Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-2025-556

Reviewer A

Comment 1: This paper reports n=79 patients based on 10 clinical variables. It is not clear that there is enough number of patients overall, or that 10 variables would be permissible for 79 patients. Moreover, the result is that 85% of patients clustered into one group, with the rest divided between two other groups, with n=6 in each group; it is not clear there is enough patients in the clusters 1 and 3 to define these groups. To obtain more patients why not include data from not only 2015-2019 but 2015-2024?

Reply: Thank you very much for your suggestion; it was an excellent one. After collecting the cases from 2019, we began organizing the data and analyzing the results, which led to our current conclusion. We subsequently attempted to validate this conclusion by increasing the sample size. However, the proportion of intermediate-high risk pulmonary embolism cases among all pulmonary embolism cases is not high, so the annual growth in cases is relatively slow. Hierarchical clustering presents all possible cluster divisions through a tree diagram, without the need to pre-specify the number of clusters as in K-means, thus avoiding bias caused by subjective Settings in small samples. Hierarchical clustering can operate stably when the sample size is ≥20 (79 cases are completely sufficient). Therefore, it may take a considerable amount of time to collect and organize the subsequent data. However, this conclusion can still provide a theoretical support for the subsequent research.

Comment 2: On this point, is it standard in China to perform repeat CTPA at 30 days in PE patients? If not, how was the patient population accrued (retrospectively)?

Reply: Thank you for pointing to the question. 30-day CTPA is not routine in China. However, as intermediate-high risk pulmonary embolism patients are currently generally treated at secondary or higher-level medical centers, they routinely undergo CTPA follow-up scans around 30 days after diagnosis, as long as they are willing to comply with follow-up.

Comment 3: Ideally the final analysis would better define the phenotype of a cluster: we are told cluster 3 has higher age, DM, and chronic heart disease but not analogous written information for clusters 1 and 2 (and note this phenotype of cluster 3 is based on n=6 patients). It would be important for authors to interpret Table 1. It seems like cluster 1 has chronic lung disease, but this is a heterogenous category and may include multiple disorders, so this is another problem with n=6 in the cluster.

Reply: We thank the Reviewer for the comment. We have explained the characteristics of different clusters (page7,Line 142-145): All patients had chronic lung disease and a few had diabetes in cluster 1.Cluster 2 was characterized by the absence of chronic lung disease and diabetes in all patients, and about two-thirds of patients had chronic heart disease. Cluster 3 was characterized in all patients with chronic heart disease and diabetes, without chronic lung disease. We agree with your view that chronic lung disease is a heterogeneous category. In our research, chronic lung diseases mainly include chronic obstructive pulmonary disease and interstitial lung disease.

Reviewer B

Comment 1: The sample size is small (n=79), with uneven cluster distribution (e.g., Cluster 1 and 3 each have only 6 cases), limiting statistical power and generalizability. Because of this, I do not think the current data cannot be used for clustering analysis.

Reply: Thanks to the Reviewer's comment. Due to the relatively small proportion of patients with medium to intermediate-high pulmonary embolism among all cases of pulmonary embolism, despite our great efforts in collection, the number of cases is still limited. Therefore, we ultimately adopted the method of hierarchical clustering. The hierarchical cluster analysis utilized is suitable for small sample sizes. Hierarchical clustering presents all possible cluster divisions through a tree diagram, without the need to pre-specify the number of clusters as in K-means, thus avoiding bias caused by subjective Settings in small samples. Hierarchical clustering can operate stably when the sample size is ≥20 (79 cases are completely sufficient). Despite the limited reciprocal size in analyzing intermediate-high-risk pulmonary embolism patients, the results offer potentially meaningful clues, such as differences in clinical characteristics between subgroups and variations in long-term prognosis. This provides theoretical support for subsequent research and verification.

Comment 2: the retrospective design introduces selection bias, and the lack of external validation cohorts hinders the reproducibility of results.

Reply: We appreciate the valuable suggestion. We acknowledge that the retrospective design poses risks of selection bias and external validation is critical for clinical translation. Currently, we are undertaking a multi-center prospective study with a larger sample size, aiming to gather more evidence on potential efficacy differences or similarities in various aspects or prognostic outcomes and verify our results on external dataset. We also look forward to getting more suggestions and Pointers from you in the future.

