“High-Risk” Clinical and Inflammatory Clusters in COPD of Chinese Descent

Pei Yee Tiew, MD; Fanny Wai San Ko, MD; Jayanth Kumar Narayana, BS, MS; Mau Ern Poh, MD; Huiying Xu, MD; Han Yee Neo, MD; Li-Cher Loh, MD; Choo Khoo Ong, MD; Micheál Mac Aogáin, PhD; Jessica Han Ying Tan, MD; Nabilah Husna Kamaruddin; Gerald Jiong Hui Sim; Therese S. Lapperre, MD, PhD; Mariko Siyue Koh, MD; David Shu Cheong Hui, MD; John Arputhan Abisheganaden, MD; Augustine Tee, MD; Krasimira Tsaneva-Atanasova, PhD; and Sanjay H. Chotirmall, MD, PhD

BACKGROUND: COPD is a heterogeneous disease demonstrating inter-individual variation. A high COPD prevalence in Chinese populations is described, but little is known about disease clusters and prognostic outcomes in the Chinese population across Southeast Asia. We aim to determine if clusters of Chinese patients with COPD exist and their association with systemic inflammation and clinical outcomes.

RESEARCH QUESTION: We aim to determine if clusters of Chinese patients with COPD exist and their association with clinical outcomes and inflammation.

STUDY DESIGN AND METHODS: Chinese patients with stable COPD were prospectively recruited into two cohorts (derivation and validation) from six hospitals across three Southeast Asian countries (Singapore, Malaysia, and Hong Kong; $n = 1,480$). Each patient was followed more than 2 years. Clinical data (including co-morbidities) were employed in unsupervised hierarchical clustering (followed by validation) to determine the existence of patient clusters and their prognostic outcome. Accompanying systemic cytokine assessments were performed in a subset ($n = 336$) of patients with COPD to determine if inflammatory patterns and associated networks characterized the derived clusters.

RESULTS: Five patient clusters were identified including: (1) ex-TB, (2) diabetic, (3) low comorbidity: low-risk, (4) low comorbidity: high-risk, and (5) cardiovascular. The cardiovascular and ex-TB clusters demonstrate highest mortality (independent of Global Initiative for Chronic Obstructive Lung Disease assessment) and illustrate diverse cytokine patterns with complex inflammatory networks.

INTERPRETATION: We describe clusters of Chinese patients with COPD, two of which represent high-risk clusters. The cardiovascular and ex-TB patient clusters exhibit high mortality, significant inflammation, and complex cytokine networks. Clinical and inflammatory risk stratification of Chinese patients with COPD should be considered for targeted intervention to improve disease outcomes.

CHEST 2020; 158(1):145-156

KEY WORDS: cardiovascular; Chinese; COPD; mortality; TB

FOR EDITORIAL COMMENT, SEE PAGE 11

ABBREVIATIONS: CVD = cardiovascular; GOLD = Global Initiative for Chronic Obstructive Lung Disease; HR = hazard ratio; LCHR = low-comorbidity high risk; LCLR = low comorbidity low risk; PDGF = platelet-derived growth factor; RDA = regularized discriminant analysis; TNF = tumor necrosis factor; VEGF = vascular endothelial growth factor

AFFILIATIONS: From the Lee Kong Chian School of Medicine (Drs Tiew, Mac Aogáin, and Chotirmall; and Mr Narayana), Nanyang Technological University, Singapore; Department of Respiratory and Critical Care Medicine (Drs Tiew, Lapperre, and Koh), Singapore General Hospital, Singapore; Department of Medicine and Therapeutics (Drs Ko and Hui), The Chinese University of Hong Kong, Hong Kong; Indian Institute of Science Education and

COPD in the Asian subcontinent is underestimated owing to poor disease awareness and significantly delayed diagnosis.^{1,2} COPD burden in Asia is expected to further increase, driven by aging populations, increased tobacco consumption, and rapid urbanization.³⁻⁵

Although an inimitable group of COPD-associated risk factors exists in Asia, a diverse range of geographic environments, climates, cultural practices, health-care policies, and resource availability influences diagnosis and management.⁶ Taken together, this contributes to clinically different COPD phenotypes in the region, for instance, higher in men, nonsmokers and fewer symptoms compared with western cohorts.^{7,8} Prior work focused on racial differences includes few Asian-based patients and is predominantly conducted outside Asia.⁹⁻¹¹

Work on ethnic Chinese populations reveals the prevalence and scale of COPD.¹² The China Pulmonary Health study evaluated a nationally representative sample of adults across mainland China. Of the 57,779 individuals studied, COPD prevalence was 8.6%,

accounting for almost 100 million individuals, a clear public health priority.¹² A second study involving 67,752 Chinese adults confirms these findings and estimates COPD burden at 13.6%.¹ Because of large COPD numbers, it is plausible that disease subgroups with differing prognostic outcomes exist necessitating individualized intervention. Significant numbers of ethnically Chinese individuals reside in countries across the Asian subcontinent including Malaysia, Singapore, and Hong Kong and few published data have assessed COPD in these populations. The limited data do suggest high mortality in Chinese patients with COPD, however.¹³

In view of high COPD burden and mortality in Chinese patients, a need to better understand potential disease clusters (or subgroups) exists to allow improved risk stratification and targeted intervention. Here, in a multicenter study across three countries in Southeast Asia, we evaluate a large group of Chinese patients with COPD and describe clusters with prognostic and inflammatory relevance.

