anzueto, A.; Barnes, P.J.; Fabbri, L.M.; Martinez, F.J.; Nishimura, M.; et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am. J. Respir. Crit. Care Med.* **2013**, *187*, 347–365. [CrossRef]

27. Rennard, S.I.; Locantore, N.; Delafont, B.; Tal-Singer, R.; Silverman, E.K.; Vestbo, J.; Miller, B.E.; Bakke, P.; Celli, B.; Calverley, P.M.; et al. Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints. Identification of five chronic obstructive pulmonary disease subgroups with different prognoses in the ECLIPSE cohort using cluster analysis. *Ann. Am. Thorac. Soc.* **2015**, *12*, 303–312. [CrossRef] [PubMed]

28. Lim, J.U.; Lee, J.H.; Kim, J.S.; Hwang, Y.I.; Kim, T.H.; Lim, S.Y.; Yoo, K.H.; Jung, K.S.; Kim, Y.K.; Rhee, C.K. Comparison of World Health Organization and Asia-Pacific body mass index classifications in COPD patients. *Int. J. Chronic Obstr. Pulm. Dis.* **2017**, *12*, 2465–2475. [CrossRef]

29. Wei, Y.F.; Tsai, Y.H.; Wang, C.C.; Kuo, P.H. Impact of overweight and obesity on acute exacerbations of COPD—Subgroup analysis of the Taiwan Obstructive Lung Disease cohort. *Int. J. Chronic Obstr. Pulm. Dis.* **2017**, *12*, 2723–2729. [CrossRef] [PubMed]

30. Sun, Y.; Milne, S.; Jaw, J.E.; Yang, C.X.; Xu, F.; Li, X.; Obeidat, M.; Sin, D.D. BMI is associated with FEV1 decline in chronic obstructive pulmonary disease: A meta-analysis of clinical trials. *Respir. Res.* **2019**, *20*, 236. [CrossRef] [PubMed]

31. Devadoss, D.; Long, C.; Langley, R.J.; Manevski, M.; Nair, M.; Campos, M.A.; Borchert, G.; Rahman, I.; Chand, H.S. Long Noncoding Transcriptome in Chronic Obstructive Pulmonary Disease. *Am. J. Respir. Cell Mol. Biol.* **2019**, *61*, 678–688. [CrossRef] [PubMed]

32. Vogel, C.F.A.; Van Winkle, L.S.; Esser, C.; Haarmann-Stemmann, T. The aryl hydrocarbon receptor as a target of environmental stressors—Implications for pollution mediated stress and inflammatory responses. *Redox Biol.* **2020**, *34*, 101530. [CrossRef]

33. Zemans, R.L.; Jacobson, S.; Keene, J.; Kechris, K.; Miller, B.E.; Tal-Singer, R.; Bowler, R.P. Multiple biomarkers predict disease severity, progression and mortality in COPD. *Respir. Res.* **2017**, *18*, 117. [CrossRef] [PubMed]

34. Schikowski, T.; Mills, I.C.; Anderson, H.R.; Cohen, A.; Hansell, A.; Kauffmann, F.; Krämer, U.; Marcon, A.; Perez, L.; Sunyer, J.; et al. Ambient air pollution: A cause of COPD? *Eur. Respir. J.* **2014**, *43*, 250–263. [CrossRef] [PubMed]

35. Kwon, S.O.; Hong, S.H.; Han, Y.J.; Bak, S.H.; Kim, J.; Lee, M.K.; London, S.J.; Kim, W.J.; Kim, S.Y. Long-term exposure to PM10 and NO₂ in relation to lung function and imaging phenotypes in a COPD cohort. *Respir. Res.* **2020**, *21*, 247. [CrossRef] [PubMed]