Reclassifying ANA associated rheumatic musculoskeletal disease, moving towards basket indications for therapeutics

Submitted in accordance with the requirements of the degree of PhD

The University of Leeds School of Medicine

Leeds Institute of Rheumatic and Musculoskeletal Medicine (LIRMM)

Dr Jack Arnold

January 2025

1 DECLARATION

I hereby declare that the work presented in this thesis has not been submitted for any other degree or professional qualification, and that it is the result of my own independent work.

Dr Jack Edward Arnold

21st January 2025

2 ABSTRACT

2.1 BACKGROUND

ANA-associated Rheumatic Musculoskeletal Diseases (ANA-RMDs), including Systemic Lupus Erythematosus, Primary Sjogren's Syndrome, Scleroderma, Myositis, and mixed/undifferentiated connective tissue disease, vary in access to therapeutics despite overlapping clinical/immunological features. These diseases are currently classified according to historical clinical criteria mated to basic laboratory tests. Many patients fulfil multiple diagnostic criteria, and a larger proportion fail to fulfil any criteria at all, being termed undifferentiated disease. These criteria are imperfect and could be replaced with more modern data driven approaches.

2.2 AIMS

This doctorate set out to address the problem of reclassification through several approaches, with increasing complexity. Basic clustering of flow cytometric parameters, derivation of the clinical basket ANA-arthritis, assessment of gene expression-based disease stratification and finally the use of deep learning-based data compression techniques in two large well phenotyped cohorts.

2.3 METHODS

Within this thesis basic clustering, Gaussian mixture modelling, variational autoencoder based data compression and linear mixed effects modelling are used to manipulate and analyse data. This is supported with causal inference approaches including directed acyclic graph analysis. Work is carried out in the multi diagnosis DEFINITION and the European PRECISESADS cohorts.

2.4 RESULTS/CONCLUSIONS

This work defines the first clinical basket indication in ANA-RMDs, the ANA-arthritis basket. It also reclassifies ANA-RMDs into 5 classes using deep learning-based approaches, with differential long-term outcomes. This represents the first application of these approaches within rheumatology and has implications both within the field and more broadly in answering similar questions in other medical specialties

3 INTELLECTUAL PROPERTY AND PUBLICATIONS ASSOCIATED WITH THIS RESEARCH

3.1 INTELLECTUAL PROPERTY AND PUBLICATION STATEMENTS:

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 14: is based on work from a jointly authored abstract by Jack Arnold, Lucy-Marie Carter, Md Yuzaiful Md Yusof, Zoe Wigston, Sabih Ul Hassan, Katie Dutton, Shouvik Dass, Antony Psarras and Edward Vital. The initial concept of the study was set by Professor Vital, processing of flow cytometry samples was carried out by Zoe Wigston. Analysis of flow cytometry data was carried out by Dr Arnold as was database organisation, processing of demographic data, clustering analysis and figure/abstract curation. The completed abstract was authored by Dr Arnold

Chapter 15: is based on work from a jointly authored publication by Jack Arnold, Lucy-Marie Carter, Md Yuzaiful Md Yusof, Katie Dutton, Zoe Wigston, Shouvik Dass, Samuel Wood, Samuel Relton and Edward Vital. The initial concept of the study was set by Drs Arnold, Relton and Vital. Processing of flow cytometry samples was carried out by Zoe Wigston. Drs Dutton and Carter assisted with data collection and recruitment. Further recruitment and all data analysis was carried out by Dr Arnold as was database organisation, processing of demographic data, clustering analysis and figure/abstract curation. Drs Yusof and Wood assisted with the graphical abstract and proof reading. Dr Dass facilitated data collection as head of department.

Chapter 16: is based on work from a jointly authored publication for submission by Jack Arnold, Lucy-Marie Carter, Md Yuzaiful Md Yusof , Zoe Wigston , Daniel Toro Dominguez, Samuel Relton, PRECISESADS Clinical Consortium, Guillermo Barturen, Marta E. Alarcón-Riquelme and Edward Vital. The initial concept of the study was set out by Drs Arnold, Relton, Barturen, Alarcón-Riquelme and Vital. Data collection and data cleaning was carried out by the PRECISESADS Clinical Consortium, Dr Toro Dominguez and Dr Barturen. All analysis was carried out by Dr Arnold under the supervision of Drs Barturen, Relton,

Alarcón-Riquelme and Vital. Manuscript curation and data visualisation was carried out by Dr Arnold.

Chapter 17: is based on work from a jointly authored manuscript being prepared for submission by Jack Arnold, Lucy-Marie Carter, Md Yuzaiful Md Yusof, Katie Dutton, Zoe Wigston, Shouvik Dass, Samuel Relton and Edward Vital. The initial concept of the study was set by Drs Arnold, Relton and Vital. Processing of flow cytometry samples was carried out by Zoe Wigston. Drs Dutton and Carter assisted with data collection and recruitment. Further recruitment and all data analysis was carried out by Dr Arnold as was database organisation, processing of demographic data, clustering analysis and figure/abstract curation. Dr Dass facilitated data collection as head of department. Professor Vital and Dr Relton had general oversight over the analysis.

4 PUBLICATION, POSTERS AND PRESENTATIONS ARISING DIRECTLY FROM THIS THESIS

4.1 PAPERS IN PEER-REVIEWED JOURNALS

- Jack Arnold, Lucy M Carter, Md Yuzaiful Md Yusof, Katherine Dutton, Zoe Wigston, Shouvik Dass, Samuel Wood, Samuel Relton, Edward M Vital, ANA-associated arthritis: clinical and biomarker characterization of a population for basket trials, *Rheumatology*, Volume 63, Issue 11, November 2024, Pages 3135–3145, https://doi.org/10.1093/rheumatology/keae269

4.2 ORAL PRESENTATIONS (FIRST AUTHOR):

- Characterisation of a basket trial population for arthritis associated with ANA-positive RMDs – Podium presentation at British Society of Rheumatology international conference 2023.
- Discovery and validation of a new classification of ANA-RMDs that better predict long term outcomes compared to legacy diagnoses. Oral presentation at EULAR international conference 2024 (EMEUNET top 10 abstract)
- Discovery and validation of a new classification of ANA-RMDs that better predict long term outcomes compared to legacy diagnoses. Oral presentation at Northern Rheumatology Meeting 2024 (Best translational science abstract)
- Discovery and validation of a new classification of ANA-RMDs that better predict long term outcomes compared to legacy diagnoses. Oral presentation at ACR international conference 2024
- Discovery and validation of a new classification of ANA-RMDs that better predict long term outcomes compared to legacy diagnoses. Accepted for presentation at Lupus Toronto international conference 2025
- Discovery and validation of a new classification of ANA-RMDs that better predict long term outcomes compared to legacy diagnoses. Accepted for presentation at Japanese College of Rheumatology international conference 2025 (JCR Travel Award winner)

4.3 POSTER PRESENTATIONS

- Characterisation of an ANA-arthritis basket population for clinical trials – Poster presentation at EULAR international conference 2023

4.4 ABSTRACTS

- Arnold J, Carter LM, MD Yusof MY, et al, AB0510 Immunophenotypic reclassification of ana associated autoimmune disease: preliminary analysis of flow cytometric data in a multi-disease cohort. Annals of the Rheumatic Diseases 2022;81:1382-1383.
- Arnold J, Carter LM, MD Yusof MY*, et al*, POS1516 Characterisation of an ANA-Arthritis basket population for clinical trials *Annals of the Rheumatic Diseases* 2023;**82:**1117.
- Arnold J, Barturen G, Relton S*, et al* OP0145 Discovery and validation of a new classification of ANA RMDs that better predict long term outcomes compared to legacy diagnoses *Annals of the Rheumatic Diseases* 2024;**83:**158.

5 ACKNOWLEDGEMENTS

I would like to acknowledge my supervisors, Professor Edward Vital and Dr Samuel Relton for their support and guidance throughout my doctorate and the composition of this thesis. I would also like to thank my collaborators at the Centro Pfizer-Universidad de Granada-Junta de Andalucía de Genómica e Investigación Oncológica (GENYO) with whom I was fortunate enough to collaborate. My sincere thanks especially go to Professor Alarcon, Dr Toro-Dominguez and Dr Barturen who were supremely welcoming to me and helped supervise my collaborative work with the PRECISESADS cohort.

Within Leeds I am indebted to my colleagues Ms Zoe Wigston for her processing of the gene expression Taqman panels and the flow cytometric data used in the synthesis of this work. I am also grateful to my colleagues Dr Sabih Hassan and Dr Katherine Dutton who aided in the recruitment into this study. Thanks to Dr Lucy Marie Carter and Dr Yuzaiful Yusof who were my guides throughout my doctorate, consistently on hand to offer sensible advice and encouragement, as well as ensuring I stayed robustly on track with my writing.

I am grateful to my funders as well as SLE-Euro for their travelling bursary support which allowed me to collaborate effectively with GENYO. I am also grateful to EULAR and JCR for their travelling bursaries which enabled me to present my collaborative work with GENYO to my colleagues and future collaborators.

In addition to the individuals above I would like to thank all my colleagues and fellow researchers at the Leeds Institute of Rheumatic and Musculoskeletal Medicine and the Leeds Biomedical Research Centre. Special thank you to all of those who helped me refine my work for publication and presentation at national and international conferences.

Finally, I would like to thank my family including my parents Alison and Paul and my wonderful wife Sarah who was a constant rock and source of encouragement. Finally thank you to my son Frederick, who was (comparatively) well behaved during the drafting of this thesis.

6 CONTENTS

7 LIST OF ABBREVIATIONS

Abbreviation	Definition
ACR	American College of Rheumatology
ADCC	Antibody-Dependent Cellular Cytotoxicity
AIC	Akaike Information Criterion
ALT	Alanine Aminotransferase
ANA	Antinuclear Antibody
ANOVA	Analysis Of Variance
APLS	Antiphospholipid Syndrome
APS	Antiphospholipid Syndrome
ARC Classes	ANA-RMD Classification Classes
AST	Aspartate Aminotransferase
AUROC	Area Under the Receiver Operating Characteristic Curve
AZ	AstraZeneca
AZA	Azathioprine
BAFF	B Cell Activating Factor
BIC	Bayesian Information Criterion
BICLA	British Isles Lupus Assessment Group-Based Combined Lupus Assessment
BILAG	British Isles Lupus Assessment Group
BRC	Biomedical Research Centre
cAPS	Catastrophic Antiphospholipid Syndrome
CBOW	Continuous Bag of Words
CCI	Charlson Comorbidity Index
CK	Creatine Kinase

Abbreviation	Definition
CL	Cardiolipin
CLE	Cutaneous Lupus Erythematosus
CNS	Central Nervous System
COVID 19	Coronavirus Disease 19
CPM	Counts Per Million
CRAN.R	Comprehensive R Archive Network
CRF	Clinical Records Form
CRP	C-Reactive Protein
CT	Computed Tomography
CTD	Connective Tissue Disease
DAG	Directed Acyclic Graph
DAS-28	Disease Activity Score 28
DM	Dermatomyositis
DMARD	Disease-Modifying Antirheumatic Drug
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
EMEUNET	Emerging EULAR Network
EMG	Electromyography
EMS	Early Morning Stiffness
ENA	Extractable Nuclear Antigen
ESR	Erythrocyte Sedimentation Rate
ESSDAI	EULAR Sjogren’s Syndrome Disease Activity Index
EU	European Union
EULAR	European League Against Rheumatism

Abbreviation	Definition
EUSTAR	European Scleroderma Trials and Research Group
FACIT-Fatigue	Functional Assessment of Chronic Illness Therapy Fatigue
FDA	Food And Drug Administration
GC	Glucocorticoid
GDPR	General Data Protection Regulation
GMM	Gaussian Mixture Modelling
GWAS	Genome-Wide Association Study
HAQ	Health Assessment Questionnaire
HC	Healthy Controls
HCQ	Hydroxychloroquine
HLA	Human Leukocyte Antigen
ICD	International Classification of Diseases
ICECAP	ICEpop Capability Measure for Adults
IFN	Interferon
IFN-I	Type 1 Interferon
IFNAR	Interferon Alpha Receptor
IIM	Idiopathic Inflammatory Myopathy
IL	Interleukin
IMD	Index Of Multiple Deprivation
ISG	Interferon Stimulated Gene
ITGAM	Integrin Alpha M
JAK	Janus Kinase
JAMA	Journal of the American Medical Association
LAMDA	Lupus Arthritis and Musculoskeletal Disease Activity Index

Abbreviation	Definition
LDH	Lactate Dehydrogenase
LLDAS	Lupus Low Disease Activity State
LMER	Linear Mixed Effects Modelling
MCP Joint	Metacarpophalangeal Joint
MCS	Mental Component Score
MCTD	Mixed Connective Tissue Disease
MFI	Mean Fluorescence Intensity
MH	Mental Health
MHC	Major Histocompatibility Complex
MICE	Multiple Imputation with Chained Equations
MITAX	Myositis Disease Activity Assessment Tool
MMF	Mycophenolate Mofetil
MR	Magnetic Resonance
MRC	Medical Research Council
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
MSK	Musculoskeletal
MTX	Methotrexate
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
NK Cell	Natural Killer Cell
NKT Cell	Natural Killer T Cell
NLP	Natural Language Processing

Abbreviation	Definition
NSAID	Nonsteroidal Anti-Inflammatory Drug
OA	Osteoarthritis
OMERACT	Outcome Measures in Rheumatology
OR	Odds Ratio
PAC	Proportion Of Ambiguously Clustered Pairs
PAH	Pulmonary Artery Hypertension
PAM	Partitioning Around Medoids
PBMC	Peripheral Blood Mononuclear Cells
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
pDC	Plasmacytoid Dendritic Cell
PF	Physical Function
PGA	Patient Global Assessment
PM	Polymyositis
PPI	Proton Pump Inhibitor
PPIA	Peptidylprolyl Isomerase AÂ
PPIE	Patient and Public Involvement and Engagement
PRECISESADS	Precise Systemic Autoimmune Diseases
PRO	Patient Reported Outcome
PSS	Primary Sjogren Syndrome
RA	Rheumatoid Arthritis
RCSI	Relative Cluster Stability Index
RCT	Randomized Controlled Trial
RE	Role Emotional

Abbreviation	Definition
REML	Restricted Maximum Likelihood
RMD	Rheumatic Musculoskeletal Disease
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RP	Role Physical
RTX	Rituximab
SIGLEC	Sialoadhesin Or CD169
SLE	Systemic Lupus Erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index
SLERPI	Systemic Lupus Erythematosus Risk Probability Index
SLICC	Systemic Lupus International Collaborating Clinics
SRI4	Systemic Lupus Erythematosus Responder Index
SS	Systemic Sclerosis
SSA	Sjogren’s Syndrome Related Antigen A antibody
SSC	Systemic Sclerosis
TNF	Tumour Necrosis Factor
TTG	Tissue Transglutaminase
UA	Usual Activities
UCTD	Undifferentiated CTD
UKPSSR	United Kingdom Primary Sjogren’s Syndrome Registry
USS	Ultrasound Scan
VAE	Variational Autoencoder
VAS	Visual Analogue Scale
VP	Variance Inflation Factor

Abbreviation	Definition
VST	Variance Stabilising Transformation
VTE	Venous Thromboembolism
WCC	White Cell Count
WPAI	Work Productivity and Activity Impairment Questionnaire
XGBoost	Extreme Gradient Boosting

8 LIST OF FIGURES

9 LIST OF TABLES

10 INTRODUCTION

10.1 BACKGROUND AND AIMS

The image on cover page of this thesis encapsulates both the primary question and several of the potential solutions. ANA-associated rheumatic musculoskeletal disease (ANA-RMD) are a family of autoimmune disorders united by a common blood test and with broad ranging clinical manifestations. The word cloud image on the cover of this thesis comprises the most common words written in the case report forms of a cohort of ANA-RMD patients within the DEFINITION study. This demonstrates the broad range of manifestations and illustrates through the size of the words their frequency. Problems such as arthritis, fatigue, rashes and ulcers are common and overlap with other features such as osteoarthritis and depression. This creates a complex and overlapping family of diseases which are difficult to subclassify.

Currently this ANA-RMD family is subdivided into historically defined diagnoses based on predominantly clinical features with the addition of several immunological markers. However, there is considerable overlap between these features and many people fulfil multiple diagnostic criteria or none at all, being termed undifferentiated disease.

More precise disease classification within ANA-RMDs could be achieved with the use of modern biomarkers and clinical data using machine learning based data compression and clustering algorithms. This aims to increase the homogeneity of each disease label in terms of disease outcomes, responsiveness to therapies and the patient experience. Classification could be based around shared biomarkers, shared clinical manifestations as explored or a hybrid approach involving both. The overarching theme throughout is to assess whether these approaches outperform legacy diagnoses and, in the future, whether these newly derived classes may be appropriate for evaluation in clinical trials, targeted therapies and risk stratification.

11 LITERATURE REVIEW

11.1 ANA-ASSOCIATED RHEUMATIC MUSCULOSKELETAL DISEASES

ANA-associated rheumatic musculoskeletal disease (ANA-RMD) are a group of diagnoses that are unified by the presence of anti-nuclear antibodies (ANA) and overlapping clinical features. These diagnoses have classification criteria based on symptoms, signs and routine immunology.

11.1.1 Anti-nuclear antibodies

Autoantibodies against nuclear antigens were first described in the context of Systemic Lupus Erythematosus in 1948 in 1957 with the description of anti-double stranded DNA antibodies(1). Since this initial description a plethora of antibodies against cell nucleus and cytoplasmic autoantigens. have been described across a range of autoimmune disease both within and outside the remit of rheumatology. Collectively this family of antibodies are termed Anti-nuclear antibodies (ANA) and they represent a key unifying diagnostic test across a broad family of autoimmune rheumatic diseases with the overarching term ANA- associated Rheumatic Musculoskeletal diseases (ANA-RMDs)(1–3).