Research (Mr Narayana), Pune, India; Department of Medicine (Drs Poh and Kamaruddin), University of Malaya, Kuala Lumpur, Malaysia; Department of Respiratory and Critical Care Medicine (Drs Xu, Neo, and Abisheganaden), Tan Tock Seng Hospital, Singapore; Department of Medicine (Drs Loh and Ong), RCSI-UCD Malaysia Campus, Georgetown, Penang, Malaysia; Department of General Medicine (Dr Tan), Sengkang General Hospital, Singapore; Department of Respiratory and Critical Care Medicine (Drs Sim and Tee), Changi General Hospital, Singapore; Department of Respiratory Medicine (Dr Laperre), Bispebjerg University Hospital, Copenhagen, Denmark; Living Systems Institute and Department of Mathematics (Dr Tsaneva-Atanasova), College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK; and PSRC Centre for Predictive Modelling in Healthcare (Dr Tsaneva-Atanasova), University of Exeter, Exeter, UK.

FUNDING/SUPPORT: This research is supported by the Singapore Ministry of Health's National Medical Research Council under its Research Training Fellowship [NMRC/Fellowship/0049/2017 to P. Y. T.] and a Clinician-Scientist Individual Research Grant [MOH-000141 to S. H. C.]; the Singapore Ministry of Education under its Singapore Ministry of Education Academic Research Fund Tier 1 [2016-T1-001-050 to S. H. C.]; the NTU Integrated Medical, Biological and Environmental Life Sciences (NIMBELS), Nanyang Technological University, Singapore [NIM/03/2018 to S. H. C.]; the Ageing Research Institute for Society and Education (ARISE), Nanyang Technological University, Singapore [ARISE/2017/6 to S. H. C.]; and Engineering and Physical Sciences Research Council (EPSRC) grant, United Kingdom [EP/N014391/1] to K. T. A.

Parts of this article have been presented in the form of an abstract at the European Respiratory Society Annual Congress 2019, Madrid, Spain.

CORRESPONDENCE TO: Sanjay H. Chotirmall, MD, PhD, Lee Kong Chian School of Medicine, Nanyang Technological University, 11 Mandalay Road, Singapore 308232; e-mail: schotirmall@ntu.edu.sg

Copyright © 2020 American College of Chest Physicians. Published by Elsevier Inc. All rights reserved.

DOI: <https://doi.org/10.1016/j.chest.2020.01.043>

Methods

COPD Patient Recruitment

Patients of Chinese ethnicity (defined as an individual where both parents were of Chinese lineage) aged ≥ 40 with stable COPD as their predominant diagnosis were prospectively recruited over a 7-year period from January 2012 through December 2018 when attending respiratory outpatient clinics at six tertiary hospital sites across three countries: (1) Singapore (three sites: Singapore General Hospital, Changi General Hospital, and Tan Tock Seng Hospital); (2) Malaysia (two sites: RCSI-UCD Malaysia Campus and University Malaya Medical Centre); and (3) Hong Kong (one site: Prince of Wales Hospital). COPD was defined according to the global initiative for chronic obstructive lung disease (Global Initiative for Chronic Obstructive Lung Disease [GOLD]) criteria.¹⁴ Bronchiectasis was excluded by chest radiography in the absence of tram tracking, ring opacities, and tubular structures.^{15,16} Disease stability was defined as the absence of exacerbation over the preceding 6 weeks before study recruitment and all patients were receiving COPD therapy (including smoking cessation counseling, inhaler assessment, COPD action plans, inhalers as long-acting β -agonists, long-acting muscarinic antagonists, inhaled corticosteroids, and/or short-acting bronchodilators in addition to vaccination as appropriate) on the basis of GOLD guidelines.¹⁴

Derivation and Validation COPD Cohorts

A complete cohort of $n = 1,480$ patients with COPD of Chinese ethnicity were recruited into the study in two separate arms: a derivation and validation cohort. A total of $n = 911$ patients made up the derivation cohort and were recruited from four different sites over the study period from 2012: Singapore General Hospital and Changi General Hospital (Singapore), RCSI-UCD Malaysia Campus (Malaysia), and Prince of Wales Hospital (Hong Kong). An independent and unrelated validation cohort of $n = 569$ patients were recruited from five different sites from 2013 onward: Singapore

General Hospital, Tan Tock Seng Hospital, Changi General Hospital (Singapore), University Malaya Hospital (Malaysia), and Prince of Wales Hospital (Hong Kong). Clinical data were obtained for all subjects and the institutional review board of all participating hospitals approved the study.

Clustering Analysis

Clinical variables and comorbidities were preprocessed with nonmetric multidimensional scaling followed by hierarchical clustering. A trained

regularized discriminant analysis (RDA) model was used to assign cluster membership for the validation cohort. Each derived cluster was defined on the basis of predominant (or lack of) clinical features, comorbidities, and outcome (mortality).

Full details on ethical approvals, patient recruitment, clinical data collation, specimen collection and processing, inflammatory assessments, and statistics are provided in the [e-Appendixes 1-3](#).

Results

Independent unrelated derivation ($n = 911$) and validation ($n = 569$) cohorts of Chinese patients with COPD were recruited. Demographic profiles between cohorts were comparable and similar proportions of patients (from each country) formed the final cohorts ([e-Table 1](#)). The majority of patients (in both cohorts) were men, and, although more current smokers and ex-smokers were identified in the derivation and validation cohorts respectively ($P < .001$), no significant difference in overall smoking pack-years between cohorts was observed ([e-Table 1](#)). Greater proportions of prior pulmonary TB (18.2% vs 8.4%; $P < .001$), osteoporosis (3.7% vs 1.4%; $P < .01$), and asthma (7.1% vs 1.9%; $P < .001$) were identified in the derivation, whereas more malignancy (8.0% vs 17.2%; $P < .001$) was detected in the validation cohort ([e-Fig 1A](#)). Prior pulmonary TB was highest in the Malaysian derivation and lowest in the Hong Kong validation cohort ($P < .001$) reflective of contrasting TB

prevalence ([e-Fig 1B](#)). Coronary artery disease was highest in both cohorts from Singapore, whereas peptic ulcer disease was greatest in the Malaysian derivation and lowest in their validation cohort. Airway microbiology demonstrates more *Streptococcus pneumoniae* and *Haemophilus influenzae* in the derivation cohort ([e-Fig 1C](#)).

