Reviewer A

It is well written and adds interesting data to the current evidence in ILD. The authors aimed to identify novel ILD phenotypes in a multi-ethnic South-East Asian population. Although there are several limitations, the findings are interesting and open new scenarios to investigate, especially in the Asian population.

Few minor comments:

Comment 1: Lines 42-43. I have some concerns about the definition of "non-IPF fibrotic ILD (F-ILD)". In the supplement, authors report "radiological findings as evaluated by a reporting thoracic radiologist and clinician during multi-disciplinary discussion were determined to be fibrotic or non-fibrotic, and then categorised […]. Only 15% of non-IPF had a UIP pattern, while the ground glass was more represented. This is also reflected by the difference in treatment, immunomodulator vs antifibrotic. The adjective "fibrotic" could be misleading, as some underlying diseases may have a major inflammatory component. "Non-IPF" ILD could be a better term if not all the patients had clear radiological signs or histology highlighting fibrosis.

Reviewer 1 Reply 1: We thank the reviewer for the comment. In our study, we included all subjects who had any evidence of fibrotic change radiologically (Figure 1) and have amended the manuscript to, "There were 305 ILD patients recruited, of which 287 had fibrotic ILD radiological changes" (Manuscript, page 6, line 96). We have also revised the methods section in the Supplementary Appendix to clarify this, "Subjects who had no evidence of any fibrotic change were excluded from the study." (Supplementary Appendix Methods, page 3, lines 56-67)

We agree that some diseases may have a major inflammatory component, particularly in those with groundglass, and one limitation of our study is that the extent of fibrosis was not quantified. We have added this to our discussion as follows, "The extent of fibrosis radiologically was not quantified and hence identifying subjects that would benefit from antifibrotic therapy was also limited." (Manuscript, page 12, lines 245-246)

Changes in the text: Revised sentence to, "There were 305 ILD patients recruited, of which 287 had fibrotic ILD radiological changes" (Manuscript, page 6, line 96).

Addition of the sentence, Subjects who had no evidence of any fibrotic change were excluded from the study." (Supplementary Appendix Methods, page 3, lines 56-67).

Addition of the sentence, "The extent of fibrosis radiologically was not quantified and hence identifying subjects that would benefit from antifibrotic therapy was also limited." (Manuscript, page 12, lines 245-246)

Comment 2: Line 67: please add the section/page of the supplementary material you are referring to.

Reviewer 1 Reply 2: Thank you for the comment, we have amended this by clarifying

that the Methods section of the Supplementary Appendix is referred to. This now reads as, "Both numerical and categorical data were selected for cluster analysis based on clinical relevance and previous literature (Supplementary Material Appendix 1 Methods)."

Changes in the text: Addition of "Methods" (Page 4, Line 62).

Reviewer B

A very creative way to classify the non-IPF fILD according to the local community. It would be interesting if you could include:

Comment 1: HRCT pattern for each cluster which had been discussed in the MDD meeting (UIP pattern in non-IPF cohort is only found in 15 patients)

Reviewer 2 Reply 1: We thank the reviewer for the comment and have included the radiological patterns in Table S1 of the Supplementary Appendix Tables section (Supplementary Appendix Tables, page 13, Table S1).

Changes in the text: Addition of the radiology patterns observed in each of the clusters is in Table S1 (Supplementary Appendix Tables, page 13, Table S1).

Addition of "and radiological patterns" in the manuscript (Manuscript, page 6, line 109).

Comment 2: Perhaps comparing the characteristic of IPF and non-IPF in cluster 1 and 4 (cluster with lowest survival, similar in IPF patients)

Reviewer 2 Reply 2: Thank you for the suggestion. We agree that it would be important to present the comparison of the characteristics of the IPF group against each of the respective clusters and have thus included a summary table of the p-values for the characteristics of each cluster compared against IPF. This is presented as Table S3 in the Supplementary Appendix Tables section on page 15 and is referenced in the manuscript on page 7, lines 133-135.