11.1.2 Systemic Lupus Erythematosus (SLE)

SLE is a multisystemic autoimmune disorder resulting from a generalized breakdown in immune tolerance and affects approximately 1 in 2000 women and 1 in 20,000 men. Generally, disease onset is in the 1st-4th decade and disproportionately affects those of African and Asian ancestry(4,5). SLE is profoundly heterogeneous in its clinical presentation ranging from mild mucocutaneous features to life threatening neurological, renal or haematological presentations.

Common manifestations of SLE include skin rashes, arthritis, alopecia and photosensitivity. The kidneys are commonly affected ranging from mild proteinuria to end stage renal failure. Serosal involvement is not uncommon with pleural effusions, pericarditis and enteritis all noted features. Autoimmune cytopenias are also common. Multiple classification criteria have been developed, including the ACR 1982, 1997 and 2019 criteria and the Systemic Lupus International Collaborating Clinics (SLICC) 2012 criteria. All these criteria attempt to preserve brevity whilst incorporating the full spectrum of SLE related pathology. The most modern of these criteria, the 2019 EULAR/ACR criteria require a positive ANA (titre ≥1:80) and a combined score of ≥10(6).

Entry criterion

Antinuclear antibodies (ANA) at a titer of ≥1:80 on HEp-2 cells or an equivalent positive test (ever)

↓

If absent, do not classify as SLE
If present, apply additive criteria

↓

Additive criteria

Do not count a criterion if there is a more likely explanation than SLE.
Occurrence of a criterion on at least one occasion is sufficient.
SLE classification requires at least one clinical criterion and ≥10 points.
Criteria need not occur simultaneously.
Within each domain, only the highest weighted criterion is counted toward the total score§.

Clinical domains and criteria	Weight	Immunology domains and criteria	Weight
Constitutional		***Antiphospholipid antibodies***	
Fever	2	Anti-cardiolipin antibodies OR	
Hematologic		Anti-β2GP1 antibodies OR	
Leukopenia	3	Lupus anticoagulant	2
Thrombocytopenia	4	***Complement proteins***	
Autoimmune hemolysis	4	Low C3 OR low C4	3
Neuropsychiatric		Low C3 AND low C4	4
Delirium	2	***SLE-specific antibodies***	
Psychosis	3	Anti-dsDNA antibody* OR	
Seizure	5	Anti-Smith antibody	6
Mucocutaneous			
Non-scarring alopecia	2		
Oral ulcers	2		
Subacute cutaneous OR discoid lupus	4		
Acute cutaneous lupus	6		
Serosal			
Pleural or pericardial effusion	5		
Acute pericarditis	6		
Musculoskeletal			
Joint involvement	6		
Renal			
Proteinuria >0.5g/24h	4		
Renal biopsy Class II or V lupus nephritis	8		
Renal biopsy Class III or IV lupus nephritis	10		

Total score:

↓

Classify as Systemic Lupus Erythematosus with a score of 10 or more if entry criterion fulfilled.

Table 1: EULAR/ACR 2019 SLE classification criteria, reproduced from the original paper by Aringer et al(6).

The key pathophysiological mechanism in SLE is the production of autoantibodies against nuclear antigen, a factor shared across the spectrum of ANA-RMDs, resulting in immune complex deposition, activation of complement mediated pathways and organ damage(7). The other key pathological mechanism is the type I IFN pathway (IFN-I). The association between IFN-I and SLE has been known for some time(8), and research has linked high IFN-I levels to increased rates of neuropsychiatric, renal and dermatological features(8–11). Key genes in this pathway are variants in HLA and Fcγ receptor genes and clinically other type I interferonopathies such as Aicardi–Goutières syndrome present with striking lupus-like presentations(12). Within SLE mRNA sequencing analysis has identified separate signatures associated with disease "susceptibility", "activity" and "severity". Severity signatures were shown to be associated with lupus nephritis and enrichment of plasmablast associated pathways. Interestingly within this work, "susceptibility" signatures were shown not to normalize, even in patients in clinical remission(13).

Morbidity and mortality as mentioned varies greatly between individuals. The general disease course is generally relapsing and remitting with periods of low activity interspersed with acute flare episodes. In the USA mortality in African American patients with SLE is higher than those with European ancestry, irrespective of socioeconomic status(4,5). Across SLE prevalent cases mortality has been estimated at 18.6 per 1000 person/year, but in incident SLE cases the mortality rate is higher at 26.5 per 10,000 person/month peaking in the first year after diagnosis(14). Five- and eight-year survival rates are estimated 91% and 89%, respectively(14). Patient reported outcomes are similarly impaired and have shown to be predictive of mortality outcomes in SLE(15). Anxiety and depression are common, associated with impaired medication adherence and have been shown to persist even after objective disease activity measurements improve(16).

11.1.3 Primary Sjogren Syndrome (pSS)

Primary Sjogren's Syndrome (pSS) is an ANA-RMD characterised by lymphocytic infiltration of the exocrine glands, xerostomia and xerophthalmia and the presence of anti-Ro and anti-La antibodies. The initial clinical triad of xerostomia, keratoconjunctivitis sicca and polyarthritis was first described by Henrik Sjögren in 1925 and the association with Ro and La antibodies was first described in 1961(17,18).

Disease onset is generally later than SLE, with a high incidence in perimenopausal women. Other features and internal organ involvement may also be present with significant clinical overlap with SLE. Approximately 50% of patients with pSS develop extra glandular features. Sjögren syndrome can exist as its own entity (pSS) or in combination with other autoimmune

diseases such as SLE or Systemic Sclerosis (Secondary SS)(19). Complicating this further is the clinical overlap with chronic pain and depressive disorders in pSS(20).

Interferon stimulated gene expression has been noted to be upregulated in pSS patients and murine models of pSS(21). This was first documented in the 1970s when type II IFN was detected in the serum of patients with SLE, pSS, systemic sclerosis and rheumatoid arthritis patients(22). Salivary gland biopsy analysis has shown high levels of IFN pathway activation and plasmacytoid dendritic cell recruitment in tissue samples from pSS patients(23). Systemic IFN-I signatures have been detected in PBMCs and whole blood samples with 3 IFN patterns detected: type 1 dominant, type 1 and 2 mixed pattern and an inactive pattern(24,25). In totality between 53-81% of pSS patients have been shown to express high levels of type 1 IFN(26). In the same manner as SLE apoptotic autoantigens are thought to provide the initial immunogenic stimulus which drives IFN-I production, possibly via plasmacytoid dendritic cell activation in exocrine gland tissue

pSS is currently classified according to the 2016 ACR/EULAR criteria which require a score of ≥4 from the following:

- Labial salivary gland with focal lymphocytic sialadenitis and focus score ≥ 1 (3 points)
- Anti-SSA (Ro) + (3 points)
- Ocular staining score ≥ 5 (or van Bijsterfeld score ≥ 4) on at least one eye (1 point)
- Schirmer ≤ 5 mm/5min on at least one eye (1 point)
- Unstimulated whole saliva flow rate ≤ 0.1 ml/min (1 point)

Anti-Ro antibody positivity is common across multiple ANA-RMDs, particularly SLE, IM and UCTD and makes up 3 points of the criteria, the remaining oral and ocular sicca features are critical(27). However, sicca features and positive Schirmer tests have been documented in SLE, RA and SSc patients within a Thai study(28). Given the porous nature of these criteria other techniques such as salivary gland ultrasound have been employed to try to differentiate glandular pSS from other ANA-RMDs and coexistent pSS with other ANA-RMD diagnoses is commonplace(29).

Therapeutic approaches focus on topical management of mucosal and ocular sicca symptoms with eyedrops, lubricants and occasionally systemic pilocarpine drops. Extra glandular involvement is commonly managed with hydroxychloroquine (HCQ), non-steroidal anti-inflammatory drugs (NSAIDs) and occasionally other disease modifying anti-rheumatic drugs (DMARDs). Refractory disease may require treatment with B-cell depleting therapy, though this is currently not licensed for pSS patients in the UK and must be applied for with

an individual funding request (IFR)(30). Lack of parity in treatment provision between SLE patients with Sicca and Ro antibodies and comparable pSS patients is an exemplar of the need for disease reclassification in ANA-RMDs.

Although for many pSS patients the disease runs an indolent course there is an association between glandular pSS and non-Hodgkin's lymphoma. Over 25 years of follow up non-Hodgkin's lymphoma incidence was 10.5% in pSS patients, a 20-fold increase over the general population. Concurrence of other autoimmune diseases is common with 15.8% of patients developing thyroid disease, 7.2% developing pulmonary fibrosis and 10.5% developing a form of vasculitis in the same study(31). Further registry data has shown that the presence of cryoglobulins and high disease activity scores are associated with increased all-cause mortality risk in pSS.

11.1.4 Inflammatory Myopathy (IIM)

Inflammatory Myopathy (IIM) comprises a group of disorders characterised by autoimmune inflammation of striated muscle. This may be associated with cutaneous features in the case of dermatomyositis (DM), and patients may also be affected by systemic features, notably interstitial lung disease. IIM is rare with an annual incidence of approximately 2-19 cases/million in the US(32). There is a bimodal peak incidence at between 5-15 years if and 45-60 years. Within the USA those of African ancestry are more commonly affected compared to European ancestry patients at a ratio of between 3 to 4:1(32).

Multiple autoantibodies are known to be associated with IIM. Both myositis associated antibodies (MAAs) and myositis specific antibodies (MSAs) have been described. Certain antibodies are associated with specific clinical syndromes. For example, antibodies against histidyl tRNA synthetase (Anti-Jo1) are associated with antisynthetase syndromes whereas anti-PL12, anti-MDA5, anti-OJ and anti-EJ antibodies are associated with higher rates of interstitial lung disease. Others such as anti-TIF1ɣ have been associated with paraneoplastic myositis overlaps(33).

Clinical diagnosis is based around a combination of clinical factors, immunological findings such as MSA and MAAs and histological appearances on muscle biopsy. MRI imaging of affected muscle groups may also show inflammatory changes. Joint classification criteria between EULAR and ACR were developed in 2017 and permit classification of definite, probable and possible IIM with further criteria aiding the diagnosis of IIM subgroups(34). The EULAR/ACR criteria for IIM are detailed in Table 2 . In those with a muscle biopsy available probable IIM requires a score ≥6.7 and <8.7, and definite IIM is defined as an aggregated

score ≥8.7. In those without a biopsy probable IIM is defined as an aggregate score ≥5.5 and <7.5 and definite IIM ≥7.

Variable	Score Points		Definition
	Without muscle biopsy	With muscle biopsy	
Age of onset			
Age of onset of first symptom assumed to be related to the disease ≥ 18 years and < 40 years	1.3	1.5	18 ≤ Age (years) at onset of first symptom assumed to be related to the disease < 40
Age of onset of first symptom assumed to be related to the disease ≥ 40 years	2.1	2.2	Age (years) at onset of first symptom assumed to be related to the disease ≥ 40
Muscle weakness			
Objective symmetric weakness, usually progressive, of the proximal upper extremities	0.7	0.7	Weakness of proximal upper extremities as defined by manual muscle testing or other objective strength testing, which is present on both sides and is usually progressive over time
Objective symmetric weakness, usually progressive, of the proximal lower extremities	0.8	0.5	Weakness of proximal lower extremities as defined by manual muscle testing or other objective strength testing, which is present on both sides and is usually progressive over time
Neck flexors are relatively weaker than neck extensors	1.9	1.6	Muscle grades for neck flexors are relatively lower than neck extensors as defined by manual muscle testing or other objective strength testing

Variable	Score Points		Definition
In the legs proximal muscles are relatively weaker than distal muscles	0.9	1.2	Muscle grades for proximal muscles in the legs are relatively lower than distal muscles in the legs as defined by manual muscle testing or other objective strength testing
Skin manifestations			
Heliotrope rash	3.1	3.2	Purple, lilac-coloured or erythematous patches over the eyelids or in a periorbital distribution, often associated with periorbital oedema
Gottron´s papules	2.1	2.7	Erythematous to violaceous papules over the extensor surfaces of joints, which are sometimes scaly. May occur over the finger joints, elbows, knees, malleoli and toes
Gottron's sign	3.3	3.7	Erythematous to violaceous macules over the extensor surfaces of joints, which are not palpable
Other clinical manifestations			
Dysphagia or oesophageal dysmotility	0.7	0.6	Difficulty in swallowing or objective evidence of abnormal motility of the oesophagus
Laboratory measurements			
Anti-Jo-1 (anti-histidyl-tRNA synthetase) autoantibody present	3.9	3.8	Autoantibody test in serum performed with standardized and validated test, showing positive result
Elevated serum levels of creatine kinase (CK)* *or* lactate dehydrogenase (LDH)* *or* aspartate aminotransferase ASAT/AST/SGOT)* *or* alanine aminotransferase(ALAT/ALT/SGPT)	1.3	1.4	The most abnormal test values during the disease course (highest absolute level of enzyme) above the relevant upper limit of normal

Variable	Score Points		Definition
Muscle biopsy features- presence of:			
Endomysial infiltration of mononuclear cells surrounding, but not invading, myofibres		1.7	Muscle biopsy reveals endomysial mononuclear cells abutting the sarcolemma of otherwise healthy, non-necrotic muscle fibres, but there is no clear invasion of the muscle fibres
Perimysial and/or perivascular infiltration of mononuclear cells		1.2	Mononuclear cells are located in the perimysium and/or located around blood vessels (in either Perimysial or endomysial vessels)
Perifascicular atrophy		1.9	Muscle biopsy reveals several rows of muscle fibres which are smaller in the perifascicular region than fibres more centrally located
Rimmed vacuoles		3.1	Rimmed vacuoles are bluish by Haematoxylin and Eosin staining and reddish by modified Gomori- Trichrome stains

Table 2: 2017 EULAR/SCR classification criteria for adult and juvenile idiopathic inflammatory myopathies

As a ready reckoner the original Peter and Bohan criteria (Used in the inception of this thesis' DEFINTION cohort) remain helpful in summarizing the key clinical features of IIM(35–38). They are as follows:

1 Symmetrical weakness of limb-girdle muscles and anterior neck flexors

2 Muscle biopsy evidence typical of myositis

3 Elevation of serum skeletal muscle enzymes, particularly CK

4 Typical EMG features of myositis

5 Typical DM rash, including heliotrope and Gottron's papules

For the diagnosis of PM		For the diagnosis of DM	
Definite	All of items 1–4	Definite	Item 5 plus 3 of items 1–4
Probable	3 of items 1–4	Probable	Item 5 plus 2 of items 1–4
Possible	2 of items 1–4	Possible	Item 5 plus 1 of items 1–4

Table 3: Bohan and Peter criteria for IIM(36)

Core features of myositis are observed in other ANA-RMDs, continuing the theme of bleed-through between diagnoses. These include elevations in muscle CK levels which occur within disease activity measurements for SLE, interstitial lung disease in IM and Systemic sclerosis and scleromyositis (overlap between Systemic Sclerosis and IIM)(39–41).

The role of type 1 IFN pathways in IIM is well documented, initially described by Isenberg et al in 1986.(42) Plasmacytoid dendritic cell (pDC) infiltration of striated muscle tissue, secretion of IFN-I and upregulation of interferon stimulated genes in muscle and peripheral blood mononuclear cells (PBMCs) has been observed(43). The skin is also affected with 21/25 of the most overexpressed genes in DM patients being ISGs(44). The exact mechanisms underpinning IFN-I upregulation are not entirely clear however from a pathological perspective in vitro studies have suggested that IFN-I is able to induce myotubule atrophy, impair capillary networks and induce mitochondrial dysfunction in dermatomyositis (45–47). Indeed, some of the most active IIM patients, the anti-MDA5

subset have been shown to have among the most pronounced elevations in their IFN-I signatures(43)

Treatment of IIM is with oral immunosuppressants and glucocorticoids in the acute phase with tapering as disease control is obtained. Trials of targeted biologic therapies such as Janus kinase inhibitors are underway(48). Intravenous immunoglobulin and rituximab are used for treatment refractory disease in the UK(49). In the UK rituximab therapy is available for IIM and indeed Scleromyositis, who fulfil IIM criteria, but not for systemic sclerosis patients. This again emphasizes disparity in treatment provision between very phenotypically similar patient groups.

Despite current therapies treatment outcomes remain poor. Estimates of 10-year survival vary greatly ranging from between 20-90%, with death primarily occurring due to malignancy, cardiovascular disease and respiratory disease(50,51).

11.1.5 Antiphospholipid syndrome (APLS)

Antiphospholipid syndrome (APLS) is an autoimmune coagulopathy which can occur as a primary disease (Primary APLS or PAPs) or secondary to coexisting autoimmunity (Secondary APLS). The most common secondary cause of APLS is concurrent SLE. It is most characterised by the presence of autoantibodies to cardiolipin, β2 glycoprotein 1 or lupus anticoagulant. RNA sequencing analysis has however identified differentially expressed genes which separate SLE with secondary APLS from non-APLS SLE. SLE with secondary APLS was characterised by downregulation of IFNα and IFNɣ expression, and venous thromboembolic events associated with upregulation of neutrophil/myeloid response modules when compared to pure SLE samples(52).

The disease can be classified according to the International Consensus Statement criteria or the 2023 EULAR/ACR classification criteria(53). For brevity the International Consensus Statement criteria require the presence of 1 clinical criterion (Vascular thrombosis or pregnancy morbidity) and positivity on 2 separate occasions ≥12 weeks apart to one of the associated antiphospholipid antibodies(54).