Unsupervised hierarchical clustering of the derivation ($n = 911$) cohort revealed five clusters of COPD in the Chinese each defined by their predominant (or lack of) clinical features: (1) prior pulmonary TB (ex-TB; $n = 156$); (2) coexisting diabetes (diabetic; $n = 109$); (3) low comorbidity: low risk (LCLR; $n = 192$); (4) low-comorbidity: high risk (LCHR; $n = 339$); and (5) coexisting cardiovascular disease (CVD; $n = 115$) ([Fig 1](#) and [e-Table 2](#)). Although two clusters demonstrate low comorbidity, the LCHR and CVD clusters closely resembled one another, separated only by dendrogram branching ([Fig 1](#)). Prognostic outcome

Figure 1 – Unsupervised clustering (of the derivation cohort) reveals five clinically relevant patient clusters of ethnic Chinese patients with COPD demonstrating variable prognostic outcome. Dendrogram illustrating the five derived COPD clusters using nonmetric multidimensional scaling followed by hierarchical clustering. Different clusters are represented by colors: ex-TB (red), diabetic (blue), low comorbidity: low-risk (gray), low comorbidity: high-risk (orange), and cardiovascular (light blue).

between clusters varied on the basis of 2-year all cause and respiratory-related mortality (Fig 2A-C). Two-year mortality was highest in the CVD (42.6%), ex-TB (34.0%), and LCHR (25.7%) clusters (log rank test, $P < .05$) (Fig 2A), which remained significant after adjustment for age, sex, BMI, FEV₁, and smoking pack-year exposure (Fig 2B). Hazard ratio for death in each cluster (compared with LCLR) was as follows: CVD (hazard ratio [HR], 2.94; 95% CI, 1.51-5.72; $P < .01$), ex-TB (HR, 2.10; 95% CI, 1.16-3.80; $P < .05$), LCHR (HR, 2.01; 95% CI, 1.18-3.43; $P = .01$), and diabetic (HR, 1.67; 95% CI, 0.82-3.40; $P = \text{NS}$). When only respiratory-related causes of death are considered, a similar pattern (to that with all-cause mortality) is observed (Fig 2C). Smoking status had no influence on mortality (e-Fig 2A) or exacerbation severity (e-Fig 2B) within each cluster however did differ between the low comorbidity clusters with predominantly current smokers in the LCHR and ex-smokers in the LCLR clusters (e-Table 2).

Having identified five clinical clusters demonstrating different prognostic outcomes, we next validated these findings in an independently recruited validation cohort ($n = 569$). All patients in the validation cohort were assigned to a cluster using RDA with a high leave-one-out cross-validation accuracy (97.8%) illustrating robustness of both our model and the previously derived clusters (e-Table 3 and e-Fig 3). The validation cohort was assigned as follows: ex-TB ($n = 102$; 17.9%), diabetic ($n = 72$; 12.7%), LCLR ($n = 88$; 15.5%), LCHR ($n = 193$; 33.9%), and CVD ($n = 114$; 20.0%). The mean RDA assigned probability of a patient from the validation cohort belonging to each of the derived clusters is as follows: ex-TB (mean, 93% $\pm$ SD 13%), diabetic (mean, 96% $\pm$ SD 8%), CVD (mean, 99% $\pm$ SD 6%), LCHR (mean, 88% $\pm$ SD 17%), and LCLR (mean, 69% $\pm$ SD 13%).

Overall proportion of patients in each cluster was comparable to the derivation cohort despite a lower

Figure 2 – A, Kaplan-Meier curve demonstrating survival differences between clusters for 2-year all-cause mortality: worst prognosis in the cardiovascular and ex-TB (TB) clusters that (B) remains significant after adjustment for age, sex, BMI, FEV₁, and smoking pack-year exposure illustrated as Cox regression survival curves. (C) Cumulative incidence curves for each of the five derived COPD clusters for respiratory causes of death: greatest incidence of respiratory cause of death observed in the cardiovascular, ex-TB, and low comorbidity: high-risk clusters. Different clusters are represented by colors: Ex-TB (red), diabetic (blue), low comorbidity: low-risk (gray), low comorbidity: high-risk (orange), and cardiovascular (light blue). See Figure 1 legend for expansion of abbreviations.

overall prevalence of comorbidities (including TB) (e-Table 2 and e-Table 4). Baseline comorbidities (Fig 3) and 2-year all-cause mortality followed the derivation cohort with poorest survival observed in the CVD (43.0%) and ex-TB (26.0%) clusters (log rank test, $P = .001$) (Fig 4A), which remained significant after adjustment for age, sex, BMI, FEV₁, and smoking pack-year exposure (Fig 4B). HR for death in each cluster (compared with LCLR) was as follows: CVD (HR, 3.08; 95% CI, 1.74-5.44; $P < .0001$), ex-TB (HR, 2.01; 95% CI, 1.07-3.80; $P < .05$), LCHR (HR, 1.59; 95% CI, 0.89-2.85; $P = \text{NS}$), and diabetic (HR, 1.73; 95% CI, 0.86-3.45; $P = \text{NS}$). Where respiratory-related causes of death are considered, the CVD, ex-TB, and LCHR clusters again illustrate the highest risk (Fig 4C). Smoking status, as in the derivation cohort, did not influence mortality (e-Fig 2c) or exacerbation severity (e-Fig 2d) in any cluster, although the predominance of current smokers in the LCHR and ex-smokers in the LCLR clusters was reproduced (e-Table 4). Finally, to further verify the accuracy obtained in clustering the validation cohort using RDA leave-one-out cross-validation, we generated a decision tree to classify

Chinese individuals with COPD into one of our proposed five clusters (e-Fig 4). This alternate classification methodology produced accuracy results of 72.4%, comparable to the RDA classification accuracies reported previously.