Changes in the text: Addition of summary table of p-values for characteristics of each cluster compared against IPF (Supplementary Appendix Tables, page 15, Table S3). Addition of the line, "The p-values for each of the cluster's characteristics compared against IPF are summarized in Supplementary Appendix Table S3." (Manuscript, page 7, lines 133-135).

Comment 3: Are cluster 1,4 ILD patients fit in PF-ILD diagnosis?

Reviewer 2 Reply 3: Although we did not examine for PF-ILD based on existing trial criteria, it is most likely that Clusters 1 and 4 subjects would fit in with a diagnosis of PF-ILD, as this is a retrospective diagnosis.

Change in the text: nil.

Comment 4: Cluster 3 patients with more emphysema but FVC, DLco value almost

similar (which I expect Dlco would be even lower in emphysema), the slowest lung function declination received more anti-fibrotic (AF) as compared to cluster 1,4. Any reason for that? Is AF initiation depending on patient's willingness, disease severity and affordability?

Reviewer 2 Reply 4: We thank the reviewer for his comment and agree with his observation that the DLco in Cluster 3 was higher than expected. We did consider that greater antifibrotic use may have been contributory to this, however as only 6/42 patients in this cluster (14.3%) were on antifibrotics, it is difficult to comment on whether this finding is due to antifibrotic use or other factors and hence further study would be required to explore this.

We agree with the reviewer that the cited factors are important considerations when starting antifibrotics. In Singapore, anti-fibrotics cost an average of SGD $3600 per month and financial subsidies are extremely limited. Hence, cost is a major consideration for patients who may decline treatment due to financial difficulties. We have added this important consideration into our discussion section (page, 13 lines 267-269)

Change in text: Addition of the sentence, "In addition, due to cost limitations, such as for the use of antifibrotics, some subjects may have declined treatment and hence deteriorate more rapidly." (Manuscript, Page 13, lines 267-269).

Comment 5: comparing cluster 3 (slow decline) and 4 (fast decline), are the treatment initiated (more AF) in cluster 3 played a role in the progression of the lung function?

Reviewer 2 Reply 5: Thank you for the question. We recognize that this is an important consideration and have performed post-hoc analysis on the annual rate of FVC decline for Clusters 3 and 4 with and without antifibrotic therapy. The results are presented in the manuscript (lines 162-165 on pages 8-9) and in Table S6 (Supplementary Appendix Tables, page 18). We found there was no significant difference in the annual rate of FVC decline between subjects who received antifibrotic therapy and those who did not, between both clusters. However, this is limited by the small number of subjects on antifibrotic therapy.

Changes in the text: Addition of the sentence, "Post-hoc analysis showed that there was no significant difference in the rate of FVC decline in subjects who received anti-fibrotic therapy and those who did not, for both Clusters 3 and 4. (Supplementary Appendix Table S6)." (Manuscript, pages 8-9, lines 162-165).

Addition of Table S6 (Supplementary Appendix, Tables, page 18).

Comment 6: This is just a suggestion:
An overview of the cluster phenotype in fILD Singapore cohort (characteristics and each of cluster survival compared to IPF)
Followed by the cluster with similar disease progression like IPF, based on:
a) lung function (FVC, Dlco)
b) demographic
c) the specific diagnosis of ILD (including the subset of CTD)
d) HRCT pattern
e) how many on AF, and actually requiring AF

Reviewer 2 Reply 6: Thank you for the suggestions. The characteristics of IPF and the 4 clusters are summarised in Tables 1 and 2. We agree that it would be important to present the comparison of the characteristics of the IPF group against each of the respective clusters and have thus included a summary table of the p-values for the characteristics of each cluster compared against IPF. This is presented as Table S3 in the Supplementary Appendix Tables section on page 15 and referenced in the manuscript on page 7, lines 133-135.