Patients may be diagnosed following presentation with an acute thrombus or pregnancy loss or may present with multiorgan failure secondary to Catastrophic APLS (cAPS). Skin rashes such as livedo reticularis are common and MSK features may be present.

Treatment is primarily with anticoagulation to reduce the risk of further thrombus formation, more severe manifestations such as cAPS can require treatment with oral glucocorticoids

and other immunosuppressants such as Rituximab. Outcomes vary widely, with cAPS carrying a poor prognosis, in comparison to single VTE events in susceptible individuals.

The type I IFN pathway also appears to be upregulated in PAPs even in the absence of concurrent SLE or other ANA-RMDs, more notably in those with anti-. β2 glycoprotein 1 antibodies(55). Therapeutic intervention with antimalarial therapy appears to reduce this. However other studies have shown this IFN-I signature not to be as durable as those noted in SLE(56). However elevated IFN-I expression has been linked to endothelial damage and atherosclerosis formation in APLS, driven by defective endothelial progenitor cell function(57). More work is needed in this area to clarify these associations.

11.1.6 Systemic Sclerosis (SSc)

Systemic sclerosis (SSc), or scleroderma, is an ANA-RMD characterised by microvascular disease and excessive fibrotic collagen deposition in the skin and internal organs. Raynaud's and skin thickening are the most common presenting symptoms. 93.4% of patients are ANA positive with the most common subtypes being anti-centromere antibody (32.3%) and anti-Scl70 (36.8%)(58).

The cause of SSc has not been fully elucidated, but the presence of anti-nuclear antibodies and activation of innate immunity suggests an autoimmune cause(59). Type 1 IFN signatures similar to other ANA-RMDs have been identified in SSc and appears to correlate with increased vascular and lung involvement, chief drivers of mortality in the disease(60–64). Moreover, skin and lung tissue in SSc patients has been shown to overexpress IFN stimulated genes(64–66).

The most common classification criteria in place for SSc are the 2013 ACR-EULAR criteria(67). Patients with a score of ≥9 from Table 4 can be classified as systemic sclerosis:

Disease Item	Sub Item	Weight
Skin thickening of the fingers of both hands extending proximal to the metacarpophalangeal joints (sufficient criterion)		9
Skin thickening of the fingers (only count the higher score)	Puffy fingers	2
	Sclerodactyly of the fingers (distal to the metacarpophalangeal joints	4

Disease Item	Sub Item	Weight
	but proximal to the proximal interphalangeal joints)	
Fingertip lesions (Only count the higher score	Digital tip ulcers	2
	Fingertip pitting scars	3
Telangiectasia		2
Abnormal nailfold capillaries		2
Pulmonary arterial hypertension and/or interstitial lung disease (maximum score is 2)	Pulmonary artery hypertension	2
	Interstitial lung disease	2
Raynaud's phenomenon		3
SSc related autoantibodies (maximum score is 3)	Anticentromere	3
	Anti-topoisomerase I	3
	Anti-RNA polymerase III	3

Table 4: 2013 ACR-EULAR classification criteria for systemic sclerosis

Approximately 39% of patients within the EULAR Scleroderma Trials and Research Group (EUSTAR) database were on treatment with glucocorticoids, 8.3% required treatment with the broadly immunosuppressive drug cyclophosphamide, a drug generally reserved for severe or refractory ANA-RMDs(58).Mortality is most commonly due to interstitial lung disease and renal disease, although angiotensin II inhibition has substantially improved outcomes in scleroderma renal crises(58,68). Despite advanced in therapy outcomes remain poor. Cumulative survival from diagnosis at 5 years is 74.9% dropping to 62.5% at 10 years with lung involvement being the main cause of mortality(69).

11.1.7 Mixed Connective Tissue Disease (MCTD)

Mixed connective tissue disease was originally described in 1972 by Sharp et al as an association of Raynaud's phenomenon, puffy fingers, oesophageal dysmotility and arthralgia/arthritis(70,71). Generally, these have been associated with the presence of anti-ribonuclear protein (anti-RNP) antibodies(71–73). However, it should be noted that review of the original Sharp cohort did not show evidence of this as a hallmark feature and moreover that anti-RNP antibodies are known to be expressed in other ANA-RMDs(74).

The aetiology of MCTD remains unclear, several HLA subtypes have been associated with an increased risk of disease onset with DRB1*04:01 noted as a major risk factor(75).

A review of the MCTD literature by Isenberg et al analysed 4 studies with >100 MCTD patients and >5 years of clinical follow up to try to address the issue of heterogeneity within MCTD and the variable expression of “cardinal” features(71). In this sense the standardisation of classification and diagnostic criteria for MCTD has been an ongoing problem with 4 different criteria proposed(76). Subsequent work has compared the specificity/sensitivity of these criteria in anti-RNP positive patients and demonstrated the Khan and Alarcón-Segovia criteria to have the maximum specificity when diagnosing MCTD(77). The Alarcón-Segovia criteria are utilised within the PRECISESADS cohort within this thesis and comprise the following:

- Positivity to anti-RNP (anti-U1RNP) antibody and at least 3/5 of the following clinical features.
    - Hand oedema
    - Synovitis
    - Histologically or biologically proven myositis
    - Raynaud phenomenon
    - Acrosclerosis with or without proximal systemic sclerosis

In terms of therapeutic approaches to MCTD these have been similarly mixed, in a 280-patient study by Hajer et al 78.2% of MCTD patients received glucocorticoids, 74.6% patients received a cytotoxic agent (Defined as methotrexate, cyclophosphamide or azathioprine) and interestingly 15% were treated with tumour necrosis factor inhibition (Ant-TNF). Anti-TNF treatment is particularly curious given the association between these agents and drug induced lupus syndromes which overlap clinically with MCTD(78). Another cohort collated by Capelli et al showed similar levels of glucocorticoid use in MCTD (82%), 58% treated with oral immunosuppressants and 45% in receipt of an oral antimalarial(79). To date no targeted biologic therapies are in use for MCTD in the UK, again this is suboptimal given that many SLE patients (Who would fulfil MCTD criteria) are eligible for such therapies.

A somewhat dated 1999 study demonstrated a favourable outcome in 62% of MCTD patients over 3-29 years of follow up with 38% having on-going disease/death(80). pulmonary artery hypertension (PAH) was a notable cause of mortality in these patients and a subsequent meta-analysis in 2023 showed a PAH prevalence of 12.53% in MCTD patients(81). A Hungarian retrospective study in 2013 of 280 patients identified 5, 10, and 15-year survival rates of 98%, 96%, and 88% respectively. Negative prognostic factors included antiphospholipid antibody positivity and malignancy rates were also noted to be higher in MCTD patients(82).

In a sense the difficulties in classification and characterisation of MCTD are a microcosm of the wider difficulties across the family of ANA-RMDs. An initial postulated phenotype in MCTD has grown to demonstrate widening heterogeneity in its clinical features and associated biomarkers as our understanding of it has increased. A factor that traditional diagnostic approaches have failed to keep pace with.

11.1.8 Undifferentiated CTD (UCTD)

Undifferentiated CTD (UCTD) comprises patients with a positive ANA and some clinical features of ANA-RMDs but not satisfying any of the existing diagnostic criteria for differentiated disease(83). The most commonly accepted criteria for UCTD were developed my Mosca et al and comprises the following(84):

- ≥1 ANA-RMD clinical disease manifestation
- A positive Antinuclear antibody
- Does not fulfil classification criteria for another ANA-RMD diagnosis.

The existence of UCTD as a distinct disease entity, or a as a progression stage to criteria ANA-RMD is debated and changes to diagnostic criteria unsurprisingly impact UCTD prevalence. A 2021 study applied the ACR/EULAR SLE criteria to an existing UCTD cohort and were able to reclassify 20.2% of patients as criteria SLE(85).

UCTD itself can retrospectively be subdivided into Evolving UCTD (eUCTD) defined as UCTD where there is development of criteria disease within 3 years and Stable UCTD (sUCTD) where there is not(83). A systematic literature review carried out in 2023 reviewed 1118 UCTD patients. Anti-Ro (35%) and anti-RNP (29%) were the most prevalent ANAs within the cohort(83). Follow-up ranged from 1-15 years and noted that 18% of UCTD patients achieved remission with others progressing to RA (29%), SLE (25%), pSS (18%) MCTD/SSc (9%) and IIM (1%)(83). Other work has suggested that 50% of systemic autoimmune diseases have an undifferentiated profile at onset with up to 40% of UCTD patients evolving into criteria ANA-RMD(84). Work in SLE applying the SLE Risk Probability Index (SLERPI) to a cohort of 422 UCTD patients identified 39 patients as non-criteria SLE(86) This algorithm has previously been developed in SLE patients where it demonstrated a 94.2% accuracy for detecting SLE within a cohort of 802 adults with SLE or control rheumatological conditions(87). These finding reinforce the idea that a large proportion of UCTD exists on a continuum between asymptomatic autoimmunity and criteria ANA-RMDs, isolated more by historical convention than pathophysiological differences.

The most common therapeutic interventions in UCTD are oral antimalarials and intermittent low dose glucocorticoids, with oral DMARDs being less common(88,89). Mortality was estimated by Mosca et al as 2% within a cohort of 91 patients over 15 years of follow up(90). In terms of long-term damage accrual and therapeutic intervention in stable UCTD a monocentric inception cohort study of 141 patients has been conducted(91). At final follow up 10% of patients were on glucocorticoids and 6% were on immunosuppressive therapy, 27.3% had ≥1 organ system damaged(91). Younger age at diagnosis, not taking antimalarial or oral immunosuppressive therapy and young age at first symptom onset were identified as markers of increased organ damage at 5-year follow up(91).

11.1.9 Overlap in disease activity scoring measures

As discussed, there is substantial overlap between the clinical phenotypes of different ANA-RMD diagnoses. Features such as arthritis and skin rashes are represented across multiple diagnoses. This is summarised in Figure 1 which shows shared clinical features from validated disease activity measures in ANA-RMDs. BILAG-2004, SLEDAI-2K, ESSDAI and MITAX overlap considerably, but heterogeneity remains with the EUSTAR activity index minimally overlapping with the other diseases.

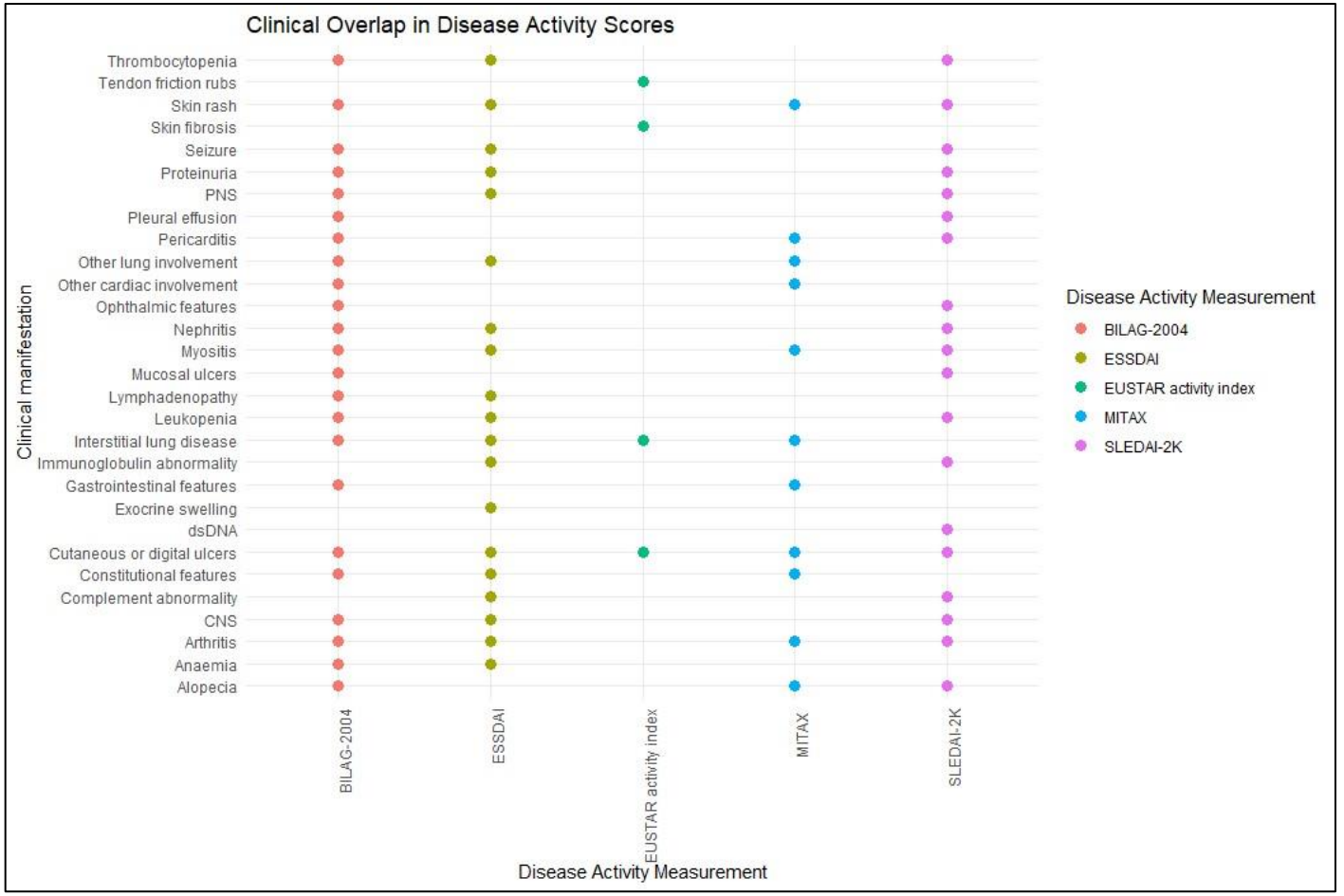


Figure 1: Overlap in ANA-RMD disease activity measurements

11.2 "AT RISK" ANA POSITIVITY

ANA associated rheumatic musculoskeletal disease comprises a broad spectrum of clinical disease as discussed. Patients with symptoms suggestive of an underlying connective tissue disorder are generally screened with an ANA test in primary care. 25% of the population can have a low titre positive ANA, but of those only a small percentage progress to clinically detectable autoimmune disease (92,93). ANA positivity is more common in women and in certain ethnic groups such as those of African and South Asian heritage suggesting a genetic component. Indeed, within established disease more aggressive SLE is noted within certain ethic subgroups, such as the African diaspora (94). However, not all those with a positive ANA go on to develop frank ANA-RMDs. In those that progress, a positive ANA test may predate the onset of clinically detectable symptoms by many years. Consequently, there is significant interest in detection and risk stratification for progression within this "At Risk" cohort.

Within the NHS the natural trajectory for these patients is initial evaluation within the community for symptoms compatible with an underlying ANA-RMD. This is followed by screening with an ANA blood test and other baseline laboratory markers before onward referral to tertiary care rheumatologists if these blood tests are suggestive of an underlying connective tissue disorder. Expedient onward referral here is critical to prevent poor outcomes associated with delays in diagnosis and immunosuppressive management.

Longitudinal work by Yusof et al identified a positive family history of ANA-RMDs and a two factor interferon gene expression score as predictor of progression to frank SLE in at risk patients(95). Other work has demonstrated IFN-ɣ associated protein and Galectin-9 as predictors of disease progression in ANA positive individuals(96). These are not central to this work but help emphasise the ANA-RMD disease continuum from At Risk through to UCTD and then criteria ANA-RMD, with somewhat arbitrary cutoffs between undifferentiated and differentiated CTD.

11.3 SHARED CLINICAL AND BIOMARKER VARIABLES

As noted previously ANA-RMDs have common features including some organ features and symptoms, onset in adolescence to middle age, a lifelong disease course and fluctuating disease activity with flares and remissions. They share common immunopathogenic features and all are treated with a similar range of mostly unlicenced immunosuppressive therapies

that range in intensity and effectiveness. Since we aim to replace them, in this thesis we term the above diagnoses as “legacy diagnoses”.

11.4 Unmet need for better therapies for ANA-RMDs

Although ANA-RMDs are uncommon (prevalence approximately 1 in 1000), they have substantial unmet needs since (i) mortality in people with SLE, SSc and IM is increased more than 3-fold compared to general population despite current best therapy(97–102); (ii) long-term complications such as cardiovascular disease are increased up to 50 times, particularly since therapy with glucocorticoids worsens long-term outcomes(103,104); (iii) SLE costs the UK economy around £8,000 per patient annually in direct and indirect healthcare costs(105,106); (iv) SLE and pSS impair quality of life significantly, comparable to other chronic diseases such as diabetes, hypertension and heart failure(107–109). Despite this widely known unmet need, development of new therapies has been slow. Clinical trials have mostly been performed in SLE. Many of these trials have been negative despite promising phase II data or have resulted in small effect sizes deemed not cost-effective by NICE(110). One reason for these poor trial outcomes may be the use of composite outcome measures that try to assess a wide range of different symptoms and signs. A recent study identified the symptoms prioritised most highly by patients with SLE and these are not always addressed in trial eligibility and endpoints. That may explain why quality of life endpoints are often not met(111). In most trials, subpopulations of higher responding patients have been identified. We believe that these problems would be better addressed by replacing these legacy diagnoses with new classes of ANA-RMD that are aligned to patients needs and therapeutic options available.