Following an unbiased semiquantitative cytokine-array screen (evaluating 120 cytokines; data not shown), six cytokines (tumor necrosis factor [TNF]-R1, TNF-R2, vascular endothelial growth factor [VEGF], platelet-derived growth factor [PDGF]-AA, PDGF-BB, and PDGF-AB) were selected for confirmatory validation between clusters and compared with a group of non-COPD (healthy) control patients ($n = 24$) (e-Table 5). Independent of cluster, elevated TNF-R2 significantly associates with symptoms and severe exacerbations (e-Fig 5). Individual cytokines relating to each cluster were as follows: TNF-R2 and PDGF-AA in the CVD cluster; VEGF in the ex-TB cluster; and PDGF-AB and PDGF-BB in the LCHR cluster (Fig 5A; e-Table 6). Given observed differences for individual cytokines between clusters, we next assessed how cytokines interact within an inflammatory network and their respective complexity (Fig 5B). The CVD cluster (highest

Figure 3 – Validation cohort (V; $n = 569$) of patients with COPD of Chinese ethnicity illustrates comparable comorbidity profiles by cluster when compared with the derivation (D; $n = 911$) cohort. Bubble size corresponds to the percentage of patients demonstrating each comorbidity within their respective cohort and bubble color represents cluster membership: ex-TB (red), diabetic (blue), low comorbidity: low-risk (gray), low comorbidity: high-risk (orange), and cardiovascular (light blue). AF = atrial fibrillation; Ca = lung, esophageal, pancreatic, or breast carcinoma; CAD = coronary artery disease; CHF = congestive heart failure; CKD = chronic kidney disease; CVA = cerebrovascular disease; DM = diabetes mellitus; other Ca = all other malignancies excluding lung, esophageal, pancreatic, or breast carcinoma; PAD = peripheral arterial disease; pTB = history of prior pulmonary TB; PUD = peptic ulcer disease.

Figure 4 – Survival outcomes between the identified clusters in the validation cohort for 2-year all-cause mortality (as demonstrated independently in the derivation (D; $n = 911$) cohort; see Fig 2). A, Kaplan-Meier curves demonstrating worst prognosis in the cardiovascular and ex-TB clusters, which (B) remains significant after adjustment for age, sex, BMI, FEV_1 , and smoking pack-year exposure illustrated as Cox regression survival curves. C, Comparable cumulative incidence curves for each of the five derived COPD clusters (as demonstrated independently in the derivation (D; $n = 911$) cohort; see Fig 2) for respiratory causes of death: greatest incidence of respiratory cause of death observed in the cardiovascular, ex-TB, and low comorbidity: high-risk clusters. Different clusters are represented by color: ex-TB (red), diabetic (blue), low comorbidity: low-risk (gray), low comorbidity: high-risk (orange), and cardiovascular (light blue). See Figure 1 legend for expansion of abbreviations.

mortality) exhibits the most complex network with the highest number of positive cytokines and cytokine interactions. In line with the observed mortality in the derivation and validation cohorts, the ex-TB cluster followed by the LCHR, diabetic, and LCLR, respectively, demonstrate gradients of decreasing cytokine network complexity (Fig 5B).

We next assessed our clusters in comparison to GOLD ABCD group and conventional GOLD staging. All patients were assigned to their respective GOLD grouping: A (17.0%; $n = 252$), B (29.1%; $n = 430$), C (12.9%; $n = 191$), and D (41.0%; $n = 607$) at study enrollment. All patients were also classified by conventional GOLD grade (FEV_1 criteria) as follows: I (6.4%; $n = 95$), II (34.5%; $n = 510$), III (43.9%; $n = 650$), and IV (15.2%; $n = 225$).¹⁴ Mortality at 2-year follow-up was assessed by cluster membership within each GOLD group (Fig 6) and GOLD grade (Fig 7). Across all GOLD groups and grades, the CVD and ex-TB clusters

demonstrate highest mortality (Fig 6A, Fig 7). On univariate analysis, stratifying patients into GOLD groups, CVD and ex-TB clusters demonstrate highest mortality in GOLD A, C, and D, whereas CVD and LCHR in GOLD B (Fig 6A). CVD and ex-TB clusters demonstrate significantly higher mortality in multivariate logistic regression after adjustment for age, sex, BMI, FEV_1 , smoking pack-year exposure, and GOLD group compared with LCLR cluster (Fig 6B). Adjusted odds ratios for mortality in the CVD and ex-TB clusters (compared with LCLR group) were 2.98 (95% CI, 1.88-4.73; $P < .001$) and 1.852 (95% CI, 1.17-2.92; $P < .01$), respectively (Fig 6B). Similar trends were seen with adjusted hazard ratios (e-Fig 6). FEV_1 was not significant in independently predicting mortality in our logistic regression model. When the clusters were stratified by GOLD grade, significantly poorer 2-year survival was observed in the CVD and ex-TB clusters irrespective of underlying (conventional) GOLD grade (Fig 7; $P < .05$). Taken