In Table S4, we compare the annual rate of FVC decline for each of the 4 clusters compared to IPF, which includes that of Cluster 1 (similar disease behaviour to IPF) and Cluster 4 (more aggressive behaviour than IPF). We have also included the radiological patterns and individual CTD diagnoses for each of the clusters in Table S1 of the Supplementary Appendix Tables section (Supplementary Appendix Tables, page 13, Table S1).

We found that survival differences between the respective clusters and the IPF group remained significant after adjustment for treatment. We have also included post-hoc analysis on the annual rate of FVC decline for clusters with and without antifibrotic therapy, although this is limited by the small number of subjects on antifibrotic therapy. We do not have information available on the number of subjects for which antifibrotics were required, however this would be important to explore in further studies.

Changes in text: Addition of summary table of p-values for characteristics of each cluster compared against IPF (Supplementary Appendix Tables, page 15, Table S3). Addition of the line, "The p-values for each of the cluster's characteristics compared against IPF are summarized in Supplementary Appendix Table S3." (Manuscript, page 7, lines 133-135).

Addition of the CTD subtypes and radiology patterns observed in each of the clusters is in Table S1 (Supplementary Appendix Tables, page 13, Table S1). Addition of "and radiological patterns" in the manuscript (Manuscript, page 6, line 110-).

Addition of the sentence, "Post-hoc analysis showed that there was no significant difference in the rate of FVC decline in subjects who received anti-fibrotic therapy and those who did not, for both Clusters 3 and 4. (Supplementary Appendix Table S6)." (Manuscript, pages 8-9, lines 162-165). Addition of Table S6 (Supplementary Appendix, Tables, page 18).

Reviewer C

This study identified the four clusters of fibrotic ILD in South-East Asia and showed different prognosis and lung function trajectories for each cluster. These findings are important for the understanding of the heterogeneity of fibrotic ILD.

Major comments
> Comment 1: Is there a difference in the prevalence of gastroesophageal reflux disease (GERD) among the four clusters? GERD may be associated with high body mass index, high exacerbation rate, disease progression of ILD.

Reviewer 3 Reply 1: We thank the reviewer for the comment and agree that GERD is an important consideration. This was initially omitted as the results were not statistically significant, however we recognize that this would still be of interest in the results and have hence included the results in the Tables section, in Table 2 (page 2, Table 2).

Changes in the text: Addition of row "GERD, Gastritis, Peptic Ulcer Disease -no. (%)" in Table 2 (Tables, page 2, Table 2).

Comment 2: The statistic values should be listed in Table 1. Alternatively, the characteristics among IPF and each Cluster should be listed in one Table. As different clusters have different prognosis, Cluster 1 has similar survival to IPF, as shown in the legends of Figure 3A. Are there any similarities between the characteristics IPF and Cluster 1?

Reviewer 3 Reply 2: We thank the reviewer for the comment and have included a column of p-values in Table 1 (Tables, Page 1, Table 1). We agree that it would be important to present the comparison of the characteristics of the IPF group against each of the respective clusters and have thus included a summary table of the p-values for the characteristics of each cluster compared against IPF. This is presented as Table S3 in the Supplementary Appendix Tables section on page 15 and referenced in the manuscript on page 7, lines 133-135.

Specific to the reviewer's query, when Cluster 1 and IPF were compared, there was no significant difference in age, BMI, proportion of subjects with a high comorbidity burden, ischemic heart disease, GERD, pulmonary hypertension or an ANA titre >1:160 and mean FVC percentage predicted and DLCO percentage predicted. However, there were significant differences in sex, Chinese and Indian ethnicity, smoking status and number of pack years, proportion of subjects with a low comorbidity burden, hypertension, previous history of pulmonary tuberculosis and radiology.