11.5 A shared feature across ANA-RMDs – Arthritis

11.5.1 Arthritis within the single disease paradigm – SLE arthritis

As noted previously, multiple clinical features are represented across the spectrum of ANA-RMDs. One of the most prominent of these is the presence of inflammatory arthritis which commonly affects the full spectrum of ANA-RMD disease. To date limited work has been done contrasting the impact of ANA-RMD associated arthritis (ANA-arthritis) across diagnoses. Formalised assessment of this area would be very helpful in guiding prospective work. Currently there are no published works assessing the utility of clinical baskets within ANA-RMDs. Therefore, I plan to assess the utility of the high prevalence feature of inflammatory joint/tendon disease as a potential basket indication.

Within the current single disease paradigm our group has published 4 key works on improving the care of arthritis within SLE. Prevalence was assessed with ultrasonography in a cohort of 88 SLE patients with inflammatory MSK symptoms and compared to 24 SLE controls. 68% of those with MSK symptoms had active inflammation on ultrasound scanning (USS) compared to 17% of asymptomatic patients. 27% of those with inflammatory symptoms had subclinical inflammation only detected on ultrasound, suggesting the utility of ultrasound assessment in SLE related arthritis(112). Building on this further ultrasound-based assessment was shown to be highly sensitive in detecting improvements in MSK SLE. In contrast SLEDAI and joint counts were shown to be less effective and may underestimate the efficacy of MSK targeted therapies in SLE. The USEFUL study by the BILAG group further demonstrated the value of ultrasound in the assessment of SLE arthritis demonstrating that only 68% of lupus arthritis patients with a positive ultrasound had clinically detectable synovitis. Those with a positive ultrasound demonstrated a superior response to glucocorticoid therapy suggesting a role for MSK USS in the enrichment of trial cohorts for future therapeutic trials.(113). Finally, data from the USEFUL study was repurposed to develop the Lupus Arthritis and Musculoskeletal Disease Activity Score (LAMDA) which was subsequently validated in lupus patients. The LAMDA composite score comprises a swollen joint count, physician and patient scored visual analogue scores and an erythrocyte sedimentation rate (114). As LAMDA was shown to outperform traditional lupus joint MSK assessments (SLEDAI-2K and BILAG-MSK domains) in terms of effect sizes it is currently under evaluation for its applicability across ANA-RMDs.

11.5.2 Inequitable access to targeted treatments across ANA-RMDs – ANA+ arthritis

Despite the clinical homogeneity within ANA+ arthritis and shared pathological mechanisms such as IFN-I pathway activation there is significant inequity in access to targeted therapies. Inclusion criteria for trials have concentrated on legacy diagnosis and therefore licencing of therapeutics has followed suit. For example, SLE patients with anti-Ro+, low complement and arthritis would have access to Belimumab, Rituximab and Anifrolumab while no licensed therapy is available for Sjogren's patients with similar antibodies and arthritis.

Table 5 is a brief summation of the prevalence of articular involvement and arthritis across the spectrum of ANA-RMD, illustrating the scale of the problem in question.

Condition	Study type	% Articular involvement	Key findings	Location	Western Prevalence
SLE	Systemic literature review(115)	69-95%	Heterogeneous presentation. Arthralgia, tenosynovitis and erosive disease(116). Poorly treated relative to other organ manifestations. Poor body of evidence and classification system for lupus arthritis in general(116).	NA	0.1% UK adults
UCTD	Retrospective cohort study(117)	30-49%	5 year follow up, n = 716(117). 49% presented with arthralgia(117). 30% presented with arthritis(117).	Hungary	Unclear
Sjogren's Syndrome	Retrospective cohort study(118)	30-60%	Bicentric French cohort study, n = 419(118). 30-60% demonstrate articular involvement. 35% had frank synovitis and 70% were treated with immunosuppression(118).	France	0.25% of Spanish adults(119)
Sjogren's Syndrome	Retrospective cohort study(120)	NA	Retrospective French cohort study, n = 57(120). 57 pSS patients with synovitis(120). Previous treatment: Steroid monotherapy (3.5%); Steroid + HCQ (79%); HCQ monotherapy (49%); MTX (35%); RTX (5.3%)	France	Unavailable
Systemic Sclerosis	Retrospective cohort study	31%	German cohort, n=57(101). 31% showed evidence of arthritis (19% clinically & 26% radiographic)(101).	Germany	307 cases per million (Data from 2013)
Systemic Sclerosis	Systemic literature review(101)	23%	Meta analysis of 7 studies, n = 547(101). Showed 23% prevalence of clinical arthritis(101).	NA	Unavailable

Condition	Study type	% Articular involvement	Key findings	Location	Western Prevalence
Inflammatory Myositis	Literature review(121)	18-53%	Literature review of 17 studies containing at least 20 pts each(121). Arthritis occurs in 18-53% of patients according to previous studies(121,122). Most commonly non erosive polyarthritis, may precede muscle involvement	NA	17.6 per 1000,000 patient years in UK
Inflammatory Myositis	10 year observational study(122)	19%	10-year observational study of Salford cohort(122). 19% had visible arthritis at presentation (122)	UK	Unavailable
MCTD	Literature review(123)	50-90%	90% describe arthralgia(123). 50-60% develop arthritis(123).	NA	3.8 per 100,000 in Norway in 2008(124)

Table 5: Scoping literature review of ANA arthritis prevalence across ANA-RMDs

11.6 THE CHALLENGE OF CLINICAL TRIAL DESIGN IN ANA-RMDS

11.6.1 Trial recruitment

Recruitment to ANA-RMD trials is a significant problem due to their comparatively low prevalence. Recruitment to these trials is often slow and total numbers of patients are often insufficient. This was highlighted in SLE during a recent Open Space meeting of experts(125). Consensus opinion was that multi-site trials involving multiple specialist centres are necessary to generate the power needed for the small, expected effect sizes. These effect sizes may be small due to the eligibility criteria and outcome measure problems in these heterogeneous populations as discussed below.

11.6.2 Subgroup analysis

Work within pSS has shown that subgroups defined by hierarchical clustering of data from the UK Primary Sjogren's Syndrome Registry (UKPSSR) within the failed TEARS and TRACTISS trials would have demonstrated a positive result(126–128). This suggests that current disease label-based stratification fails to identify those most likely to benefit from specific targeted interventions.

11.6.3 Outcome measures are problematic

The use of composite endpoints to represent the wide clinical phenotypes of diseases such as SLE, IM and pSS in clinical trials is a key feature underpinning disappointing results in promising therapeutics. In SLE Rituximab has had 2 large phase 3 studies which failed to meet their primary endpoints in, EXPLORER and LUNAR.

EXPLORER was a phase III randomised controlled trial assessing non-renal and non-CNS lupus in 257 North American patients. Patients were randomised to receive either Rituximab (2 x 1000mg doses 2 weeks apart) or placebo alongside 0.5-1mg/kg prednisolone tapering to 10mg daily by week 10 of the study(129). The primary endpoint was remission and maintenance of remission by week 52 as defined by the BILAG criteria (BILAG C or better in all domains). Unfortunately, this study failed to meet its primary endpoint as no significant difference was shown in major or partial clinical responses at 6 months. Subgroup analysis did show some RTX superiority for post hoc endpoints, RTX was shown to be superior for African American and Hispanic groups and there were reduced incidences of BILAG A category disease flares.

LUNAR was another phase 3 randomised controlled trial designed to investigate the use of RTX in lupus nephritis(130,131). 144 patients were randomised 1:1 to receive either RTX or

placebo in combination with mycophenolate mofetil (MMF) and prednisolone. This too failed to meet its primary endpoint of renal response at week 52. The failure of this trial has been discussed extensively. The relatively short follow up time of 52 weeks has been criticised as RTX has been shown to improve proteinuria at 72 weeks in other studies(130,132). Additionally, the relatively small sample size in LUNAR prevented statistically significant analysis of differences in partial treatment response between the treatment arms.

There were issues common to both studies which may also explain their failure to meet their primary endpoints. Both studies allowed high prednisolone doses and LUNAR additionally added in 3g MMF alongside RTX. This is understandable in the face of organ threatening disease but may have masked any RTX related treatment effect(129–132).

Disease activity scores have been an issue in lupus trials for some time, involving anti-CD20 monoclonals and other biologic therapies. The British Isles Lupus Assessment Group (BILAG) scoring system was utilised in both trials. In EXPLORER many patients had far greater disease activity well above the BILAG 1A and BILAG 2B scores required as inclusion criteria(129,131,133). BILAG B scores were shown to be poor at differentiating partial RTX responders. A good example of this was a mild malar rash scoring the same as a scarring discoid lupus rash. Likewise, the SLEDAI scoring system can also be poor at picking up partial drug responders. Consequentially composite scores such as the BILAG-based combined lupus assessment (BICLA) and systemic lupus erythematosus responder index (SRI) have been developed to address this issue. The BICLA was used successfully in the TULIP2 trial of anifrolumab in SLE when the parent study TULIP1 failed to meet an SRI response as its primary endpoint. The BICLA is based upon improvement of baseline BILAG disease activity without any worsening of individual BILAG domains. In comparison the SRI requires a ≥4 point reduction in the SLEDAI score with no new BILAG- A features and no more than 1 new BILAG-B score. In addition, an SRI response precludes any significant deterioration in the physician global assessment. Practically these differences mean that the SRI demonstrates complete response in at least 1 organ system more effectively than the BICLA. The BICLA more effectively demonstrates composites of partial response in several organ systems better than the SRI-4(134)

As major organ involvement such as active lupus nephritis and neuropsychiatric disease are excluded from trials in SLE, drug efficacy tends to be determined by the response in musculoskeletal and mucocutaneous organ systems within clinical trials(135). These patients represent the cohort of stably unstable disease amenable to clinical trial inclusion but do not represent the real-world spectrum of SLE one might encounter in clinic. Moreover,

certain features such as immune mediated fatigue which predominate in pSS patients are poorly characterised by existing disease activity scores or absent altogether as in the case of the EULAR Sjögren's syndrome disease activity index (ESSDAI), a common composite endpoint in pSS trials(136).

11.6.4 Classification and labelling bias

Classification of disease is a common problem in pSS trials wherein patients with extra glandular features are commonly mislabelled as SLE due to significant clinical overlap and therefore excluded from trials(137). This was noted in a 2015 study which determined that more than half of the most unwell pSS patients were incorrectly labelled as RA or SLE(138).

Additionally, patients with UCTD who make up a substantial proportion of the ANA-RMD population do not currently qualify for any clinical trials and are instead treated in an off-label manner as they fail to fulfil any diagnostic criteria. Trials in rarer diseases within the ANA-RMD spectrum such as IM have previously had difficulty recruiting requisite numbers and therefore may benefit from broader basket-based approaches.

11.7 THE CASE TO RE-CLASSIFY THE SPECTRUM OF ANA-RMDS

11.7.1 Clinical and demographic variables stratify ANA-RMDs

Previous studies evaluated predictors of clinical outcomes within individual ANA-RMD diagnoses. Demographic, clinical, immunological characteristics and co-morbid conditions predicted many of these outcomes, such as mortality in myositis, systemic features in pSS and fatigue across multiple diseases. Many of these predictors were similar in each diagnosis (139–142). The effect of ancestry is particularly important with African ancestry patients responding worse to licensed therapy despite experiencing more severe disease(143). These results suggest that a common set of clinical and demographic features may have value in defining clinically relevant subgroups across the disease spectrum. This is important to address healthcare inequalities.

11.7.2 Subdivision of existing diseases linked to increased efficacy of therapy

Specific biologic therapies are more efficacious within existing subgroups of ANA-RMD diagnoses

Belimumab and anifrolumab are the only two drugs licensed for ANA-RMD, both for SLE. Belimumab targets BAFF and is more effective in patients with high disease activity (measured by SLEDAI), positive DNA-antibodies and low complement(144). Anifrolumab

targets Type I interferon and is more effective in patients with a positive interferon gene signature (IFNGS)(145–147). In belimumab-treated patients IFNGS had only a trend to association with response(148). Rituximab is not licensed but is commissioned by NHSE for SLE, myositis, and sometimes used in individual cases of other ANA-RMDs. Predictors of better clinical response to rituximab in SLE were investigated in an MRC-funded consortium. In this, our group showed that number of ANA subtypes and an optimised interferon score predict better response to rituximab(149). Older patients, patients with non-European ancestry, higher disease activity and cutaneous manifestations were associated with inferior responses(150). In pSS, although the primary endpoint of a trial of rituximab was not met, re-analysis of trial data into four subgroups by Prof Ng suggested a treatment effect with hydroxychloroquine in the High Symptom Burden subgroup and with rituximab in the Dryness-Dominant with Fatigue subgroup compared with placebo(151).

11.7.3 Shared genetic signatures across ANA-RMDs

There is considerable evidence for shared genetic signatures across ANA-RMDs. Previous genome wide association studies (GWAS) have identified shared risk loci spanning SLE, pSS and SSc(152–154) . Moreover, there are multiple HLA associations with ANA-RMD, and specific polymorphisms have been associated with SLE, pSS and SSc in the same manner(155–163). Therefore, given this significant homogeneity within 3 disease groups it is logical to assume some further overlap with UCTD, IM and MCTD.

Interferon (IFN) signatures have shown to be elevated in peripheral blood samples of patients with multiple different ANA-RMDs. Multiple studies have demonstrated an elevated IFN signature in SLE, pSS and SSc(164,165). A more recent epigenome wide comparative study has demonstrated a similar signature in MCTD(166). Additional markers of IFN activity such as cell surface SIGLEC-1 and Tetherin have been shown to be elevated in ANA-RMDs and, in the case of pSS, SIGLEC-1 expression may help discriminate between glandular and extra glandular involvement(167,168). Indeed, high interferon expression has been linked to increase rates of progression to active ANA-RMDs in asymptomatic ANA+ patients(169).

11.7.4 The same clinical subpopulations may exist within multiple diagnoses

Subpopulations who may be most clinically relevant or respond best to therapy in SLE trials may be represented in each legacy diagnosis. For example, belimumab and anifrolumab are known to be effective for synovitis in SLE patients. There are patients with synovitis within each of the other legacy diagnoses who are not currently eligible for these treatments despite the same clinical problem. Grouping these together may result in a trial population that is more homogenous than the legacy diagnosis approach. This facilitates more effective

trial designs. For example, in a basket trial for ANA-RMD patients with arthritis, appropriate eligibility criteria, standard of care, and a more responsive organ-specific outcome measure could be used, as known to be successful in rheumatoid arthritis trials.

11.7.5 There are common pathogenic processes in each disease with corresponding therapies and biomarkers

There are many common pathogenic mechanisms in ANA-RMDs, such as B cells that lead to production of autoantibodies, or type I interferon pathway activation(170). Analysis of the B cell-targeted therapies belimumab and rituximab both show that markers of B cell activation, such as autoantibody profiles, complement levels, or plasmablast levels, can predict clinical response. Such profiles can be identified within many ANA-RMD legacy diagnoses. IFNGS is a biomarker for better response to anifrolumab. Each legacy diagnosis also contains an IFNGS-positive and -negative subpopulation so grouping these together may result in higher effect sizes since all patients included are the “biomarker-positive” subset of their legacy diagnosis(171). This was the original objective of the DEFINITION study used in this application.

11.7.6 Machine learning has been used to reclassify patients within existing diagnoses

Both supervised and unsupervised machine learning approaches have been applied within individual ANA-RMD disease cohorts. For example: in rheumatoid arthritis machine learning approaches have been applied to predict disease flares on tapering biologic drugs(172). Professor Morand’s group applied machine learning methods to subtyping within SLE; analysis of whole blood RNAseq data revealed four overlapping clusters that differed in terms of SLE classification criteria and flares, with some suggestion of appropriate therapeutic targets(173). In a second study, analysis of routinely available clinical data was used to identify high-disease activity SLE patients in need to escalation of immunosuppression(174). Professor Bertsias’ lab have applied similar machine learning approaches including K-means and extreme gradient boosting to identify subsets of SLE patients at greater risk of early damage accrual(175). Whole blood RNA sequencing and hierarchical clustering techniques have also been used to identify 5 distinct SLE endotypes (Neutrophilic, B-cell, metabolic and 2 mild disease subtypes). Interestingly targeted intervention with Bortezomib in the neutrophilic cluster, azathioprine/ixazomib in the B-cell cluster and fostamatinib in the metabolic cluster was predicted to reverse these derangements, suggesting a potential avenue for targeted intervention(176). Professor Jury’s team applied a k-means clustering approach in the immune cell profile of juvenile SLE

patients and stratified the disease into 4 distinct subgroups(177). Logistic Regression and Decision Trees methods with the forward wrapper selection technique were used to identify factors associated with ultrasound detected damage in 120 adult SLE patients(178). As well as machine learning, within SLE there have been several studies describing the use of gene expression has to define subsets of SLE patients that differ in disease activity progression(179–182).

11.7.7 Undifferentiated connective tissue disease not addressed by research or trials in any legacy diagnosis

Many patients with ANA-RMDs do not meet the classification criteria for any one legacy diagnosis. These patients are termed “undifferentiated connective tissue disease”. This population has been studied in detail by Professor Marta Mosca(183–185). These patients have a diverse range of clinical features, poor quality of life, and are not eligible for any trial or licensed therapy. Reclassification would allow the option for these patients to be included in new patient groupings and a large number were included in DEFINITION and PRECISESADS.

11.7.8 Gene expression-based ANA-RMD reclassification

The need for stratification within and across ANA-RMDs, and for basket trial approaches was emphasised in a recent Open Space meeting of international lupus experts(23). Given this novel approach there has been considerable interest in further application of these methods to the field with several review articles summarising the strengths of such an approach(186–188).