A

B

Figure 5 – The five derived COPD clusters of Chinese ethnicity illustrate inflammatory signatures that associate with all cause and respiratory-related mortality. A, Radar plot illustrating variation in systemic cytokine profile between each of the five derived COPD clusters. The median normalized score for each selected cytokine (in the final assessed panel) is plotted on the radar chart for comparison between clusters. Each dot represents the median normalized value of the cytokine and the color indicates the specific cluster: ex-TB (red), diabetic (blue), low comorbidity: low-risk (gray), low comorbidity: high-risk (orange), and cardiovascular (light blue). PDGF = platelet-derived growth factor; TNF-R1 = tumor necrosis factor receptor 1; TNF-R2 = tumor necrosis factor receptor 2; VEGF = vascular endothelial growth factor. * $P \leq 0.05$, # $P < .1$. B, Network plots demonstrating inflammatory grids detected within each cluster. Increased cytokine interaction and a greater number of positive cytokines are detectable in the cardiovascular, ex-TB and low comorbidity: high-risk clusters compared with the diabetic and low comorbidity: low-risk clusters (indicated by greater than [$>$] symbols). Each circle (node) represents a cytokine from the final assessed panel: circle size corresponds to percentage of patients within that cluster demonstrating a positive value. Lines connecting two nodes indicate positive detection of both cytokines and line thickness illustrates the proportion of patients with a positive result for both cytokines. Node color corresponds to the respective clinical cluster: ex-TB (red), diabetic (blue), low comorbidity: low-risk (gray), low comorbidity: high-risk (orange), and cardiovascular (light blue).

together, these data suggest that the CVD and ex-TB clusters perform poorly despite classification as low risk by conventional GOLD group and staging necessitating early identification in Chinese populations.

Discussion

In this multicenter study across Southeast Asia, we evaluated 1,480 Chinese patients with COPD and describe five validated clusters with prognostic relevance. The two highest-risk clusters were CVD and ex-TB, which demonstrate high mortality risk. Our cluster classification demonstrates differences in mortality outcome and associates with inflammatory signatures and cytokine network complexity.

CVD and diabetes are well-recognized COPD comorbidities and therefore identification of these

clusters was foreseen in view of existing evidence.^{17–22} In the Asian subcontinent however, rapid urbanization with improved socioeconomic status has resulted in higher risks of cardiovascular consequences in COPD, with poorer prognosis, a finding consistent in our study.^{23,24} The highest mortality risk in our CVD cluster warrants attention because prior Asian data suggest the undertreatment of COPD with coexisting CVD.²⁵ Several studies report increased mortality with concomitant diabetes in COPD.^{20,26} Interestingly, however, in our work, the presence of diabetes illustrated better prognosis compared with the high-risk CVD and ex-TB clusters, with some reports suggesting ethnic differences in diabetes-related mortality in South Asians and in particular Chinese.²⁷ Two low comorbidity clusters were identified that differed in mortality outcomes. The LCHR cluster contained high proportions of current smokers in contrast to the LCLR

Figure 6 – Cardiovascular and ex-TB clusters illustrate highest 2-year mortality. A, Tree map illustrating cluster-related mortality within each GOLD groups A through D. Rectangles represent the proportion of deceased patients within each group, presented as percentages. Rectangle color indicates cluster membership: ex-TB (red), diabetic (blue), low comorbidity: low-risk (gray), low comorbidity: high-risk (orange), and cardiovascular (light blue). B, The mortality differences remain significant after adjustment for age, sex, BMI, smoking pack-year exposure, lung function (by FEV₁), and GOLD group illustrated by forest plot using multivariate logistic regression. The dot represents the OR with color indicating significance levels: green ($P < .05$), gray ($P > .05$; NS). Error bar indicates the 95% CI. See Figure 1 legend for expansion of abbreviations.

cluster with predominance of ex-smokers illustrating the benefits of smoking cessation to the natural course of COPD and its outcomes.^{28,29}

Unlike some of the identified clusters, the high-risk ex-TB cluster is novel and likely unique to Asians and other regions where TB is endemic. Patients in this cluster completed treatment with clinical and microbiological resolution; however, recognized long-term sequelae of TB such as the high-risk of subsequent pulmonary obstruction persist.^{30,31} Post-TB-related airway disease is commonly recognized in Asia; however, precise mechanisms remain unclear.³¹⁻³³ Structural lung damage as a consequence of TB with pulmonary cavitation, bronchiectasis, and endobronchial disease associates with increased risks of airflow obstruction, and, even with minimal change on chest radiography,

risks of obstruction persist suggesting alternate mechanisms contributing to COPD development.^{34,35} One possibility is that pulmonary TB “primes” lung host defenses dysregulating responses to inhaled toxins, pathogens, and cigarette smoke and increasing susceptibility to chronic lung disease despite infection clearance.^{35,36} Once chronic lung disease such as COPD has developed post-TB, host genetics and dysregulated immunity may play further roles in determining disease trajectory, progression, and outcome illustrating why many patients with post-TB-related chronic lung disease demonstrate aggressive clinical phenotypes with poorer outcome.³⁵ A prior history of TB associates with morbidity and mortality in COPD, findings consistent with our work.³⁷ Our clinical and inflammatory analyses further demonstrate complex networks and interactions in the ex-TB cluster, highlighting the possibility of

Figure 7 – Cardiovascular and ex-TB clusters demonstrate poorest 2-year survival irrespective of underlying COPD grade. Kaplan-Meier curves illustrating 2-year mortality of each cluster by conventional COPD staging (defined by FEV₁). *P ≤ .05, **P = .001 by log-rank test. Different clusters are represented by colors: ex-TB (red), diabetic (blue), low comorbidity: low-risk (gray), low comorbidity: high-risk (orange), and cardiovascular (light blue). See Figure 1 legend for expansion of abbreviations.

ongoing systemic inflammation postinfection (some of which may be attributed to higher exacerbation rates) contributing to their poorer outcome. Patient ethnicity influences TB-induced inflammation and therefore is of relevance of COPD in the Chinese.^{38,39}

