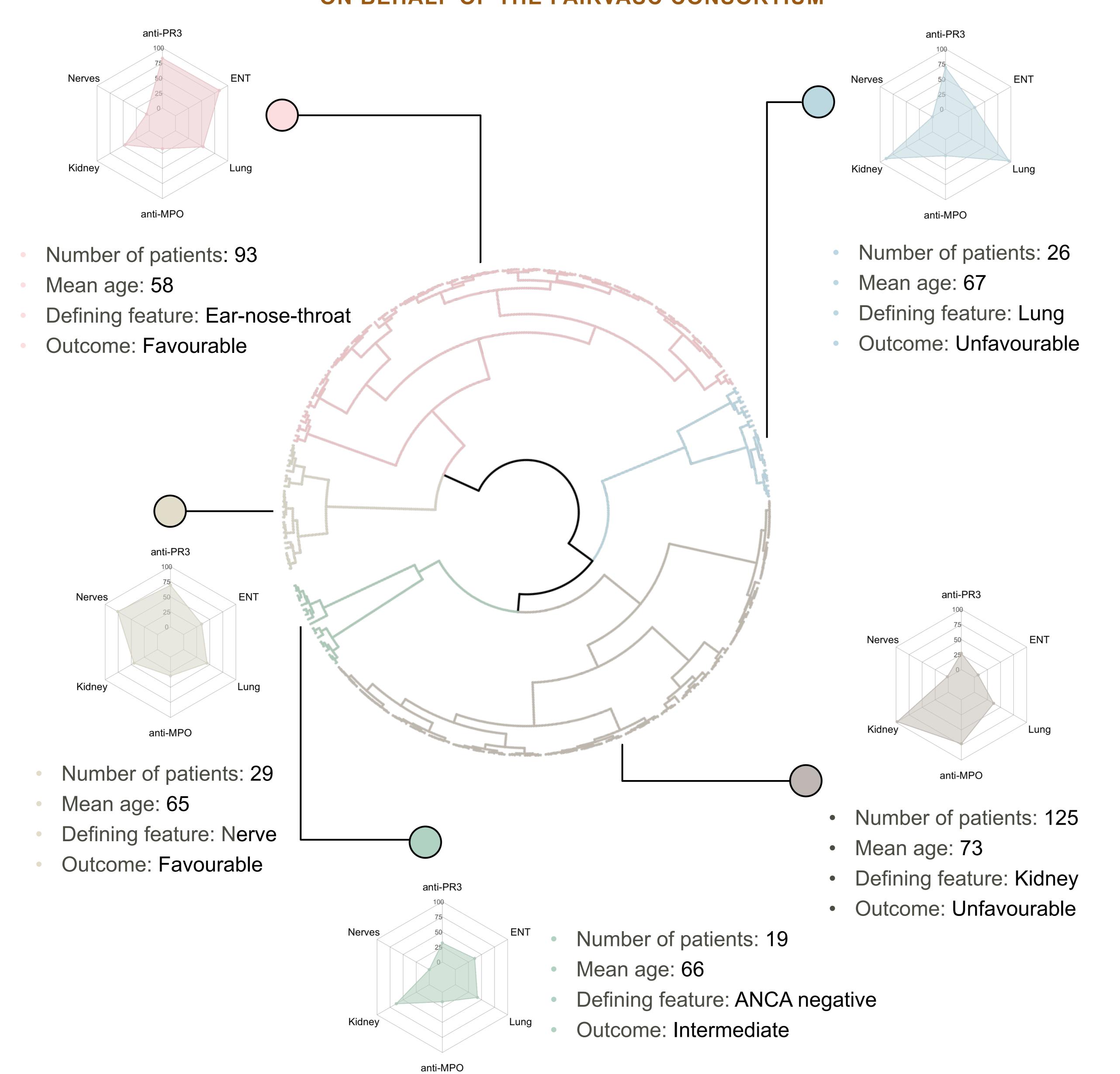


Clustering of anti-neutrophil cytoplasmic antibody-associated vasculitis - using a pre-processed harmonised dataset

KARL GISSLANDER¹, ARTHUR WHITE², MÅRTEN SEGELMARK¹, MARK A. LITTLE², ALADDIN J. MOHAMMAD^{1,3} ON BEHALF OF THE FAIRVASC CONSORTIUM



Background and method

- Dataset of 292 patients from southern Sweden with granulomatosis with polyangiitis (GPA) or microscopic polyangiitis (MPA).
- Pre-processed from a relational database format to a resource descriptive framework (RDF) graph-based data model.
- Harmonised to a FAIRVASC standard to allow the future inclusion of additional AAV patients from the FAIRVASC collaboration registries to the model.
- Factor analysis of mixed data (FAMD) and agglomerative hierarchical cluster analysis on principal components (HCPC) used to develop the cluster model.
- Model including organ pattern, ANCA status, serum creatinine,
 C-reactive protein, gender, and age at diagnosis.

Conclusion

- Our analysis suggests five clusters of AAV patients based on baseline features, associated with different mortality and renal outcome.
- The investigation acts as a proof of concept of the FAIRVASC ontology and infrastructure for the harmonisation of heterogeneous AAV datasets.
- The cluster model may in the future readily include an unprecedented number of European AAV patients (over 6000).

Contact

Karl Gisslander karl.gisslander@med.lu.se

¹ Lund University, Lund, Sweden
 ² Trinity College Dublin, Dublin, Ireland
 ³ University of Cambridge, Cambridge, United Kingdom