The PRECISESADS EU-funded project (2014-19) led by Prof Alarcon-Riquelme, aimed to reclassify autoimmune rheumatic disease according to a molecular taxonomy(154). 2500 autoimmune rheumatic disease patients were compared to 600 healthy controls using genotype, transcriptomic and metabolic analysis followed by unsupervised clustering methods. This defined four patient clusters: “inflammatory”, “lymphoid”, “interferon” and “low-disease activity”. However, this study did not analyse patient types in terms of clinical significance such as quality of life or requirement for glucocorticoids or other therapies, which is the key focus of this work. In another study, analysis of flow cytometric data from pSS and SLE using unsupervised machine learning revealed subtypes that spanned diagnoses, although again detailed clinical phenotyping was not included in the analysis(189,190). Machine learning methods have been used to identify a specific inflammatory signature of interferon stimulated genes in patients with several ANA-RMDs (SLE, pSS and UCTD) with a common positive blood test (Ro60)(191). This demonstrates

the application of these techniques in smaller sample sizes and the utility of a combined clinical, imaging and laboratory-based ML model.

Certain gene expression signatures have attracted significant interest due to shared pathological mechanisms spanning multiple ANA-RMDs. Dysregulation of type 1 interferons is a cardinal feature of multiple ANA-RMDs. Interferons are subdivided into 3 families (Type I, type II and type III). Type I IFNs have been implicated in the body's response to viral pathogens and tumour surveillance. In the context of ANA-RMDs type I IFNs have been shown to be potent immunostimulators, driving B-cell activation, antibody production and immunoglobulin class switching(192–195). Anifrolumab is a type I IFN blocking medication which works by blocking the IFNAR receptor. Anifrolumab has shown positive trial and real-world evidence in SLE and is currently undergoing clinical trials in SSc and pSS(196–199).

Within the single disease paradigm of SLE microarray analysis of the transcriptome has allowed for the assembly of functionally annotated modules that distinguish by disease activity, autoantibody status, renal involvement and cutaneous disease(200–203). Within my lab two IFN scores (IFN Score A and IFN Score B) have been derived using factor analysis from these interferon annotated modules, with IFN Score B predicting progression from the At-Risk state to SLE(95,204). Ancestry has been shown to be a strong influence on these gene expression and may in part explain the failure of the EMBRACE trial for Belimumab, which is known to be efficacious in those of European ancestry(205,206). Within this thesis modules annotated to interferon, myeloid, inflammation, plasmablast and erythropoiesis function were included based on this preceding work. Indeed, rituximab response has been stratified by hierarchical clustering analysis of these gene expression scores in non-European ancestry patients(203). Additionally, within European ancestry patients hierarchical clustering based on these scores produce clusters which were able to segregate global disease activity and renal involvement, but not responsiveness to rituximab(203).

The above suggests that gene expression-based disease reclassification is a viable and effective tool in better subdividing ANA-RMD. However currently hybridised methods taking in more clinical and basic laboratory data may have a role in enhancing its efficacy.

11.8 THE APPLICATION OF MACHINE LEARNING TECHNIQUES TO RECLASSIFY CLINICAL DATASETS

11.8.1 K-means, hierarchical clustering and Gaussian mixture modelling

Traditional k-means clustering has been in used in clinical and non-clinical reclassification projects for many years having been first described by MacQueen in 1967(207). It is a partitional clustering technique which divides data points up into non-overlapping clusters(208,209). Each cluster is represented by a mean of cluster members (centroid) and a similarity measure that quantifies the difference between multivariable data. The expectation maximisation algorithm is used to calculate the cluster membership that minimises the Sum of Squared Errors (SSE). A key limitation of k-means clustering is its inability to handle uncertainty when a single datapoint sits on the boundary between two centroids(208,209). Additionally, k-means clustering can only generate circular or hypersphere-based clusters, therefore when clusters are ellipsoid in shape they are poorly handled by k-means(210). Finally the process by which clusters are derived by k-means clustering is inherently random, and therefore multiple iterations of the same process can yield slightly different results.

Gaussian mixture modelling (GMM) is a partitional clustering technique which generates density based ellipsoidal clusters based on probability density estimations using expectation maximisation(208,209). Clusters are modelled in a Gaussian distribution. The chief disadvantage of GMM is its higher computational cost compared to k-means. A recent study showed k-means to require approximately 1/6 the computational cost of GMM(208). The key advantage of GMM over k-means is that it assigns probabilities to each data point belonging to a specific cluster. This is termed “soft clustering” and allows greater nuance within clustering attribution in comparison to the hard probabilistic attribution of k-means.

Within the paradigm of rheumatology k-means clustering has been used extensively. It has been applied to identify osteoarthritis endotypes based on biomarker stratification, identifying 3 dominant endotypes with 3 different disease phenotypes which may aid in patient stratification for clinical trials(211). It has been applied to multiple different datasets including clinical , flow cytometric and metabolomic data to good effect(212–214) Within ANA-RMDs it was recently applied to identify antibody profiles predictive of SLE disease outcomes in a cohort of 805 patients(215). GMM has also been utilised to address a similar clinical question to this doctorate in reclassification efforts using data from 6 national registries of

ANCA-Associated Vasculitis patients(216). It has also been used in Juvenile Idiopathic Arthritis and Rheumatoid Arthritis research(217,218).

As discussed previously conventional hierarchical clustering methods were applied to 5 key clinical symptoms in registry data from the UK Primary Sjogren's syndrome Registry (UKPSSR) in a highly cited paper(151). In addition to the clinical and trial data outcomes, this study suggests that to an extent reclassification and sub stratification of ANA-RMDs can be carried out using commonly described variables, leading to potentially significant treatment implications.

11.8.2 Spectral clustering

An alternative approach to k-means and GMM is to apply spectral clustering techniques. These generally allow greater flexibility within the distance metrics leading to improved performance over k-means and GMM, at the cost of greater computational requirements.(219,220).

Spectral clustering techniques generally involve the creation of a similarity matrix from the dataset which is then represented graphically. A Laplacian matrix representing the relationships between data points is then generated, which allows for eigenvectors to be derived and used with traditional clustering methods(219,221,222).

High computational requirements involved in the generation of Laplacian matrices can limit the application of spectral clustering to large datasets without the use of dedicated high power computing systems. Hybrid approaches to spectral clustering have been applied which substitute the eigenvector decomposition component for deep learning techniques but further increase computational requirements.(219,223). Therefore, due to their high computational cost and limited utility over traditional and deep learning-based approaches a spectral clustering approach was considered but ultimately not pursued within this doctorate.

11.8.3 Deep learning and variational autoencoders (VAEs)

Machine learning models at their core discover rules within data to perform a processing task(224). This is achieved through the manipulation of input data with a-priori examples of the expected output and a metric by which the effectiveness of the model can be measured known as a loss function(224). This is achieved by sequential geometric transformations of the data within a latent space measured against a loss function with the aim of sequentially reducing loss and therefore optimising the model against a known expected output. A large volume of pre-defined data transformations can be applied limited only by computational power.

An extension of this process is deep learning whereby these data transformations are stacked on top of each other within successive layers of nodes to form a neural network. Models built in this way with 1-2 layers of data are termed “shallow learning” whereas models with more layers are termed “deep learning”(224). The data transformations carried out within each processing layer are termed “weights” or less frequently “parameters”. Within deep learning models these weights are adjusted by an optimiser based the loss function of the network as described previously. This function generally computes a distance score from the predictions created by a network from a known gold standard target.

There are two key algorithms which are central to the application of deep learning models in general. Backpropagation is central to how the optimiser can sequentially adjust and improve the model by reducing the loss function. Backpropagation was initially described by Werbos in 1974 and later expanded upon in a paper by Rumelhart in 1986(225–227). Initial weight attributions within a deep learning model are random and are sequentially improved through backpropagation. In mathematical terms this process computes the gradient of a loss function within a neural network with respect to the weights of said network(228). Computing this loss is termed a “backward pass”, following which parameters are adjusted in the opposite direction to the gradient to reduce loss. This can be performed in high dimensionality latent spaces and optimisers such as stochastic gradient descent can be helpful when dealing with large and highly dimensional datasets(228,229). Within a neural network the chain rule can be applied to compute the contribution of each weight to the ultimate loss value and therefore act to reduce it through optimisation. From a mechanistic perspective this process is summarised in Figure 2, adapted from(224).

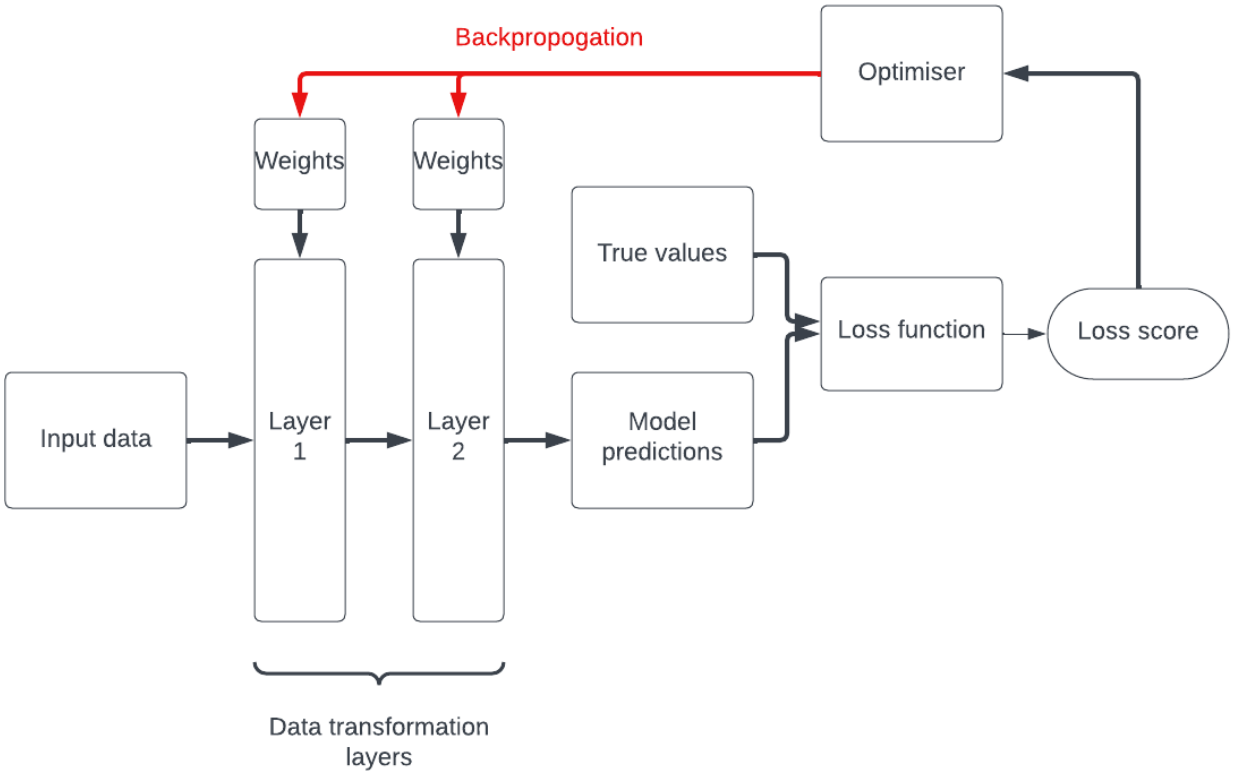


Figure 2: Basic neural network architecture

Variational autoencoders (VAEs) are a subtype of neural network first described in 2014(225,230,231). Autoencoders in macro are a combination of two neural networks, an encoder and a decoder. The encoder aims to compress the data into a lower dimensionality latent space before the decoder attempts to reconstruct the original input. In themselves autoencoders compress to unstructured latent spaces, so aren't generally favoured for data compression where generalisation to unseen data is required(232,233).

Variational autoencoders build on this process by combining Bayesian variational inference with autoencoders(234). VAEs transform the input data into a mean value "μ" and a variance value "σ" rather than fixed code, thereafter during the decoding process the model samples an element of the distribution and can decode it back to its original input(232,235). This adjustment reduces overfitting and ensures that the latent space generated during encoding is sufficiently structured to allow for effective decoding. Because of the incorporation of variational inference, the loss function in VAEs is comprised of a reconstruction loss on the final layer and a regularisation loss in the latent layer which drives the distributions generated by the encoder toward a normal standard distribution. This regularisation loss is

commonly expressed as the Kullback-Leibler (KL) divergence(233,235). The basic structure of VAEs is summarised in Figure 3. Data compression using VAEs in this way can be useful for a myriad range of tasks including financial options trading, electrocardiogram analysis and image denoising(234). Within the biomedical field this approach has been applied to identify small molecule metabolites within complex samples using the DarkChem algorithm(236) and the analysis of single cell gene expression data(237) among others.

Oncology research has utilised VAEs extensively in disease prediction and stratification(238). Combined datasets comprising gene expression, DNA methylation, and Micro-RNA data were analysed by VAEs and other autoencoder types before clustering was performed using partitioning around medoids (PAM) and k-means clustering on the bottleneck layer. Autoencoders in general (Including VAEs) were shown to outperform traditional data compression techniques such as PCA, kernel PCA, and sparse PCA in the subtype detection of four malignancies: Glioblastoma multiforme, colonic adenocarcinoma, renal clear cell carcinoma, and invasive breast cancer(239). A similar approach of combining VAEs with k-means clustering was applied to multi-omics data in neuroblastoma, identifying two prognostic subtypes which were validated in two independent datasets(240). The application of VAE architecture on a transcriptomic dataset (DR.VAE algorithm) was also shown to outperform standard classification methods in predicting drug responsiveness in cancer patients in a Toronto cohort(241). This process of data compression using the VAE followed by Gaussian mixture modelling (GMM) based clustering techniques to the latent space has also been replicated elsewhere(242). A combined approach using a VAE and GMM has been used with single-cell RNA sequencing data to determine the cell developmental trajectories of human preimplantation embryos(243).

Another approach to data compression and clustering is to use deep generative models such as VAEs to preserve hidden hierarchical dependencies in data structures and then use GMM based clustering on the resultant latent space. This is termed deep embedded clustering (DEC)(219). Given the unsupervised nature of the application the network is trained using an iterative auxiliary target distribution derived from the present soft cluster assignment. This utilises a method similar to t-distributed stochastic neighbour embedding (t-SNE) whereby a centroid-based probability distribution is derived, and the KL divergence is minimised in order to improve cluster assignment and feature representation(219).

The extensive application of VAE models within oncology and their inherent superiority to classical autoencoders in latent space projection and subsequent clustering analysis suggest they are an appropriate methodological approach to the problem of ANA-RMD

classification. Within rheumatology the application of VAE models is more limited with a PubMed search yielding only one result for the terms “variational autoencoder” and “rheumatology”. This being an analysis of salivary gland DNA methylation profiles using VAE encoding and hierarchical clustering applied to the bottleneck latent space. This was able to identify a high and low disease activity cluster with associated differences in DNA methylation status(244). This and the synergy between the roles of targeted therapeutics in rheumatology and oncology support the novelty and suitability of applying this methodology to ANA-RMD reclassification.

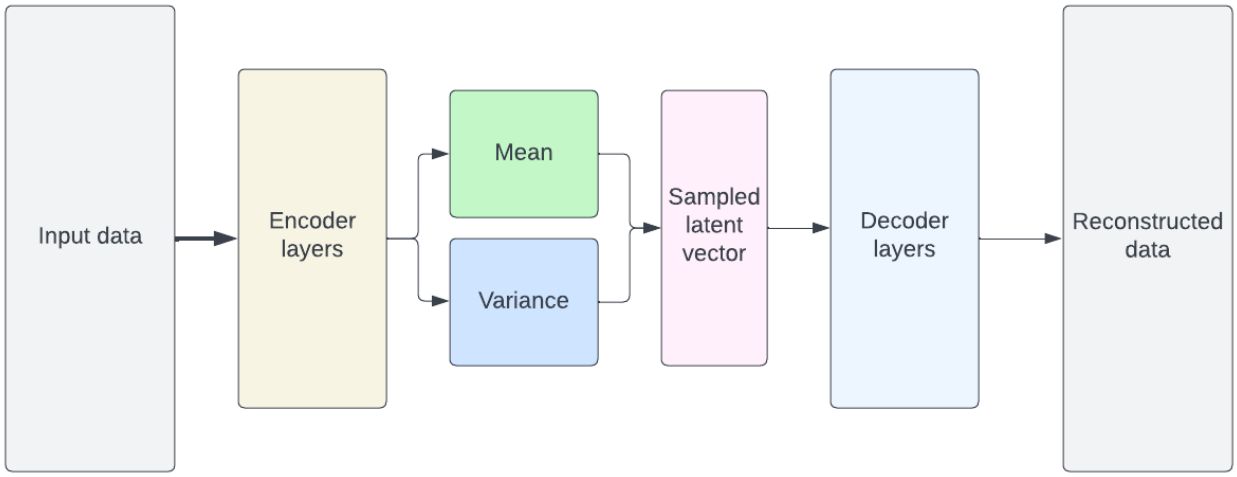


Figure 3: Variational autoencoder basic structure

11.8.4 Traditional dimensionality reduction techniques : Principal Component Analysis

Dimensionality reduction and data compression can be achieved without the use of advanced computational methods through principal component analysis (PCA). This process applies linear transformations to a dataset to create projections of the original data and is discussed in more detail within the general methods section of this thesis.

11.8.5 Natural language processing as an adjunct to biomarker-based reclassification

To date the few attempts made to reclassify ANA-RMDs, and indeed many other disease areas focus largely on biomarker-based strategies. These approaches overlook the significant quantity of clinical data gathered within clinical consultations.