Low-grade systemic inflammation is reported in COPD; however, only a defined group of patients exhibit a persistent systemic inflammatory state with high mortality.⁴⁰ Additionally, COPD can demonstrate Th-2 inflammatory responses with coexisting airway diseases such as asthma in asthma-COPD overlap syndrome.⁴¹ Systemic inflammation in COPD is complex and involves interactions between multiple cytokines and their pathways.⁴⁰ Using network analyses, we illustrate cytokine interactions occurring within each cluster. Interestingly, network complexity correlates with cluster mortality. Systemic inflammation is established in cardiovascular disease and when coexisting with COPD likely contributes to even greater levels in this cluster, where TNF is identified. TNF promotes tissue inflammation, injury, and reactive oxygen species, which in turn drives poorer prognostic outcomes. VEGF, an angiogenesis promoter, is described to have paradoxical roles in

COPD: decreased in emphysematous and increased in bronchitis phenotypes. Interestingly, we detected significant VEGF levels in our ex-TB cluster. An established body of evidence indicates that VEGF plays an essential role in TB pathogenesis enhancing granulomatous inflammation and its associated angiogenesis.^{42,43} VEGF increases in active TB, correlating with disease severity and although its levels are thought to improve posttreatment, its precise role in ex-TB COPD is unclear.^{44,45} The role of platelets and their functional consequence in COPD is of interest.^{46,47} PDGF-related cytokines associate with the CVD and LCHR clusters and have roles in cell signaling, lung development, and cardiopulmonary disease.⁴⁸ As a potent stimulant of smooth muscle proliferation, PDGF associates with small airway remodeling^{49,50} and, although its role in COPD is uncertain, our findings suggest it to be an important systemic marker (perhaps associated with active smoking) for consideration in future studies in Chinese patients with COPD.

Our derived clusters demonstrate differences in mortality outcome despite conventional stratification by GOLD approaches used for COPD grouping and

staging. This is considering our highest-risk clusters: CVD and ex-TB, which demonstrate high mortality and inflammatory complexity unrelated to underlying GOLD group or grade. Importantly, this indicates that some Chinese patients with COPD within each GOLD group and grade (including that defined as low risk) do poorly and require attention. Furthermore, our defined clusters did not differ on the basis of clinical COPD features alone (symptoms and lung function), with the exception of ex-TB group with consistently low BMIs. A comprehensive tool incorporating cardiovascular and TB assessment may be of value to improve risk stratification in Chinese COPD populations.

Here, we describe five clusters of Chinese COPD, at least three of which (cardiovascular, diabetic/metabolic, and low comorbidity) are broadly consistent with comparable studies in non-Chinese COPD populations.¹⁷⁻²¹ This work, unlike prior studies, did not identify any one cluster enriched for anxiety that overall demonstrates low prevalence in our dataset, potentially explained by used diagnostic criteria. Similarly, we did not identify an underweight cluster; however, did observe lower BMIs in the ex-TB group, a novel cluster with high mortality and relevant in any country with high TB prevalence. Understanding pathophysiological mechanisms in this cluster is an important avenue for future studies.

Our work demonstrates clear strengths: it is the largest multicenter study to cluster Chinese patients with COPD using strict diagnostic criteria including longitudinal follow-up and assessment of separate validation and derivation cohorts.⁵¹ We used robust statistical models to derive and validate clusters, adjusted for potential confounders, and demonstrate high predictive accuracies. Despite strengths, our study does have limitations including being restricted to tertiary hospitals making data less generalizable to community-based patients with milder disease. More than 90% of our COPD cohort was men, and although this is largely representative of the COPD population across Asia,^{7,8} it restricts applicability of our findings to women. It remains unclear and not addressed by this work whether Chinese women smoke less or have different susceptibilities to developing COPD. Although all patients had chest radiography demonstrating no

evidence of coexisting bronchiectasis, gold standard imaging of chest CT was not available in all patients, hence potential particularly TB-related sequelae may not have been identified. Furthermore, the precise timing of a TB diagnosis was not available in all patients; hence, we cannot fully differentiate post-TB fixed airflow limitation from COPD in appropriate patients. Although we present a simplified Classification and Regression Tree model to classify patients into our described five clusters, its accuracy was significantly lower than that from our RDA model, hence further research is clearly necessary before simplified models for patient classification can be routinely implemented into clinical practice. For inflammatory work, we selected a restricted final panel of six cytokines on the basis of strict semiquantitative screening criteria (at least 2.5-fold differences between clusters). If we decreased fold difference cutoffs, we may have identified more cytokines of relevance to our clusters. All clinical data used for derivation of clusters were collated at time of recruitment and therefore stability of the clusters or patient membership within them over time was not examined and is an interesting area for follow-up. Validation of our clusters in Chinese patients with COPD outside Asia (including BODE [BMI, airflow obstruction, dyspnoea, exercise capacity] assessment) and among other ethnic populations (eg, Malays, Indians) within Asia should be pursued. Despite demonstrating the robustness of our clustering algorithm, clustering in COPD does have its limitations and has been previously shown to result in marked heterogeneity.⁵² When comparing our cluster mortality with classic COPD mortality prediction models such as the age, dyspnoea, airflow obstruction index,⁵³ we found no significant differences. Future work should assess the additive value of our derived clusters in Chinese patients with COPD using classic COPD mortality prediction models including BODE.

We describe, for the first time, validated clusters of Chinese patients with COPD including two high-risk patient groups. TB and COPD, although associated, are independent key public health concerns demonstrating an unmet need by our work. Overall, our findings improve risk stratification in Chinese patients with COPD identifying those at highest risk requiring close monitoring and appropriate intervention.

