**Further Details on Data to be Collected from Primary Care**

Using a combination of diagnostic, symptom and monitoring READ codes, we will identify patients with selected RMD conditions from primary care.

**The RMD conditions of interest** include: Rheumatoid arthritis, psoriatic arthritis, axial spondyloarthritis, gout, juvenile inflammatory arthritis, systemic vasculitis, polymyalgia rheumatica, lupus, systemic sclerosis, scleroderma, Sjogren’s, dermatomyositis, polymyositis, mixed connective tissue disease, osteoarthritis, fibromyalgia/chronic persistent pain, osteoporosis, Paget’s disease, osteogenesis imperfecta, joint hypermobility syndrome, Ehlers-Danlos.

We will also collect socio-demographic data, information on primary care healthcare use and relevant co-morbidity data where available.

**Socio-demographic data**: smoking status, alcohol consumption, BMI, ethnicity, distance from home to GP and nearest specialist service (in mile bands), Scottish Index of Multiple Deprivation (SIMD), rural/urban classification, whether in supported living, care giver, disability status, sickness benefit, FIT note/sickness certificate issued, and employment status.

**Primary care healthcare use data**: date of event, type of healthcare professional seen, type of appointment, time of appointment.

**Presence of relevant co-morbidity data**: congestive heart disease, dementia, depression, diabetes, diabetes with complications, dyslipidaemia, HIV, hypertension, hypothyroidism. liver disease, myocardial infarction, osteoporosis, paraplegia, peptic ulcer disease, peripheral vascular disease, renal disease, valvular disease, venous thromboembolism, electronic frailty index (eFI)

**Data to be collected on shielding and vaccination**: shielding status, influenza vaccination status, pneumococcal vaccination status, COVID-19 1st vaccine, COVID-19 2nd vaccine

Albasoft (a Trusted Third Party provider) will identify eligible patients and extract primary care data required and transfer the data securely to the National Safe Haven. The cohort will then be linked to NHS databases within the National Safe Haven (eDRIS) (outpatient attendances, hospital and out of hours care and community prescribing). We are gathering similar data in Wales in collaboration with the University of Swansea.

This study has been approved by the Public and Privacy Panel for Health and Social Care in Scotland (REF 1819-0286). You do not have to notify individual patients. This is because technical and organisational measures that respect the principle of data minimisation are in place.

Patient identifying information will be removed/pseudonymised and will not be available to researchers. Only named research staff with information governance training will be able to access the pseudonymised dataset within the National Data Safe Haven. All analyses released will meet strict disclosure control requirements. Output tabulations or modelling results will also be checked to ensure that they meet with disclosure requirements. Final outputs, including papers and conference presentations must be cleared by Safe Haven staff before entering the public domain. No individual or small-number outputs that may be identifiable are released.