At its core NLP maps the statistical structure of language by examining relationships between words, sentences, and paragraphs. Preparation of text for processing in this way

involves segmentation of the text into words, characters or overlapping groups of multiple words or characters termed N-grams(245). This process is termed tokenization and generally uses one of two approaches, 1-hot encoding or word embedding.

11.8.6 Text tokenization approaches

1-hot encoding is where a unique integer index is associated with each word within a text string and the integer index is converted into a binary vector of size N where N is the size of the vocabulary in question(245). This generates very sparsely populated datasets with very high dimensionality data which may not be desirable where data compression is preferred. However, in some applications this may be sufficient.

Alternatively, tokenisation with word-embedding techniques results in lower dimensionality floating point vectors which are densely populated. Word embeddings can be generated using pre-trained model architectures such as Word2Vec(246) or Global Vectors for Word Representation (GloVe)(247). Word2vec was developed in 2016 by Mikolov et al(246). At its core it is a feedforward unsupervised shallow neural network consisting of a single hidden layer and a single dense layer trained with backpropagation. Word2vec identifies semantic relationships between using two key algorithms. With continuous bag of words (CBOW) the surrounding words are used to predict the word of interest. Both approaches can be applied to tokenise text for subsequent analysis. Data input validation approaches such as spell checking can be applied when using these algorithms, however these may struggle with spelling errors in complex medical terminology, which can limit their effectiveness.

Natural language processing (NLP) through 1-hot encoding, pre-trained embeddings such as Word2Vec or via tailored embedding in neural networks provides the means with which this data can be incorporated into data compression approaches (VAEs) and subsequent traditional clustering algorithms such as k-means and hierarchical clustering.

11.8.7 Applications of NLP within rheumatology

Within the field of rheumatology this technique has been used to extract key rheumatoid arthritis outcomes from clinical notes and to process patient narrative data within a Japanese SLE cohort(248,249). NLP pipelines were shown to extract rheumatoid arthritis outcomes from multiple electronic record systems effectively(248). Within the SLE cohort the incorporation of NLP applied to patient narratives provided additional insight into the patient experience by identifying recurring key concerns such as pain, disability and mobility issues(249).

The closest parallel to this project was the application of NLP techniques to a 7853-patient cohort of potential axial spondyloarthritis patients from two large Boston tertiary care hospitals(250). Within this study traditional ICD-10 codes for diagnosing axial spondyloarthritis were compared to NLP based algorithms. Both supervised and unsupervised algorithms were created to identify high probability axial spondyloarthritis patients from this cohort. Both supervised and unsupervised algorithms incorporating NLP outperformed traditional ICD-10 criteria (AUROC 0.80-0.87)(250). This suggests that this methodology either alone, or in combination with biomarker variables could be effective in reclassifying ANA-RMDs in a similar manner.

11.8.8 The limitations of a deep learning-based approach

Although deep learning and machine learning techniques in general are well suited to the proposed task of reclassification they are not without limitations. It is important to note that deep learning models have no functional understanding of their inputs, they merely minimise a loss function through modifying the weights ascribed to different functional layers via backpropagation(251).

Moreover, any outputs from a deep learning model need to be analysed in the context of their clinical and patient reported outcomes. Adversarial examples can be fed into a deep learning network with the specific intention of tricking the network into misclassifying them and may be useful when assessing the generalisability and robustness of any generated model(251). Deep learning models also contain no capacity for abstraction where multiple known concepts are merged to conceptualise something not previously experienced by the model(251). In the proposed application this means that any model generated will be unable to abstractly predict beyond the confines of its training dataset. This process is termed ‘extreme generalisation’ by Francois Chollet and contrasts with the ‘local generalisation’ possible within a deep learning model(251).

11.8.9 Comparison of clustering efficacy

In contrast to hierarchical techniques k-means clustering necessitates a priori determination of cluster number (k). Multiple methods exist to estimate this including silhouette analysis, Monti consensus clustering, GAP-statistic, progeny clustering and CLEST(252). Monti consensus clustering is widely utilised but has been shown capable of producing false positive structures in k=1 model data(252,253). Consequently, the M3C package was created to incorporate Monte Carlo Simulations and address other limitations in Monti consensus clustering. These approaches are discussed in detail within the general methods section.

11.9 MOLECULAR DISEASE STRATIFICATION AND BASKET TRIAL APPROACHES ARE WELL ESTABLISHED IN OTHER MEDICAL SPECIALTIES

11.9.1 Molecular disease stratification and basket trial application in oncology

There is a global drive toward precision medicine as a function to improve disease outcomes and better allocate resources. Within the UK this was noted in the 2019 NHS long term plan which incorporates a plan to offer whole genome sequencing to children with malignancy to expedite access to targeted therapies(254,255). There is considerable divergence between medical specialties to what extent this approach has been applied. Perhaps the most effective example of its application to date has been within oncology.

Within oncology the emphasis on the carcinomatous organ or origin is increasingly felt to be less relevant as targeted treatments emerge(256). In conjunction with this the traditional model of randomised controlled trials is increasingly felt to be inappropriate when evaluating treatments for rare disease types in fast moving fields such as oncology and, by extension, rheumatology. In addition to overlapping pathological pathways, recruitment of sufficient numbers in trials of rare diseases to provide power to results is often impractical. As a result, switching to a basket trial design where malignancies with a common pathological marker but different primaries or histology is an established concept in oncology(255). This has expanded dramatically within oncology over the last decade with a 2019 landscape analysis showing an increase from 2 basket trials in 2009 to 49 examples by 2019(255,257).

Key examples of large basket trials in the field of oncology include the NCI-MATCH trial, a phase 2 basket trial for advanced solid tumours, lymphoma or multiple myeloma which has progressed on previous therapy. NCI-MATCH screened nearly 6000 patients and recruited 1593 patients to 38 sub studies. Each sub study consisted of a phase 2 trial of a therapy matched to a specific genomic alteration. Within the first 27 of these sub studies the signal seeking objective was met in 7 cases (25.9%)(258,259).

Another NCI basket trial within oncology was the NCI-MPACT trial, carried out in refractory solid organ malignancies(260). Here patients were assigned to receive a targeted therapy (Veliparib + Carboplatin, Veliparib with temozolomide or Adavosertib with Carboplatin, Evorlimus or Trametinib) based on 3 aberrant gene signalling pathways of interest (DNA repair, PI3K signalling or RAS/RAF/MEK pathway) or a non-paired therapy. The results of this trial were disappointing with only 2% of the 49 patients within the treatment arm showing a partial response(260). Within their discussion the authors note that better annotation of predictive biomarker and more effective targeted therapies were required to improve further

study design. Nevertheless, other trials in oncology have led to basket therapeutic indications as shown below.

11.9.2 Molecular disease stratification and basket trial application in rheumatology

Molecular stratification of disease in rheumatology has been carried out within individual disease groups. The best example of individual ANA-RMD sub stratification in this way is the work carried out by Daniel Toro in SLE within the PRECISESADS group(261). Here whole blood samples of SLE patients were analysed and sub stratified into 9 separate modules based on gene expression scores. 3 of these modules (Interferon, Neutrophil/Inflammatory and Plasma cell/Cell cycle) showed high levels of dysregulation relative to healthy controls. Additional complexity was noted as some modules were more frequently dysregulated in the presence of others. Interferon for example was noted to be more frequently dysregulated in the context of concurrent Neutrophil module dysregulation. Using further data from a therapeutic trial of Tabalumab certain associations with clinical outcomes were noted. The neutrophil/inflammatory module was noted to be associated with renal manifestations of SLE.

Broader cross diagnostic disease stratification has been carried out by Barturen et al as referred to previously(262). There has however been little translational work taking this concept towards basket trial design currently, with the only current example being a small-scale trail of low dose interleukin-2 across 13 different autoimmune diseases(263).

11.9.3 Basket indications lead to licenced therapies

Within oncology several targeted therapies have been licenced for basket indications based on the pooled results of multiple individual trials. Pembrolizumab is a PD1 inhibitor licenced by the FDA for all metastatic tumour with microsatellite instability biomarker (MSI-H) or deficient mismatch repair (dMMR)(255,264,265). In a similar manner basket indications have been approved for the TRK inhibitor Larotrectinib(266–268), the TKI inhibitor Entrectinib(269) and the BRAF inhibitor vemurafenib(270).

Although there have not yet been any trials in basket indications in ANA-RMDs, there is precedent for such an approach in a related disease area. Nintedanib was evaluated in a phase III clinical trial in a basket of patients with a common feature: interstitial lung disease, which could be secondary to any of a range of diseases(271). This has led to licensing of nintedanib for this indication.

These examples illustrate that licencing of targeted therapeutics in a similar manner within rheumatology is a reasonable prospect. However, at present no such indications exist, and therefore the suggested approach has the potential to be transformative within the field.

11.10 SUMMARY OF RATIONALE FOR THIS APPROACH

Currently we conduct trials of new therapies by in patients classified by legacy clinical diagnoses. This is exemplified by the introduction of rituximab in which the same drug was trialled in 3 ANA-RMDs, and all 3 trials were negative, albeit with good-responding subgroups in each(272–274). The problems of trials in legacy diagnoses are therefore:

1. Conducting one trial for each disease is difficult since these diseases are uncommon
2. Many patients who do not meet any classification criteria are not eligible for trials and have no evidence-based therapy options
3. Trials each include a mixture of good and poor responders, since eligibility selected by legacy clinical diagnosis rather than by a label-agnostic predictor or biomarker

In terms of analytical techniques, deep learning techniques provide additional utility over traditional methods in terms of better preservation of non-linear and hierarchical relationships between covariates during data compression. Variational autoencoders have been extensively used in this role and allow for effective integration of multiple data formats. Where possible the introduction of NLP components to the modelling would allow better compression of the available data and would avoid significant feature engineering by hard coding N-grams or high dimensionality sparse data created through 1-hot encoding of categorical variables.

In terms of path to patient benefit there is an established precedent within oncology whereby multiple basket trials have been completed or are in progress. Moreover, there are multiple examples of therapeutics with basket indications within oncology and a further example in the form of Nintedinib for pulmonary fibrosis. Therefore, identification of therapeutic baskets in the manner proposed would be novel within rheumatology but has a precedent within medicine generally.

I hypothesise that we can re-group patients to undertake basket trials of all patients suitable for a certain therapy. An ideal example of this is ANA+ arthritis, identifiable through clinical or ultrasonographic examination. There are patients with inflammatory arthritis within each diagnosis. We could instead create a “basket” containing these patients from each disease which has the following advantages:

1. One trial would cover a large proportion of patients who will benefit, which is advantageous in uncommon diseases, more people will get access to effective treatments.
2. Patients currently ineligible for trials would attain proven treatment options
3. Trial populations would be more homogenous than legacy diagnoses (e.g. SLE trials include both skin and joint disease). This allows more appropriate inclusion criteria (based on joint swelling or ultrasound) and better outcome measures (based on a single organ of interest).
4. We may enrich the trial population for patients predicted to respond well (e.g. biomarker-positive patients)

This single-variable example illustrates the concept, but these diseases are likely too complex for such a simple classification to be applicable across the board. Also, some therapeutic approaches are not biologics with accompanying biomarkers. For example, patients with low disease activity may require rehabilitative or pain-management approaches.

12 AIMS, OBJECTIVES AND HYPOTHESES

The overarching theme of the project is to better define ANA-RMDs using deep learning and traditional approaches. To account for all the reclassification features described above, my approach may involve sub stratification by flow cytometric parameters, gene expression scores, clinical parameters in a hybrid clinical and biomarker approach across 2 large cohorts.

12.1 AIMS

To improve access to appropriate treatment for people with ANA-RMDs by defining basket populations for better clinical trials.

12.2 HYPOTHESES

The overall hypothesis guiding this research is that the ANA-RMD population can be better subdivided using advanced computational techniques than via the existing classifications in place.

Within this paradigm the following specific sub-hypotheses are offered:

1. Existing diagnoses fail to classify all individuals with clinically significant ANA-RMDs
2. Existing diagnoses fail to adequately subdivide patients into discrete groups
3. Classification of people according to their predominant clinical manifestation, immunophenotype or therapy requirement will lead to more homogenous subgroups (ARC classes).
4. ARC classes predict long term outcome and PROs better than legacy diagnoses.
5. Cardinal features of the most the most clinically impactful ARC classes will be responsive to therapy, irrespective of their underlying legacy diagnosis

12.3 OBJECTIVES

12.3.1 Results 1: Preliminary Pure flow cytometry based reclassification of ANA-RMDs using established clustering techniques

1. To achieve familiarity with basic data manipulation and clustering techniques in R prior to subsequent more detailed analysis
2. To evaluate the utility of pure flow cytometry based clustering in DEFINITION cohort B using partitioning around medoids (PAM) clustering as part of a preliminary analysis.

12.3.2

12.3.3 Results 2: Unsupervised machine learning based reclassification of ANA-RMDs. Defining ANA-Arthritis, a clinical and biomarker defined population suitable for clinical trials.

1. To evaluate the prevalence of ANA+ arthritis as a potential clinical basket in DEFINITION
2. To determine whether clinical impact is comparable across ANA-RMD diagnoses in DEFINITION
3. To determine whether immunological signatures are similar across ANA-RMD diagnosis in patients with ANA+ arthritis
4. To use Principal Component Analysis and Gaussian Mixture Modelling to identify enriched baskets of ANA+ arthritis patients for future clinical trials

12.3.4 Results 3: Variational Autoencoder based reclassification of ANA-RMD across a large European cohort, hybrid clinical and biomarker-based clustering.

1. To use a variational autoencoder (VAE) based approach to compress data from the PRECISESADS cohort.
2. To apply traditional clustering techniques to the compressed dataset and derive new ANA-RMD classes (ARC) within PRECISESADS
3. To replicate this process within DEFINITION
4. To assess the cross-sectional clinical impact of ARC membership and compare to legacy diagnosis membership in both cohorts
5. To assess the gene expression impact of ARC membership and compare to legacy diagnosis membership in DEFINITION
6. To assess the flow cytometric impact of ARC membership and compare to legacy diagnosis membership in DEFINITION
7. To assess the patient reported impact of ARC membership and compare to legacy diagnosis membership in DEFINITION
8. To assess the 5-year long term follow up impact of ARC membership and compare to legacy diagnosis membership in DEFINITION

12.3.5 Results 4: Preliminary exploration of causal and predictive models using gene expression for longitudinal disease activity across ANA-RMDs

1. To use DAG analysis to identify measured and unmeasured factors governing the impact of key gene expression scores on disease activity and flare then construct a causal model
2. Evaluate the explanatory power of this model using linear mixed effects modelling
3. To construct a predictive model including demographic and clinical variables, IFN and other gene expression scores to predict 3- and 6-month disease activity
4. To assess the responsiveness of specific gene expression scores to changes in therapeutics and glucocorticoid exposure using linear mixed effects modelling.

13 GENERAL METHODS

13.1 INTRODUCTION

This chapter outlines the methodologies applied within the analysis of the DEFINITION and PRECISESADS cohorts within the scope of this doctorate.

13.2 THE DEFINITION & PRECISESADS COHORTS

13.2.1 The DEFINITION cohort

The DEFINITION cohort was set up to address the utility of Interferon gene expression signatures in disease stratification across the spectrum of ANA-RMDs. Its initial stated aims were to evaluate the two previously validated IFN I scores (IFN Score A and IFN Score B) derived from factor analysis of a previous dataset(95). To account for variation in numbers of IFN producing cells with changes in disease activity, flow cytometric markers of IFN expression such as Tetherin and SIGLEC-1 were also collected.

As a function of the above detailed composite disease activity measurements, patient reported outcomes, gene expression scores, laboratory data and flow cytometric data were collected. Clinical data collection CRFs are included within the appendix of this thesis.

The overall aim of the DEFINTION study was to better define the role of type I IFN in ANA-RMDs by assessing IFN-I status at 3 timepoints.

- Cohort A assessed whether IFN-I scores A and B were associated with increased risk of progression from ANA “At Risk” status to classifiable ANA-RMDs.
- Cohort B assessed whether IFN-I expression was associated with specific disease phenotypes or disease severity in a cross-sectional study
- Cohort C measured change in IFN-I expression and disease activity at baseline, 3 and 6 months to assess whether IFN-I expression had prognostic benefit or dynamically responded to changes in disease activity over longer term follow up.

13.2.2 The PRECISESADS cohort

The PRECISESADS cohort was set up to utilise OMIC and bioinformatic approaches to derive new disease classifications across the spectrum of autoimmune disease. The project recruited 2656 patients between December 2014 and October 2017 from 19 sites in 9

European countries. These included sites in Austria, Belgium, France, Germany, Hungary, Italy, Portugal, Spain and Switzerland(275).

In total the project received €23,098,292 in funding and recruited 377 Rheumatoid Arthritis, 470 SLE, 402 SSc, 385 pSS, 99 MCTD, 106 PAPs, 166 UCTD patients and 651 healthy controls. Excluding the rheumatoid arthritis component this represents the largest mixed cohort of ANA-RMD patients in Europe.

Rich clinical data was collected with the PRECISESADS study. In terms of OMIC data, genetic, epigenetic, transcriptomic, flow cytometric, metabolomic and proteomic data was collected.

To date outputs from this cohort have included work in pSS which has demonstrated risk loci via GWAS analysis, Ro60 antibody associated inflammatory gene expression signatures spanning multiple ANA-RMDs, IFN upregulation and B and plasma cell gene dysregulation patterns in lupus nephritis(276–278). Cross diagnostic reclassification using pure omics approaches has also been trialled as noted earlier within the introduction, but to date no clinical data has been incorporated into reclassification approaches(262).

