A silent revolution is taking place in asthma. Although it has been known to clinicians for a while that asthma presents in various guises and forms, the value of such knowledge has remained very much a pastime of clinicians in categorising asthma. Various clinical forms of asthma have been described such as aspirin-sensitive asthma, late-onset non-asthma with an indolent course, early-onset asthma in childhood and severe asthma. However, only recently, with the application of cluster analysis, a statistical technique to grouping sets of data that have a degree of closeness has there been real progress made in defining phenotypes. Clustering requires approaches that will group a set of objects depending on degree of closeness to each other so that objects in the same group are more similar to each other than those in other groups (or clusters). It is a major test of explanatory data mining, and a common technique for statistical data analysis used in many fields including more recently the determination of phenotypes of asthma dependent on disease characteristics. Of the many clustering algorithms in use, the most commonly applied to asthma phenotyping has been connectivity models such as hierarchical clustering based on distance connectivity or k-means algorithm base by single mean cluster. Thus, in one of the first clustering analyses from the NIH-funded Severe Asthma Research Project (SARP), the use of Ward’s minimum variance hierarchical clustering method using 17 composite clinical and physiological variables derived from 63 separate binary questions applied to an adolescent and adult cohort of mild to severe asthma of 726 subjects led to the definition of five clusters of asthma.¹ In their tree analysis, subjects could be assigned to five clusters that ranged from milder to more severe disease, using only three variables: baseline FEV₁, post-bronchodilator FEV₁, and age of onset of asthma.

In the current issue of the Portuguese Journal of Pulmonology, Loreiro and colleagues report the first cluster analysis applied to a cohort of Portuguese patients using Ward’s clustering method (Loreiro et al. Cluster analysis in phenotyping a Portuguese population. Rev Port Pneumol 2015; in press). They describe five clusters that were differentiated by age of onset of asthma, gender predominance, body mass index, symptoms of asthma, allergic status and presence of eosinophilia. The interesting clusters of the Portuguese cohort are particularly those related to the onset of disease. Early onset clusters include either those with mild allergic asthma or those with allergic brittle asthma, while late onset asthma was associated with severe asthma either with chronic airflow obstruction and eosinophilic inflammation, or highly symptomatic obese women with non-eosinophilic inflammation. One can start to see clusters associated with eosinophilic inflammation, particularly with chronic airflow obstruction.

It is reassuring that these clusters derived from a relatively small number of asthmatic participants (total of 57) do have similarities with those previously described in the SARP cohort and in the Leicester cohort.² In fact, the late onset severe asthma cluster with eosinophilia is also described in a later analysis of SARP that included sputum eosinophilia.³ This reproducibility provides some support to the validity of the clustering approach. It would have been strengthened if these clusters were replicated in a separate independent cohort. Overall, one wonders with the availability of multiple national and international cohorts of asthma whether a large analysis of all these cohorts together to produce a meta-cluster analysis should not be entertained. This would need international collaboration for which there should be great enthusiasm.

The value of these cluster analyses is whether these will help the clinicians in terms of precision/personalised medicine, particularly when it applies to severe asthma.⁴ By defining these five clusters, could we now programme different treatment paradigms specific for each cluster?⁵ However, in order to reach this stage, we need to have some understanding the mechanisms underlying asthma and add mechanistic biomarker analysis to see how these biomarkers distribute across the cluster, which is the process of determining the endotypes of asthma.⁶ For
example, if Cluster 5 of the Portuguese cohort is associated with eosinophilic inflammation, and perhaps part of this population is associated with a Th2 high expression, new treatments based on suppressing activity of Th2 cytokines such as IL-4, IL-5 and IL-13 might be beneficial. On the other hand, Cluster 4 patients could be more appropriately treated by blocking non-Th2 pathways, but these pathways remain to be defined. Much more molecular phenotyping will be needed before reaching this stage of defining clusters with an accurate precision. Clustering with the use of omics data (transcriptome, metabolome, lipidome, microbiome and proteome) is the most promising route of getting to the endotypes that will be more useful in pinpointing the ‘right’ treatment for the ‘right’ patient, an approach that has been taken by the European Unbiased BIOMarkers in PREDiction of respiratory disease outcomes (U-BIOPRED) project.

At the end of the day, with the complexity of the analysis of omics data with clinical and physiological data backed by sophisticated bioinformatics analyses, the objective has to be to provide clinically-relevant phenotypes or more precisely endotypes, that will help optimise and provide targeted treatments, the whole basis of precision/personalised medicine.

References


K.F. Chung

National Heart and Lung Institute, Imperial College
London, London, United Kingdom

E-mail address: f.chung@imperial.ac.uk