

Using genetics of biomarkers and Bayesian shrinkage prediction to identify genetic factors of giant cell arteritis and its complications

Amara Nwagbata (a.i.nwagbata@ed.ac.uk)¹, Andrii Iakovliev¹, Ann Morgan², James Peters³, Paul McKeigue¹, Athina Spiliopoulou^{1,4}



THE UNIVERSITY of EDINBURGH

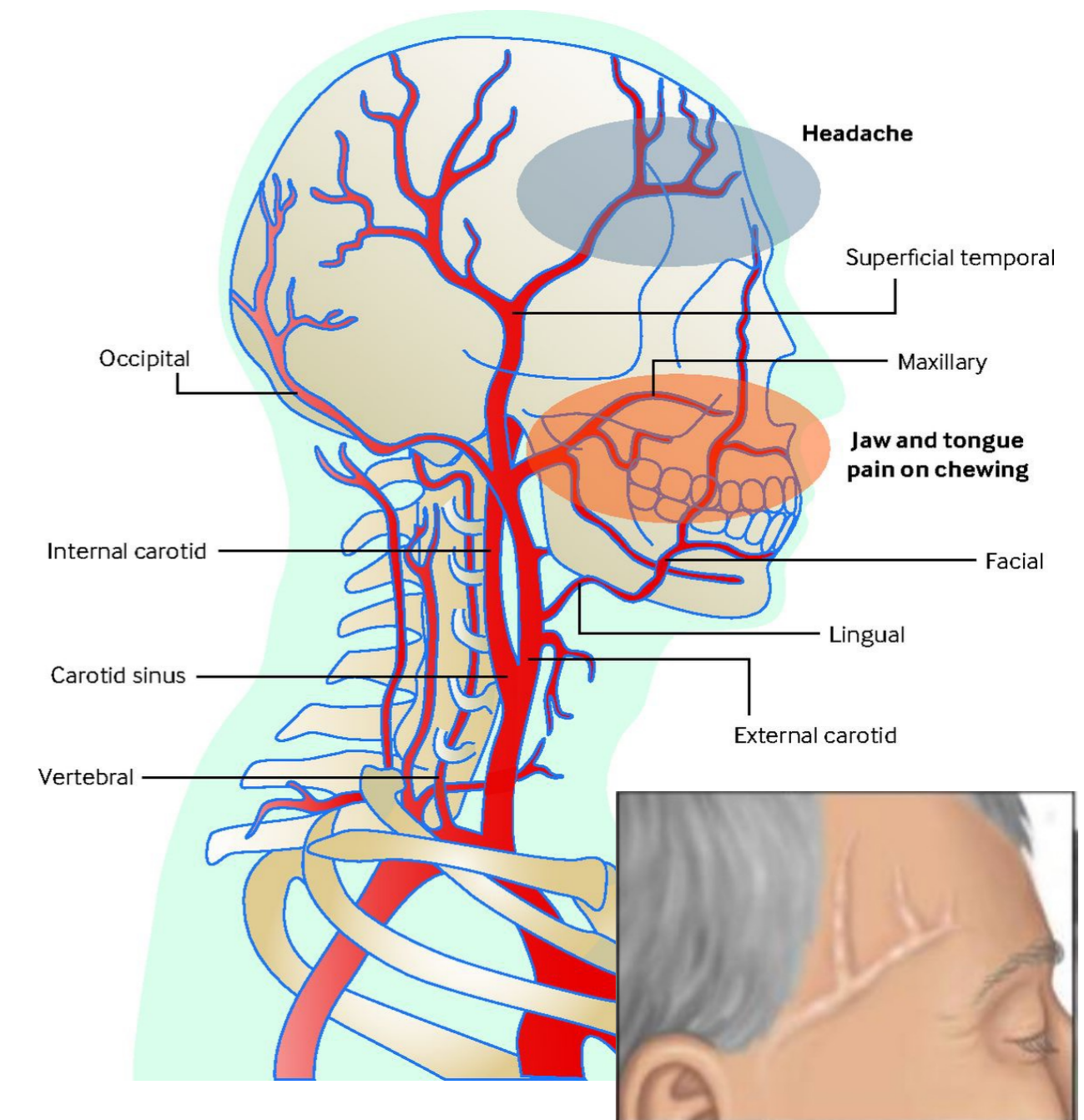
Usher institute

Background

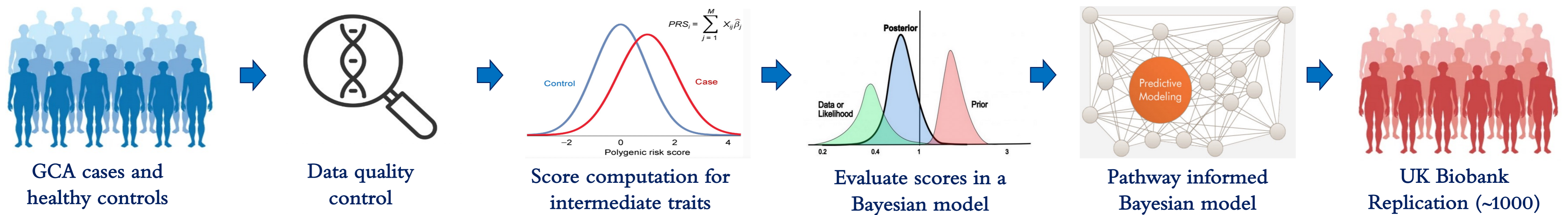
- Giant Cell Arteritis (GCA) is a rare vasculitis characterized by granulomatous inflammation of walls of large- and medium-sized arteries in the body, especially the temporal artery.
- Occurs exclusively in adults over 50 years