13.3 ETHICAL APPROVAL

13.3.1 DEFINITION ethical approval

All individuals provided informed written consent and this research was carried out in compliance with the Declaration of Helsinki. This study was approved by the National Health Service Health Research Authority (REC Ref: 17/YH/0166). Healthy control participants' peripheral blood was collected under the study number 04/Q1206/107. All experiments were performed in accordance with relevant guidelines and regulations. University of Leeds was contracted with administrative sponsorship.

13.3.2 PRECISESADS ethical approval

The PRECISESADS project was approved by multiple ethical committees as cited in previous work(279). These include: Comitato Etico Area 2 (Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico di Milano and University of Milan); approval no. 425bis Nov 19, 2014, and no. 671_2018 Sep 19, 2018; Klinikum der Universitaet zu Koeln, Cologne, Germany. Geschaftsstelle Ethikkommission; Pôle de pathologies rhumatismales systémiques et inflammatoires, Institut de Recherche Expérimentale et Clinique, Université catholique de Louvain, Brussels, Belgium. Comité d'Èthique Hospitalo-Facultaire; University

of Szeged, Szeged, Hungary. Csongrad Megyei Kormanyhivatal; Hospital Clinic I Provicia, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain. Comité Ética de Investigación Clínica del Hospital Clínic de Barcelona. Hospital Clinic del Barcelona; Servicio Andaluz de Salud, Hospital Universitario Reina Sofía Córdoba, Spain. Comité de Ética e la Investigación de Centro de Granada (CEI – Granada); Centro Hospitalar do Porto, Portugal. Comissao de ética para a Saude – CES do CHP; Centre Hospitalier Universitaire de Brest, Hospital de la Cavale Blanche, Avenue Tanguy Prigent 29609, Brest, France. Comite de Protection des Personnes Ouest VI; Hospitaux Universitaires de Genève, Switzerland. DEAS – Commission Cantonale d'ethique de la recherche Hopitaux universitaires de Geneve ; Andalusian Public Health System Biobank, Granada, Spain; Katholieke Universiteit Leuven, Belgium. Commissie Medische Ethiek UZ KU Leuven /Onderzoek; Charite, Berlin, Germany. Ethikkommission; Medizinische Hochschule Hannover, Germany. Ethikkommission.

13.4 DEFINITION STUDY DESIGN

Patients within DEFINITION were recruited to 3 arms (A,B and C). Cohort A comprised an 'At-risk' population of patients with risk factors and incomplete symptoms of ANA-RMD which were not utilised within this project. Cohort B comprised patients a known diagnosis of UCTD (SLE, pSS, IM, MCTD, SSc or UCTD) and evidence of stable disease over the preceding 12 months. These patients were recruited to a cross-sectional observational study. Those with evidence of labile disease activity, defined as physician defined evidence of labile disease activity within the preceding 12 months.

13.4.1 DEFINITION sample size calculations

13.4.1.1 DEFINITION Cohort B: Cross-sectional study

The sample size for the cross-sectional study was pragmatic, to ensure there are enough patients eligible to enter the longitudinal study (n=150 – see below). We estimated that approximately 50% of patients will be eligible for the cross-sectional study therefore the cross-sectional study required 300 patients, 60 of whom could be recruited from study A (at risk CTD).

300 patients were recruited. Of these, 294 had baseline clinical and biomarker data for analysis. Diagnoses were SLE (n=104), UCTD (n=111), pSS (n=33), MCTD (n=13), SSC (n=14), Inflammatory Myositis (n=19). Baseline characteristics are shown in Table 10.

13.4.1.2 DEFINITION Cohort C: Longitudinal study

50% of patients were estimated to have active disease at 6 months. Based on rules of thumb stating that there should be 5-10 patients in the smallest outcome category per predictor variable, with 12 predictor variables this necessitated a total of 150 patients. Recruitment was substantially affected by the COVID-19 pandemic, leading to global under recruitment and reduction in follow up visit attendance.

From an overall sample size of 124 ANA-RMD patients (SLE n=46, UCTD n=49, pSS n=16, IM n=7, SSc n=4, MTCD n=2). Key baseline variables of DEFINTION Cohort C are summarized in Table 16 within Results Chapter 4.

13.5 PHYSICIAN ASSESSMENTS

Clinical assessments in both the DEFINITION and PRECISESADS cohorts were carried out by recruiting rheumatologists with appropriate training in scoring the indices in question. Within PRECISESADS this encompassed scoring the Physician Global Assessment VAS. Within DEFINITION this comprised multiple additional metrics detailed below.

13.5.1 British Isles Lupus Assessment Group (BILAG) 2004 Index

The revised British Isles Lupus Assessment Group Index (BILAG-2004) index is a composite disease activity score which measures 97 disease activity items across 9 clinical domains (Constitutional, mucocutaneous, musculoskeletal, neuropsychiatric, cardiorespiratory, gastrointestinal, ophthalmic, renal and haematological). Activity within each item is scored over the preceding month and compared to the month prior, graded as either new, worse, same, improving, or absent. Each item is graded from A (Severe disease), B (Moderate disease), C (Mild disease), D (Inactive disease) and E (Never affected). An overall grade can be calculated for each domain by taking the highest score of each domain. Additionally, numeric BILAG scores can be calculated as follows: A = 12, B = 8, C = 1 , D/E = 0. This process allows for accurate global assessment of disease activity within SLE. BILAG-2004 has been extensively validated in SLE trials and is currently a component of the Systemic Lupus Erythematosus Responder Index (SRI-4) which has been approved by the FDA for drug trials in SLE(280).

13.5.2 Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K)

The Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K) is another composite disease activity score in SLE developed in 2002 as a progression of the earlier SLEDAI disease activity measurement(281). SLEDAI-2K comprises a list of 24 organ manifestations of SLE and laboratory variables with different score weightings, the score

being the sum of the responses. It has been widely used in clinical trials, being a core element of the SLE Responder Index (SRI-4)(280). Unfortunately, it has been shown to be poorly responsive to partial changes in clinical statues and work comparing the ability of the SLEDAI-2K and BILAG in detection of active disease requiring changes to therapy suggested that the BILAG-2004 index was superior in this respect(282). Nevertheless, the SLEDAI-2K allows for rapid documentation of a patient's SLE activity captured as a snapshot in time.

13.5.3 Myositis Disease Activity Assessment Tool (MITAX) 2009

MITAX is a disease activity assessment tool used to classify disease activity in IIM patients and to help guide therapeutic adjustments. It was first published in 2004 and took a derivation of the BILAG-2004 index whereby patients are scored with an intention to treat approach(283–285). It was subsequently revised in 2008 and comprises 7 organ domains (Constitutional, cutaneous, muscular, skeletal, gastrointestinal, cardiac and pulmonary) which mirroring BILAG are scored as either not present, improving, same, worse or new compared to the previous month. This allows for the assessment of muscular and extra muscular disease activity. MITAX has been validated and has been used in several clinical trials, notably the MYOJAK trial(286).

13.5.4 EULAR Sjögren's syndrome disease activity index (ESSDAI)

The EULAR Sjogren's Syndrome Disease Activity Index (ESSDAI) is a composite disease activity score developed in 2009 for Primary Sjogren's Syndrome covering multiple organ manifestations. This addresses the heterogeneity of presentations within pSS. It comprises 12 domain items (Cutaneous, pulmonary, renal, articular, muscular, peripheral nervous system, central nervous system, haematological, glandular, constitutional, lymphadenopathy and biological)(287).

The ESSDAI has been used as a primary endpoint in several pSS trials(288,289). It has also demonstrated good efficacy in phase 2 trials of Ianalumab in pSS as part of the NEPTUNUS-2 trial(290).

ESSDAI scoring represents the best validated and most widely utilised composite disease scoring system for pSS, overlapping clinically with the MITAX, SLEDAI-2K and BILAG-2004 index in many domains. This allows for composite disease activity across the spectrum of ANA-RMDs.

13.5.5 Physician Global Assessment (PGA) score

A global assessment of disease activity was scored by the physician on a 100mm scale ranging from nil disease (0) to the most active the disease could be (100). This scoring system was not patient specific but instead represents the highest possible disease activity for any given patient with the same diagnosis. Scores were then scaled from 0-10cm for further data analysis. This scoring system has been used effectively in SLE and rarer autoimmune connective tissue diseases and is a component in both the ESSDAI and LLDAS composite disease scores(291–294).

13.6 PATIENT REPORTED OUTCOMES

13.6.1 36-Item Short Form Health Survey (SF-36)

The 36 Item Short Form Health Survey Questionnaire (SF-36) is an instrument designed to evaluate patient reported quality of life outcomes developed initially in 1992(295). It measures 8 dimensions (Physical functioning (PF), role physical (RP), bodily pain (BP), general health (GH), vitality (VT), social functioning (SF), role emotional (RE) and mental health (MH). Additionally, two composite scores, the Physical Component Summary (PCS) and the Mental Component Summary (MCS) representing the composite physical and mental health impacts of the disease(296). Multiple R packages exist to perform these calculations in research(297). The SF-36 has been used in many trials with and outside of rheumatology, within the field of ANA-RMDs it has been used extensively across multiple existing diagnoses(298–300).

13.6.2 Euroqol 5 Dimensions (EQ-5D) score

The Euroqol 5 Dimension score is a patient reported outcome measure which measures health status, defined as “the value assigned to duration of life as modified by the impairments, functional status, perceptions and social opportunities that are influenced by disease, injury, treatment or policy”(301). This impact is measured in 5 dimensions (Mobility, Usual Activities, Self-Care, Pain & Discomfort and Anxiety & Depression) (302). The EQ-5D was designed to be broadly applicable across healthcare applications and questions reflect basic activities of daily living such as washing and dressing. Each question is scored on a scale ranging from no impairment to mild, moderate, severe, extreme impairment or inability to complete the task. A visual analogue scale is also incorporate and global index scores can be calculated representing the global impact of disease on the individual ranging from 0 (dead) to 1 (perfect health)(303).

13.6.3 Functional Assessment of Chronic Illness Therapy (FACIT) – Fatigue

FACIT-Fatigue is a patient reported outcome score initially developed to quantify fatigue related to anaemia in cancer patients(304). It comprises a 13-item questionnaire completed by the rater and has been used in multiple studies across a wide array of medical specialties. In total since its development, it has been used in over 150 studies and over 40,000 people(304). In the context of rheumatology, it was shown to be superior to the SF-36 vitality subdomain in describing disease related fatigue in SSc patients, hence it's inclusion in this work (305).

13.6.4 ICECAP-A (ICEpop CAPability measure for Adults)

The ICEpop CAPability measure for adults (ICECAP-A) is a patient reported outcome measure which attempts to quantify a person's perceived ability to achieve key life outcomes and capabilities(306,307). This is quantified across 5 domains ranging from stability to attachment, autonomy, achievement, and enjoyment. It is delivered as a patient completed questionnaire and scores can be converted to a single index value for analysis. These index values are subsequently measured against country specific tariffs for analysis(308,309) . Within rheumatology this measure has been correlated with MSK health problems at the population level but has seen limited use within ANA-RMDs(310).

13.6.5 Work Productivity and Activity Impairment (WPAI) instrument

The WPAI was developed in 1993 and is a composite patient reported score used to quantify the impact of disease on an individual's employment productivity. It collates metrics on 4 key areas, absenteeism, presenteeism, global work productivity loss and the impairment of one's daily activities outside of work(311). It has been applied broadly to multiple research questions across ranging from multiple sclerosis to chronic obesity(312,313). Within rheumatology it has been used and validated in both Spondyloarthritis and Psoriatic Arthritis(314,315). Metrics produced are given as percentages of the 4 key metrics.

13.6.6 Visual Analogue Scales (VAS)

Visual analogue scores allow patients to rate poorly quantifiable qualities such as pain and other quality of life outcomes. They have been in use since the 1920s. In the context of rheumatology, they have been used to describe pain, global health, fatigue, early morning stiffness among many other applications(316,317). They also form part of composite disease scoring systems such as the DAS-28 in rheumatoid arthritis (318)Within the context of the DEFINITION cohort VAS were applied to quantify pain, early morning stiffness, fatigue, global health and arthritis severity; based on their previous use in the literature(319–323).

Issues with VAS measurements are well known, older people are known to struggle to report VAS scores as effectively as young people and many patients report pain on extremes of the scale rather than as a true linear construct(324,325). Despite these weaknesses VAS scores are rapid to complete and in the context of this work are used in the context of overlapping patient reported outcome measures such as the FACIT-Fatigue, EQ5D and SF-36 which adds additional redundancy.

13.7 GENE EXPRESSION SCORES

13.7.1 Gene expression score composition

Gene expression scores used within this thesis were derived from those initially described by Chiche, Bachereau and Jourde-Chiche as described in the introduction and subsequently utilised within work by Carter et al (200–203). A summary table of the gene expression scores is shown in Table 6.

Gene Expression Score	Genes
Interferon Score A	*IFI44L, CCL8, ISG15, XAF1, IFI44, GBP1, IFI27, IRF7, CXCL10, CEACAM1, RSAD2, IFIT1*
Interferon Score B	*SERPING1, IFIH1, IFI16, PHF11, LAMP3, NT5C3B, TRIM38, UNC93B1, BST2, SOCS1, SP100, TAP1, STAT1, UBE2L6*
Plasmablast M7.7 Score	*ATP5G1, CDC2.CDK1, CHPF, CIT, GLICCI1, GTSE1*
Plasmablast M4.11 Score	*TNFRSF17, TXNDC5, IGJ.JCHAIN*
Plasmablast Score	*ATP5G1, CDC2.CDK1, CHPF, CIT, GLICCI1, GTSE1, TNFRSF17, TXNDC5, IGJ.JCHAIN*
Myeloid Score	*CD55, BST1, LILRA3, CEACAM4, MKNK1, FLJ20273.RBM47., HCK, FLOT1, CKAP4, BCL6, GPR97, ITGAM, AQP9, ADAM19, HMGB2, NCF4, IL1RN, CD63, IL17R.IL17RA., MMP25, PBEF1.NAMPT, NFIL3*
Erythropoiesis Score	*DNAJA4, IFIT1L.IFIT1B, UBXD1, RAB2B, BCL2L1, GATA1, UBE2O, MAP2K3, PIP5K2A, FLCN, HPS1, MCOLN1*
Inflammation Score	*SOCS3, IRAK3, IL18R1, CLEC4D, MCEMP1, SERPINA1, IL1R2, MCTP2, FKBP5, OSM, KREMEN1, FCAR, CR1, SLPI, MAPK14, PGLYRP1, ANXA3, S100P, MMP9, S100A12, GRB10*

Table 6: Gene expression score composition

Quantification of gene expression varied between the DEFINITION and PRECISESADS cohorts. Within DEFINITION was measured using a 96 panel Taqman array on blood from ANA-RMD patients collected in Tempus tubes. ΔCT values were normalised against the reference gene peptidylprolyl isomerase A (PPIA). Within PRECISESADS RNA-Seq data pertaining to the gene expression scores was analysed for quality control as per Figure 4. Data was then normalised using the Voom package in R to account for high heteroscedasticity within RNA Sequencing data. This process involved normalisation of data and log transformation of the normalised counts in Voom then estimation of the mean-variance relationship prior to the assignment of precision weights to each observation using the Voom algorithm(326). Median gene expression scores were then calculated from each of the DEFINITION and PRECISESADS samples and taken forward for analysis.

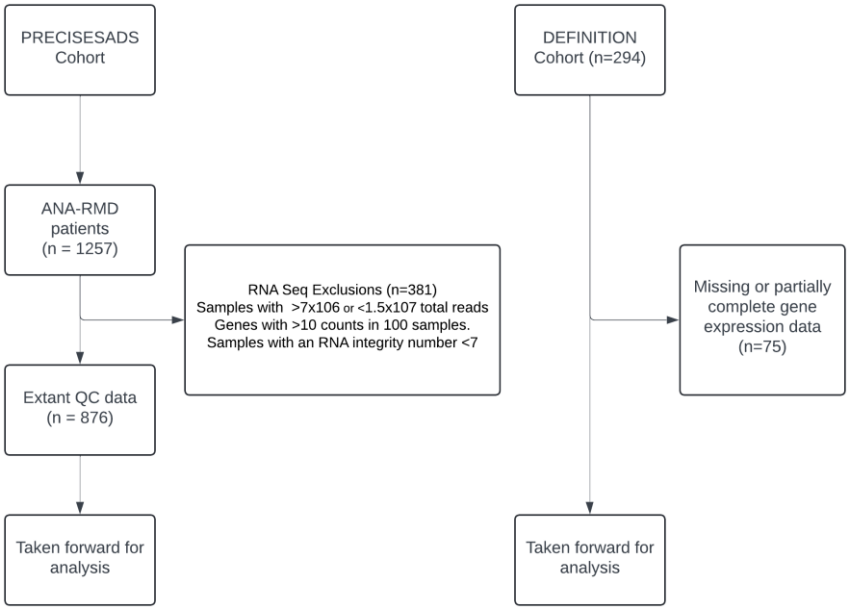


Figure 4: Quality control analysis of PRECISESADS RNA-Seq data

13.8 Flow Cytometric Variables

Within DEFINITION flow cytometric analysis was carried out with the primary aim of assessing levels of CD317 (Tetherin) and CD169 (SIGLEC-1) expression in patients with ANA-RMDs. These markers have been associated with increased IFN-I and SLE disease activity in previous work within this group(327).

Processing samples was carried out according to the following protocol

1. Peripheral blood mononuclear cells (PBMCs) isolated from EDTA by density gradient media (Leucosep).
2. Remaining RBCs lysed with ammonium chloride and washed with Phosphate buffered saline (PBS)
3. $1x10^6$ cells used for flow cytometry, incubated in IgG+ mouse serum blocking buffer
4. Antibodies added as per manufacturer recommendations
5. Cells washed a suspended in 400µL buffer and run on CytoFLEX S system.