Acknowledgments

Author contributions: P. Y. T and S. H. C. take full responsibility for the content of the manuscript, including the data analysis. P. Y. T.: Study design, patient recruitment and performance of experimental work, data collection, interpretation and analysis including writing of the final manuscript. J. K. N., K. T. A., M. M. A.: data interpretation and statistical analysis. F. W. S. K., H. X., M. E. P., N. H. K., D. S. C. H., A. T.: patient recruitment, clinical data, and specimen collection. H. Y. N., L. C. L., C. K. O., J. H. Y. T., G. J. H. S., T. S. L., M. S. K., J. A. A.: patient recruitment and clinical data collection. S. H. C.: Study design and conception of experiments, obtained study funding, interpretation of results, data analysis and writing of the final manuscript. All authors reviewed and approved the final draft of the manuscript.

Financial/nonfinancial disclosures: None declared.

Other contributions: The authors acknowledge The Academic Respiratory Initiative for Pulmonary Health (TARIPH) for collaboration support.

Additional information: The e-Appendixes, e-Figures, and e-Tables can be found in the Supplemental Materials section of the online article.

References

- Fang L, Gao P, Bao H, et al. Chronic obstructive pulmonary disease in China: a nationwide prevalence study. *Lancet Respir Med*. 2018;6(6):421-430.
- Lim S, Lam DC, Muttalif AR, et al. Impact of chronic obstructive pulmonary disease (COPD) in the Asia-Pacific region: the EPIC Asia population-based survey. *Asia Pac Fam Med*. 2015;14(1):4.
- Teramoto S, Yamamoto H, Yamaguchi Y, Matsuse T, Ouchi Y. Global burden of COPD in Japan and Asia. *Lancet*. 2003;362(9397):1764-1765.
- Oh YM, Bhome AB, Boonsawat W, et al. Characteristics of stable chronic obstructive pulmonary disease patients in the pulmonology clinics of seven Asian cities. *Int J Chron Obstruct Pulmon Dis*. 2013;8:31-39.
- Guo C, Zhang Z, Lau AKH, 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(3):e114-e125.
- Tan WC, Ng TP. COPD in Asia: where East meets West. *Chest*. 2008;133(2):517-527.
- Martin A, Badrick E, Mathur R, Hull S. Effect of ethnicity on the prevalence, severity, and management of COPD in general practice. *Br J Gen Pract*. 2012;62(595):e76-e81.
- Wedzicha JA, Zhong N, Ichinose M, et al. Indacaterol/glycopyrronium versus salmeterol/fluticasone in Asian patients with COPD at a high risk of exacerbations: results from the FLAME study. *Int J Chron Obstruct Pulmon Dis*. 2017;12:339-349.
- Gilkes A, Ashworth M, Schofield P, et al. Does COPD risk vary by ethnicity? A retrospective cross-sectional study. *Int J Chron Obstruct Pulmon Dis*. 2016;11:739-746.
- Tran HN, Siu S, Iribarren C, Udaltsova N, Klatsky AL. Ethnicity and risk of hospitalization for asthma and chronic obstructive pulmonary disease. *Ann Epidemiol*. 2011;21(8):615-622.
- Davis J, Tam E, Taira D. Disparate rates of utilization and progression to combined heart failure and chronic obstructive pulmonary disease among Asians and Pacific Islanders in Hawai'i. *Hawaii J Med Public Health*. 2016;75(8):228-234.
- Wang C, Xu J, Yang L, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. *Lancet*. 2018;391(10131):1706-1717.
- Lopez AD, Shibuya K, Rao C, et al. Chronic obstructive pulmonary disease: current burden and future projections. *Eur Respir J*. 2006;27(2):397-412.
- GOLD. Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD). 2018. Available at: <https://goldcopd.org/wp-content/uploads/2018/02/WMS-GOLD-2018-Feb-Final-to-print-v2.pdf>. Accessed April 12, 2020.
- Cantin L, Bankier AA, Eisenberg RL. Bronchiectasis. *AJR Am J Roentgenol*. 2009;193(3):W158-W171.
- Milliron B, Henry TS, Veeraraghavan S, Little BP. Bronchiectasis: mechanisms and imaging clues of associated common and uncommon diseases. *Radiographics*. 2015;35(4):1011-1130.
- Vanfleteren LE, Spruit MA, Groenen M, et al. Clusters of comorbidities based on validated objective measurements and systemic inflammation in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2013;187(7):728-735.
- Chubachi S, Sato M, Kameyama N, et al. Identification of five clusters of comorbidities in a longitudinal Japanese chronic obstructive pulmonary disease cohort. *Respir Med*. 2016;117:272-279.
- Garcia-Aymerich J, Gomez FP, Benet M, et al. Identification and prospective validation of clinically relevant chronic obstructive pulmonary disease (COPD) subtypes. *Thorax*. 2011;66(5):430-437.
- Burgel PR, Paillasseur JL, Janssens W, et al. A simple algorithm for the identification of clinical COPD phenotypes. *Eur Respir J*. 2017;50(5):1701034.
- Raherison C, Ouailaya EH, Bernady A, et al. Comorbidities and COPD severity in a clinic-based cohort. *BMC Pulm Med*. 2018;18(1):117.
- Cavaillès A, Brinchault-Rabin G, Dixmier A, et al. Comorbidities of COPD. *Eur Respir Rev*. 2013;22(130):454-475.
- Rabe KF, Hurst JR, Suissa S. Cardiovascular disease and COPD: dangerous liaisons? *Eur Respir Rev*. 2018;27(149):185057.
- Miller J, Edwards LD, Agusti A, Bakke P, Calverley PM, Celli B, et al. Comorbidity, systemic inflammation and outcomes in the ECLIPSE cohort. *Respir Med*. 2013;107(9):1376-1384.
- Kubota Y, Tay WT, Asai K, et al. Chronic obstructive pulmonary disease and beta-blocker treatment in Asian patients with heart failure. *ESC Heart Fail*. 2018;5(2):297-305.
- Mannino DM, Thorn D, Swensen A, Holguin F. Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in COPD. *Eur Respir J*. 2008;32(4):962-969.
- Khan NA, Wang H, Anand S, Jin Y, Campbell NR, Pilote L, et al. Ethnicity and sex affect diabetes incidence and outcomes. *Diabetes Care*. 2011;34(1):96-101.
- Godtfredsen NS, Lam TH, Hansel TT, et al. COPD-related morbidity and mortality after smoking cessation: status of the evidence. *Eur Respir J*. 2008;32(4):844-853.
- Anthonisen NR, Skeans MA, Wise RA, et al. The effects of a smoking cessation intervention on 14.5-year mortality: a randomized clinical trial. *Ann Intern Med*. 2005;142(4):233-239.
- Menezes AM, Hallal PC, Perez-Padilla R, et al. Tuberculosis and airflow obstruction: evidence from the PLATINO study in Latin America. *Eur Respir J*. 2007;30(6):1180-1185.
- Leung JM, Tiew PY, Mac Aogain M, et al. The role of acute and chronic respiratory colonization and infections in the pathogenesis of COPD. *Respirology*. 2017;22(4):634-650.
- Chotirmall SH, Gellatly SL, Budden KF, et al. Microbiomes in respiratory health and disease: an Asia-Pacific perspective. *Respirology*. 2017;22(2):240-250.
- Budden KF, Shukla SD, Rehman SF, et al. Functional effects of the microbiota in chronic respiratory disease. *Lancet Respir Med*. 2019;7(10):907-920.
- Jung KH, Kim SJ, Shin C, Kim JH. The considerable, often neglected, impact of pulmonary tuberculosis on the prevalence of COPD. *Am J Respir Crit Care Med*. 2008;178(4):431-433.
- Ravimohan S, Kornfeld H, Weissman D, Bisson GP. Tuberculosis and lung damage: from epidemiology to pathophysiology. *Eur Respir Rev*. 2018;27(147):170077.
- O'Leary SM, Coleman MM, Chew WM, et al. Cigarette smoking impairs human pulmonary immunity to *Mycobacterium tuberculosis*. *Am J Respir Crit Care Med*. 2014;190(12):1430-1436.
- Park HJ, Byun MK, Kim HJ, et al. History of pulmonary tuberculosis affects the