- Ischaemic complications of GCA includes blindness, stroke and aortic aneurysm.
- Evidence of genetic predisposition but lacks robust characterization of these genetic factors contributing to disease risk due to lack of statistical power owing to small sample size¹.

Aims

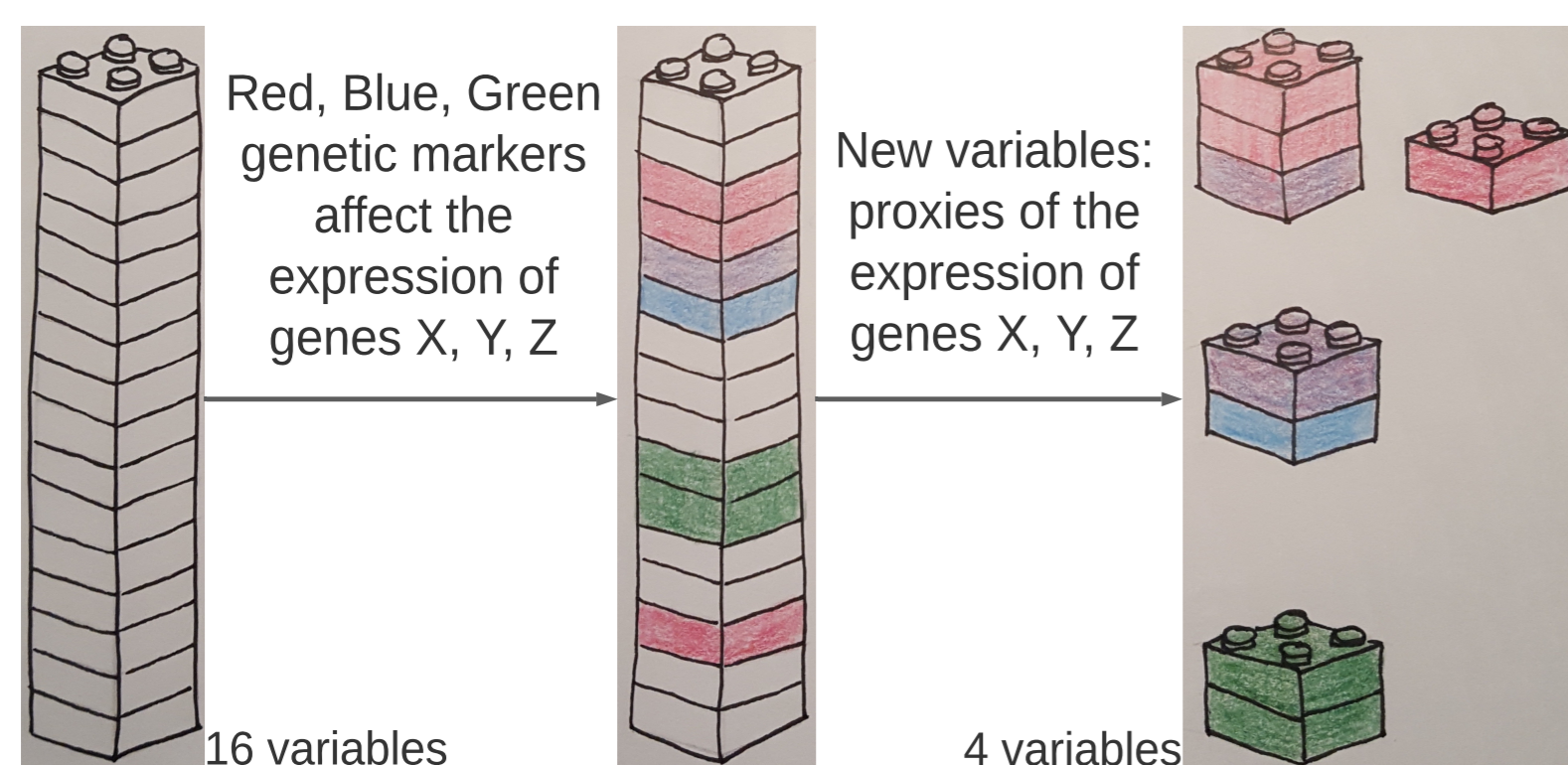
- Discover core genes and pathways involved in the risk of GCA and its complications using relevant intermediate traits (genomic, proteomics, transcriptomics, metabolites) and novel statistical methods.
- Identify molecular markers mediating these genetic effects as they could potentially aid identification of new treatment targets for subsequent translational research



