Development of allergic response data through childhood

Keywords: Bayesian nonparametric clustering, Metropolis-coupled MCMC, healthcare.

The temporal patterns of allergic sensitisation can now be inspected at a greater resolution as a result of Component Resolved Diagnostics (CRD) (Treudler and Simon, 2013). We are using this data to group similar allergic sensitisation patterns with the aim is to better understand co-sensitisation patterns, relate them to allergy-related disease risk, and inform the development of a more personalised approach to disease diagnosis, management and treatment.

Here, we scale up previous analyses of immune response data (component-specific IgE (sIgE)) to small subsets of the available allergen data by considering the patterns of response to allergens from all available sources, and across six time points from infancy to adolescence. We measured sIgE immune responses in participants from a well-characterised population-based birth cohort at six follow-ups between the ages of 1 to 16 years (ages 1, 3, 5, 8, 11 and 16). We used a Bernoulli mixture model with a Bayesian MCMC algorithm to learn clusters of sIgE components from binarised sensitisation data, i.e. each cluster contains allergen-related components with a similar sensitisation profile across the children. Model parameters and optimal number of clusters were inferred at each age. The flow of allergens between clusters across time is shown in Figure 1, showing clear and consistent patterns in the data, and allergens cluster into increasingly specialised groups according to associated child responses. Though each age was clustered independently of the others, the clusters were biologically meaningful, had exceptionally high mean assignment probabilities, and – as Figure 1 shows – displayed a high degree of consistency and stability across time points.

The cluster-based sensitisation profiles of participants across these ages were then related to asthma and hay fever variables at age 16. When subject responses are stratified appropriately (taking into account the heterogeneous nature of both the subjects and the diseases themselves), the allergic response at age 5 can be strongly associated with the development of asthma and hay fever at age 16. We identified combinations of cluster, time point and degree of cluster sensitisation that were clearly linked to an increased risk of asthma and hay fever development (an example of which is shown in Figure 2), as well as putative “lead” components (e.g. Fel d 1, from cat). Further application of this Bayesian clustering approach to similar data, and the continued exploration of the resulting clusters ought to facilitate the development of better diagnostic and prognostic biomarkers for allergic diseases.

Figure 1: Allergen component clusters derived from the clustering of immune responses.

Figure 2: Incidence of ever having asthma by age 16 modelled on response total to age 5’s “Grass/cat” cluster.