Comment 3: the study relies on baseline data without long-term follow-up, precluding assessment of long-term outcomes.

Reply: We appreciate the valuable suggestion. We conducted an analysis of the long-term prognosis at 1 year and 3 years. The results showed significant differences in 1-year and 3-year PE mortality among clusters ($p<0.001$, $p=0.003$, respectively). The Fisher's exact test analysis also demonstrated significant differences in 1-year and 3-year PE mortality ($p=0.005$, $p=0.031$, respectively). Following the reviewer's guidance, we plan to conduct more in-depth research and verification.

Comment 4: key variables like lung function and treatment details are missing, potentially affecting cluster interpretation.

Reply: Thank you for pointing to the question. Lung function data were unavailable retrospectively but will be prioritized prospectively. All the patients are under our follow-up. This is a long process. Thank you for your understanding and support. Treatment details (e.g., anticoagulation regimens) are now in Table1.

Comment 5:the use of hierarchical clustering, while suitable for exploration, may not capture all clinical complexities, and the optimal number of clusters requires validation in larger cohorts.

Reply: Thank you for pointing to the question. We agree with the reviewer's opinion. In the study, our choice of hierarchical clustering over alternative algorithms was predicated on its dual advantages in handling limited sample sizes and providing clinically interpretable stratification. Currently, hierarchical clustering was the method we chose after extensively reviewing numerous methodologies and repeatedly comparing options. We believe it is particularly well-suited to our current data type. For less common types of cases, expanding the sample size is indeed challenging. Nevertheless, the ability to derive meaningful results from our work serves as a powerful motivation driving us forward. We truly appreciate your suggestion and sincerely hope to receive more of your insights.

Comment 6: In the abstract, the background did not what risk strata is, the methods did not describe the variables used for the clustering analysis and what the prognosis outcomes are, the results should specify which group had the poorest prognosis, and the conclusion needs to be tone down because this is a failed study of clustering analysis. In the introduction, the authors did not explain why there is a clinical need to regroup the risk of PE and what the limitations of existing three levels of PE are. It is also important to explain why the used variables could generate subgroups with different prognosis.

Reply: We appreciate the valuable suggestions. The abstract and introduction have been revised accordingly.

Comment 7: This rather important for this study. In the methods, the primary research methodology should be corrected describe as a clustering analysis. The authors need to describe the sample size estimation, the consideration for these clinical outcomes, and why the selected variables could generate prognosis subgroups. In statistics, it is wrong to describe "P<0.1 is considered statistical difference". Please ensure P<0.05 is two-sided.

Reply: Thanks to the Reviewer's comment. As a retrospective study, the sample size was not pre-calculated. Hierarchical clustering (agglomerative, average linkage) was selected for its suitability in exploratory phenotyping with limited samples. Variables were chosen based on ESC prognostic indicators and comorbidities linked to PE outcomes. we have corrected statistical threshold: P<0.05 (two-sided) defined significance.

Comment 8: Please cite several related papers: 1. Wang L, Yu C, Hu K, Guo Y, Li Y, Yu S, Wang W. Research progress in interventional therapy for acute intermediate-high-risk and high-risk pulmonary embolism. J Thorac Dis 2024;16(11):7958-7977. doi: 10.21037/jtd-24-1049. 2. Yang G, Nie S. Risk factors for pulmonary embolism: a case-control study. J Thorac Dis 2025;17(3):1552-1560. doi: 10.21037/jtd-24-1293. 3. Qiao Y, Gao Y, Chen Y, Ye X, Yan C, Zeng M. Quantitative assessment and risk stratification of random acute pulmonary embolism cases using a deep learning model based on computed tomography pulmonary angiography images. Quant Imaging Med Surg 2025;15(3):1950-1962. doi: 10.21037/qims-24-1412. 4. Wang L, Yu C, Hu K, Guo Y, Li Y, Yu S, Wang W. Research progress in interventional therapy for acute intermediate-high-risk and high-risk pulmonary embolism. J Thorac Dis 2024;16(11):7958-7977. doi: 10.21037/jtd-24-1049.

Reply: We appreciate the valuable suggestion. The several related papers have been cited in the paper.