Methods



Step 1: Locus-specific genotypic scores (LocusGRS)



Description: Using summary statistics of relevant intermediate traits, significant loci are selected at a threshold of 10^{-6} and regions of 1mb are created to compute the locus-specific risk scores.

Step 3: Prediction model I

Bayesian hierarchical prediction model

- Posterior means of effect sizes estimated from the regularised horseshoe prior.
- Shrinks small effects to zero leaving out true large effect sizes

Step 2: Score computation workflow

$$\begin{matrix} \text{from input PLINK file} \\ \text{Target genotypes} \end{matrix} \begin{matrix} \text{from reference panel} \\ \text{LD matrix} \end{matrix} \begin{matrix} \text{from DB} \\ \text{GWAS of } t \end{matrix}$$

$$\begin{matrix} \text{id1} & s_1 \\ \text{id2} & s_2 \\ \dots & \dots \end{matrix} = \begin{matrix} \text{rs1} & \text{rs2} & \dots \\ \text{id1} & & \\ \text{id2} & & \\ \dots & & \end{matrix} \times \begin{matrix} \text{rs1} & \text{rs2} & \dots \\ 1 & & \\ \text{rs2} & 1 & \\ \dots & & \end{matrix}^{-1} \times \begin{matrix} \text{rs1} & \beta_1 \\ \text{rs2} & \beta_2 \\ \dots & \dots \end{matrix}$$

s g R^{-1} β_t

Description: β is the weighted effect of each intermediate trait used in the score computation, the target genotypes, g , are the genotypes of the GCA cohort. 1000 genome data is used as LD matrix, R^{-1} .

Step 4: Prediction model II

Bayesian hierarchical model with group based prior

- Pathway-informed hierarchical shrinkage priors.
- Scores grouped based on KEGG/Reactome mappings. If a score in a pathway is selected, then other scores in that pathway are penalised less, with the degree of shrinkage being learned from the data.
- Another prior distribution that shrinks risk scores based on pathway is introduced.

Results

- 716 GCA cases from UKGCA consortium and 2632 controls from WTCCC².
- Cases aged between 50-89 years with mean age of 71 years. 70% female and 99% Europeans. 72% positive biopsy result.
- Approximately 7 million variants after quality control and imputation of genetic dataset.
- Pre-processing of publicly available intermediate trait summary statistics and uploading to GENOSCORES database which has the functionality of computing locusGRS.
- Scores computation with LD adjustment currently being computed on the cluster.
- Next steps, preliminary association of these scores to the case-control outcome of the subjects.

Discussion

- We anticipate to find genetic effects of GCA not found by conventional GWAS due to stringent threshold and small sample size; potentially aiding our understanding of the disease mechanism.
- Complications, severity and response to treatment in GCA can have a major impact on quality of life of patients. Where available, the method would be applied to identify relevant predictive and prognostic biomarkers to aid targeted therapeutic intervention thereby reducing toxicities associated Steroids which is the mainstay treatment for GCA patients.
- The proposed statistical method has been used to identify genetic effects of response to TNF inhibitors in rheumatoid arthritis³.

Affiliations:

- Usher institute, University of Edinburgh.
- Institute of Cardiovascular and Metabolic Medicine, University of Leeds.
- Department of Immunology and Inflammation, Imperial College London.
- Institute of Genetics and Cancer, University of Edinburgh.

References

- Carmona et al. AJHG 2017
- WTCCC. Nature 2017
- Spiliopoulou et al. Rheumatoid arthritis 2019.
- Genoscores: <https://genoscores.cphs.mym.ed.ac.uk/>
- Images: Uffelmann et al, Nature 2021; Lazarewicz K, BMJ 2019



Medical Research Council