Changes in the text: Addition of p-values column in Table 1 (Tables, Page 1, Table 1). Addition of summary table of p-values for characteristics of each cluster compared against IPF (Supplementary Appendix Tables, page 15, Table S3). Addition of the line, "The p-values for each of the cluster's characteristics compared against IPF are summarized in Supplementary Appendix Table S3." (Manuscript, page 7, lines 133-135).

Comment 3: Table 3. Although only 90 patients (31.4%) had BAL cell counts, the table shows the prevalence of each cellular BAL findings in all patients for each cluster. For example, eosinophilic BAL findings as 5.66% (3/53) of Cluster 1 in table 1, but only 29 patients underwent BAL, so eosinophilic should be 10.3% (3/29). This is misleading in the calculation and should be recalculated and make sure statistical values.

Reviewer 3 Reply 3: Thank you for the comment. We have revised the table accordance with the suggestion. We have recomputed and performed the chi-square tests again, and there is no change in the significance of the statistical values. The revised table is in Supplementary Appendix, under the Tables section, as Table S2 (page 14, Table S2). For greater clarity in presentation due to the different "Total n" values, we have separated the BAL cell count patterns and treatment received into two separate tables, which are Table S2 in the Supplementary Appendix (page 14) and Table 3 in the Tables section (page 3), respectively.

Changes in the text: Table S2 (Supplementary Appendix, page 14, Table S2), Table 3 (Tables, page 3, Table 3).

Comment 4: Table S1 shows CTD-ILD is the major underlying diagnosis in Cluster 2. This should be stated in the results.

Reviewer 3 Reply 4: Thank you for the comment. We agree that is an important finding to state in the results and have revised the manuscript as advised. It now reads, "The predominant diagnosis was CTD-ILD (Supplementary Appendix Table S1)." (Manuscript, Page 6, lines 117-118)

Changes in the text: Addition of the sentence, "The predominant diagnosis was CTD-ILD (Supplementary Appendix Table S1)." (Manuscript, Page 6, lines 117-118)

Minor comments
Comment 5: Table 2. Pulmonary tuberculosis should be classified as previous medical history, not as comorbidity if it was a previous history.

Reviewer 3 Reply 5: We thank the reviewer for the comment and agree with the reviewer. This has been clarified in the Tables section, in Table 1 and Table 2, where it is now described as "Previous History of Pulmonary Tuberculosis" (page 1-2, Table 1, Table 2).

Changes in the text: Renaming of row "Previous History of Pulmonary Tuberculosis – no. (%)"in Table 1 and Table 2 (Tables, Page 1-2, Table 1 and Table 2).

Comment 6: Page 3. STROBE should be spelled out as Strengthening the reporting of observational studies in epidemiology.

Reviewer 3 Reply 6: Thank you for the suggestion. We have made the advised amendment as follows, "This study aims to use cluster analysis to describe clinical phenotypes in a South-East Asian population of non-IPF fibrotic ILD (F-ILD) patients and is presented in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting checklist." (Manuscript, Page 3, Lines 36-39).

Changes in the text: Addition of the phrase "Strengthening the Reporting of Observational Studies in Epidemiology" (Manuscript, Page 3, Lines 38-39).

Comment 7: Page 8. Fig. 4A should be Figure S4A. Figure 5A and 5B should be Figure S5A and S5B, respectively.

Reviewer 3 Reply 7: Thank you for highlighting this. We have made the corrections as advised.

Changes in the text: "Fig. 4A" changed to "Figure S4A" (Manuscript, page 8, line 152). "Figure 5A" changed to "Figure S5A" (Manuscript, page 9, line 167). "Figure 5B" changed to "Figure S5B" (Manuscript, page 9, line 170).

Comment 8: References style does not meet the Journal of Thoracic Disease. For reports

with up to three authors, all the author names should be listed. However, if a report has more than three authors, the first three authors should be listed followed by "et al."

Reviewer 3 Reply 8: Thank you for the comment, we have revised the citation formatting for the references per requirements.

Changes to text: Formatting change for References (pages 19-23, lines 313-422)