- severity and clinical outcomes of COPD. *Respirology*. 2018;23(1):100-106.
38. Coussens AK, Wilkinson RJ, Nikolayevskyy V, et al. Ethnic variation in inflammatory profile in tuberculosis. *PLoS Pathog*. 2013;9(7):e1003468.
 39. Pareek M, Evans J, Innes J, et al. Ethnicity and mycobacterial lineage as determinants of tuberculosis disease phenotype. *Thorax*. 2013;68(3):221-229.
 40. Agusti A, Edwards LD, Rennard SI, et al. Persistent systemic inflammation is associated with poor clinical outcomes in COPD: a novel phenotype. *PLoS One*. 2012;7(5):e37483.
 41. Poh TY, Mac Aogain M, Chan AK, Yui AC, Yong VF, Tiew PY, et al. Understanding COPD-overlap syndromes. *Expert Rev Respir Med*. 2017;11(4):285-298.
 42. Harding JS, Herbath M, Chen Y, et al. VEGF-A from granuloma macrophages regulates granulomatous inflammation by a non-angiogenic pathway during mycobacterial infection. *Cell Rep*. 2019;27(7):2119-2131.
 43. Datta M, Via LE, Kamoun WS, et al. Anti-vascular endothelial growth factor treatment normalizes tuberculosis granuloma vasculature and improves small molecule delivery. *Proc Natl Acad Sci U S A*. 2015;112(6):1827-1832.
 44. Matsuyama W, Hashiguchi T, Matsumuro K, et al. Increased serum level of vascular endothelial growth factor in pulmonary tuberculosis. *Am J Respir Crit Care Med*. 2000;162(3 Pt 1):1120-1122.
 45. Kumar NP, Banurekha VV, Nair D, Babu S. Circulating angiogenic factors as biomarkers of disease severity and bacterial burden in pulmonary tuberculosis. *PLoS One*. 2016;11(1):e0146318.
 46. Harrison MT, Short P, Williamson PA, Singanayagam A, Chalmers JD, Schembri S. Thrombocytosis is associated with increased short and long term mortality after exacerbation of chronic obstructive pulmonary disease: a role for antiplatelet therapy? *Thorax*. 2014;69(7):609-615.
 47. Fawzy A, Putcha N, Paulin LM, et al. Association of thrombocytosis with COPD morbidity: the SPIROMICS and COPDGene cohorts. *Respir Res*. 2018;19(1):20.
 48. Noskovicova N, Petrek M, Eickelberg O, Heinzlmann K. Platelet-derived growth factor signaling in the lung. From lung development and disease to clinical studies. *Am J Respir Cell Mol Biol*. 2015;52(3):263-284.
 49. Raines EW. PDGF and cardiovascular disease. *Cytokine Growth Factor Rev*. 2004;15(4):237-254.
 50. Churg A, Tai H, Coulthard T, Wang R, Wright JL. Cigarette smoke drives small airway remodeling by induction of growth factors in the airway wall. *Am J Respir Crit Care Med*. 2006;174(12):1327-1334.
 51. Bhatt SP, Balte PP, Schwartz JE, et al. Discriminative accuracy of FEV1:FVC thresholds for COPD-related hospitalization and mortality. *JAMA*. 2019;321(24):2438-2447.
 52. Castaldi PJ, Benet M, Petersen H, et al. Do COPD subtypes really exist? COPD heterogeneity and clustering in 10 independent cohorts. *Thorax*. 2017;72(11):998-1006.
 53. Puhan MA, Hansel NN, Sobradillo P, et al. Large-scale international validation of the ADO index in subjects with COPD: an individual subject data analysis of 10 cohorts. *BMJ Open*. 2012;2(6):e002152.