Gating strategies identified cell populations as per the following markers:

Cell population	Cell Surface Marker
NK bright cells	CD3- CD56++
NK dim cells	CD3- CD56 intermediate
NKT cells	CD3+ CD56+
T-cells	CD3+
B-cells	CD19+
Naïve B cells	CD19+ CD27-
Memory B-cells	CD19+ CD27+
Plasmablasts	CD19+/- CD27+ CD38++ CD3-CD14-
Monocytes	CD14+
Non-Classical Monocytes	CD14+ CD16+
Intermediate Monocytes	CD14++ CD16+
Classical Monocytes	CD14++ CD16-)
SIGLEC-1 on Classical monocytes	CD169 MFI on CD14++CD16-
Tetherin on memory B cells	CD317 MFI on CD19+ CD27+ cells
Tetherin on classical monocytes	CD317 MFI CD14+ CD16+

Table 7: Flow cytometry cell markers

Mean Fluorescence Intensity (MFI) measurement was used to quantify cell surface expression of Tetherin and SIGLEC-1. Cell populations were analysed as a proportion of total PBMC count, to account for variations in PBMC populations between ANA-RMDs. Tetherin measurements were measured on Memory B cells as per previous work(328).

13.9 Traditional Data Compression Approaches

13.9.1 Principal Component Analysis

Principal component analysis (PCA) is a non-machine learning based dimensionality reduction technique whereby principal components are constructed through linear manipulations of the initial variables within a dataset(329–331). The resultant principal components are orthogonal, and the majority of the variance within the dataset is generally accounted for within the first few components(329,331).

To perform this process data is standardised and a covariance matrix is constructed comprising the covariances of all the possible pairs of initial variables. For each pair a positive value is associated with positive correlation whereas a negative value is associated with an inverse correlation. Eigenvectors and eigenvalues are then derived from this covariance matrix and principal components are ranked by their respective eigenvalues. Scree plots can be used to determine how many principal components are to be retained and what proportion of the total variance is explained by the retained principal components.

Subsequently the eigenvectors (loadings) can be analysed for each principal component to understand which of the original variables contribute maximally to each principal component.

PCA has been widely utilised in clinical and non-clinical research(330). It is computationally cheap, simple and requires no hyperparameter adjustment. However, it has some notable limitations, chiefly being that it is unable to perform non-linear transformations so is unable to represent complex nonlinear relationships in large datasets with high dimensionality. It is also sensitive to outliers and within high dimensionality datasets examination of loadings to determine maximal contributors to each principal component can be difficult.

13.10 Autoencoder and Variational Autoencoder Based Data Compression

13.10.1 Data preparation

Prior to construction of a VAE the data is pre-processed by normalisation of values where appropriate and split into training and test datasets, typically utilising an 80:20 split. Models were trained within the training set with the 20% test set used to evaluate performance of the model in data compression and reconstruction. To compare two datasets with different

measurements of gene expression data, PRECISESADS data was normalised as described in section 13.7.

13.10.2 Autoencoder Structure

Simple autoencoder structures are not applied within this thesis but are important to consider in the context of their variational counterparts. At its core an autoencoder aims to compress a dataset into a low dimensionality space and then reconstruct this as effectively as possible as in

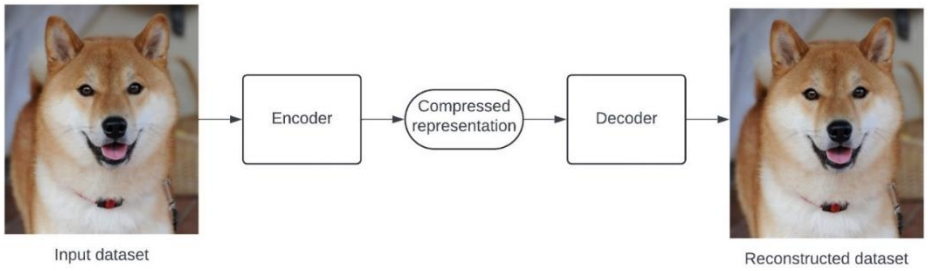


Figure 5.

Practically autoencoders are effective in data compression but are less effective at generalising to unseen data so have largely been displaced by VAEs(332).

Figure 5: Basic Autoencoder structure

13.10.3 Variational Autoencoder Structure

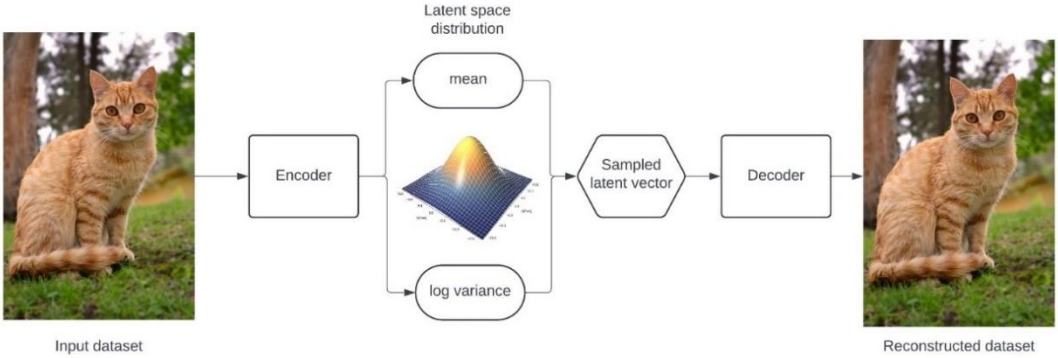
To recap the introduction, VAEs augment the classical autoencoder approach by compressing data to parameters of a statistical distribution, a mean and variance instead of compressing to fixed code layers. These variables are then used to reconstruct each element of the distribution back to the original input(332). The process of encoding reduces the dataset to a mean and log variance within the latent space, the model then randomly samples a point from the latent normal distribution and a decoder which maps this point back to the initial representation, as in Figure 6(332).

Figure 6: Variational autoencoder structure

13.10.4 Activation function selection

Neural networks and by extension VAEs consist of multiple layers of neurons beginning with an input layer and progressing to an output layer. Each individual node has a weight which is applied during data transformation. Activation functions allow the network to select which nodes will be activated at a given time. Failure to incorporate this would result in a purely linear function akin to a linear regression model(333). Activation functions apply non-linear transformations, and many exist within the field. Two of the most common are rectified linear unit (ReLU) and LeakyReLU(333).

The ReLU function is widely used and deactivates the neuron only when the output of the linear transformation is zero. ReLU has shown superior performance over the sigmoid and other functions and has been widely applied in deep learning(333–336). However, it has a key issue termed the “dead neuron problem”. As all negative values in ReLU are set to zero, specific neurons can become persistently inactive during model training(333,336). This results in persistent zero outputs in response to the weighted sum of inputs to this neuron. This causes significant reductions in the gradients of the loss function which can impair model learning and arrest the training process.



An approach to address this problem is a modification of ReLU termed LeakyReLU which incorporates a small nonzero gradient for negative input values through ReLU. This approach retains the computational simplicity of ReLU whilst preventing neuron death during the training process of the model(333,336). Given the composite datasets of sparsely populated binary and continuous variables in this thesis, LeakyReLU was preferred as the activation function of choice.

13.10.5 Optimiser selection

Within VAE models the optimiser algorithm adjusts the weight of the model in order to minimise reconstruction and regularisation loss. Optimisers work using backpropagation as described previously to adjust these weights. Through model training this process is iterated until a desired loss level is reached, a requisite number of computational cycles (Epochs) is completed, or overfitting and other complications occur.

Within the application of VAEs multiple optimisers are in use, with most implementing a variant of stochastic gradient descent(337). Adaptive moment estimation (Adam) is an example of a commonly used optimiser in VAE applications due to its rapid convergence and adaptive learning rates(338). Adam is particularly effective when applied to large datasets with high dimensionality spaces and has been shown to deal effectively with noisy and sparse gradients(338,339).

Adam achieves this efficacy by the combination of introducing an adaptive learning rate for each parameter and combining this with momentum methods which aid the optimiser to accelerate the learning process by considering previous gradients. Momentum is a method which can counteract local minima within a global training curve whereby simple stochastic gradient descent would become stuck. This is represented in Figure 7 and is overcome through the incorporation of momentum which allows the optimiser to update based on current gradient and previous parameter updates(337).

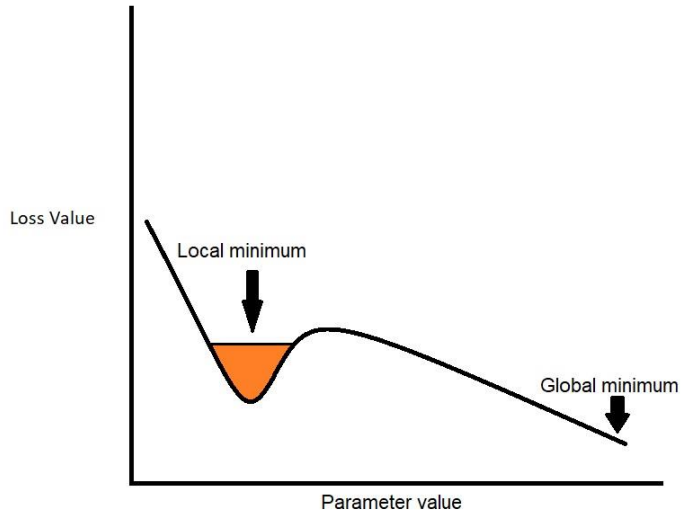


Figure 7: Local and global minima. Momentum based methods overcome these by incorporating previous gradients to each update and ensuring the optimiser retains inertia, allowing it to train past local minima in contrast to methods which rely totally on stochastic gradient descent.

As mentioned, Adam incorporates momentum and adaptive learning rates making it commonly the optimiser of choice for VAEs. As a result, it has been applied extensively to multiple tasks within the field of medical machine learning including deep learning based asthma attack detection and retinal disease photo classification among others(340,341).

13.10.6 Loss function selection

VAEs are trained using two loss functions, reconstruction loss which measures the fidelity of the decoded data to the original input and regularisation loss which measures the ability of the encoder to generate a regularised latent space which is as close as possible to normal distribution(332).

13.10.6.1 Reconstruction loss functions

Within VAE construction two common techniques for quantifying reconstruction loss are binary crossentropy(BCE) and mean squared error (MSE) calculations. Binary crossentropy is commonly utilised in autoencoders with binary outputs and measures the difference between the probability distribution of the original dataset and the reconstructed dataset(342). In contrast MSE is more appropriate for continuous data and simply measures the average of the squared differences between the original and reconstructed datasets(343). Given the mixed datasets here comprising binary and continuous covariates with a goal to simply compress the data prior to subsequent cluster analysis, MSE was preferred over BCE in this application.

13.10.6.2 Regularisation loss functions

In the context of variational autoencoders regularisation loss, commonly Kullback Leibler (KL) divergence is used to quantify to what extent the predicted probability distribution from the autoencoder differs from the target probability distribution. A KL divergence values of zero indicates identical probability distributions. In summary the larger the KL divergence the greater the difference between the predicted and target property distributions.

13.11 CLUSTERING TECHNIQUES

Clustering techniques aim to divide datasets into classes with similar behaviours and features. An effective clustering process yields clusters with high intra-class similarity and low inter-class similarity(344). Multiple clustering techniques and their development have been discussed within the introduction. This section will focus on those applied within this thesis.

13.11.1 K-means clustering

K-means clustering is an unsupervised iterative clustering algorithm which partitions observations into k clusters where each observation is assigned to the cluster with the nearest mean(344–346). The process is as follows:

1. Selection of k points as centroids
2. Computation of squared Euclidian distance for each observation from centroids
3. Assignment of observations to their nearest centroid
4. Recompute of the centroid of each cluster until the centroids do not change

K-means clustering attempts to minimise the mean square error (MSE) for each k cluster through this process. This approach is termed Expectation-Maximisation (EM) and is repeated until model convergence.

Multiple algorithms exist to perform K-means clustering, within baseR the default is currently the method described by Hartigan and Wong(347,348). A key weakness of K-means clustering is its inability to define non-circular (Or hyperspherical) clusters. Moreover K-means has no facility for dealing with uncertainty when one datapoint is close to multiple centroids and is affected by the algorithm's initial guess of centroids(349).

13.11.2 Partitioning Around Medoids (PAM)

Partitioning around medoids is another clustering algorithm that differs from K-means in that it selects an actual data point within each cluster as the cluster centre rather than simply a defined point. PAM maps a distance matrix into a specified k. The algorithm acts to

minimise the average dissimilarity of objects to their closest selected object by identifying these centrally located medoids within each cluster(350).

PAM has several advantages over K-means, it is considered less sensitive to outliers as it can utilise non-Euclidian distance matrices which reduce their influence(350). It is however more computationally expensive in larger datasets as it evaluates every single data point as a potential medoid when iterating. Therefore, it must compute distances between every single data point and every single medoid within a dataset. This can limit its utility in large datasets(350).

13.11.3 Gaussian mixture modelling (GMM)

Gaussian Mixture Modelling (GMM) is a probabilistic unsupervised clustering technique. The key assumption being that a dataset in totality can be divided into multiple Gaussian distributions, each representing a distinct group of observations(349,351). This approach allows for the generation of ellipsoidal clusters modelled in a Gaussian distribution.

GMM employs Expectation-Maximisation to estimate the means, covariances (Σk) and mixing coefficients. This is calculated in a 2-step fashion:

Expectation (E) step:

The model calculates the posterior probabilities (Responsibilities) that each observation belongs to a specific Gaussian distribution.

Maximisation (M) step: Using the calculated responsibilities as weights for each observation, the means, covariances and mixing coefficients are updated to maximise the expected log-likelihood of the data

These processes are repeated in an iterative manner until model convergence. Chief advantages of GMM over K-means are the ability to define ellipsoid clusters, and probabilistic assignment. These come with the caveat of greater computational cost and complexity relative to K-means

13.11.4 Spectral and hierarchical clustering

Spectral clustering and hierarchical clustering are only briefly applied in this thesis using the M3C R package and do not feature in the results output(352). Therefore, these are not covered in depth within this section but have been discussed within the introduction.

13.11.5 Comparison of clustering efficacy

In contrast to hierarchical techniques K-means clustering necessitates a priori determination of cluster number (k). Multiple methods exist to estimate this including Silhouette analysis, Monti consensus clustering, Proportion of Ambiguous Clustering (PAC) score calculation and the application of the M3C algorithm to generate a Relative Cluster Stability Index (RCSI).

13.11.5.1 Silhouette analysis

Silhouette analysis and Average Silhouette Width (ASW) are a commonly used cluster validation index to aid determination of the optimum k value for a dataset. They were initially developed by Rousseauw in 1987(353) . Silhouette analysis is a used to help determine the optimal k value for unsupervised clustering techniques. It provides quantitative assessment of how effectively assigned a datapoint is to its designated cluster(354). This metric ranges from -1 to +1 where +1 represents perfect integration of that data point to the assigned cluster and -1 suggests the observation should be assigned to a neighbouring cluster. This is termed the Silhouette coefficient and is calculated as follows:

$$\boldsymbol{S(i) = b(i) - a(i)/(max\{a(i), b(i)\})}$$

Where:

- ***S(i)*** is the silhouette coefficient for the data point i
- ***a(i)*** is the inter-cluster distance.
- ***b(i)*** is the inter-cluster distance(354).

The averages silhouette coefficient for different k values can be compared alongside heuristic/clinical evaluation of cluster significance to help determine the optimum k value. Weaknesses of the silhouette score method are its sensitivity to cluster outliers which can drastically affect the average silhouette score. It can also struggle with high dimensionality datasets as distances between datapoints become more uniform. Finally, irregularly shaped non-convex clusters may lead to poor silhouette scores even in well-defined clusters(355).

13.11.5.2 Monti consensus clustering

Monti consensus clustering was developed in 2003 and presents a resampling based method of k determination(356). Consensus clustering is a resampling technique that repeatedly performs clustering on different subsets of the dataset using different initialisation parameters to produce a consensus matrix summarising the co-occurrence of observations within different structures across all the relevant iterations(357). This technique was

pioneered in microarray gene expression data and has seen extensive use within gene expression analyses.

Weaknesses of the Monti consensus clustering was its difficulty in dividing unimodal data into stable k clusters and in datasets with a known k value it was shown to be poor at identifying the true data structure by Şenbabaoğlu et al(252) .

13.11.5.3 Proportion of Ambiguous Clustering (PAC) score

The Proportion of Ambiguous Clustering (PAC) score was developed by Şenbabaoğlu et al to address limitations within the Monti consensus clustering approach(252). PAC scores progress consensus clustering further by quantifying the proportion of observations within a dataset that are ambiguously clustered, defined as observations that appear in different clusters in different iterations. This allows for further assessment of cluster durability in addition to consensus clustering approach and silhouette analysis.

Although PAC scoring represents an improvement over pure consensus clustering, limitations remain. PAC scores do not consider null reference distributions and do not test the null hypothesis k=1(352). Additionally, there is no formal hypothesis testing incorporated into consensus clustering and PAC scoring(352). Moreover, when using the PAC metric there is a tendency toward overfitting and high k values

13.11.5.4 Monte Carlo Reference-Based Clustering (M3C) and the Relative Cluster Stability Index (RCSI)

Monte Carlo reference-based consensus clustering (M3C) was developed to address some of the limitations associated with consensus clustering and PAC score utilisation(352). M3C calculates null distributions of PAC scores for each k through Monte Carlo based simulation of k=1 null datasets. This permits formalised testing of the null hypothesis that the PAC score comes from a single Gaussian cluster against the alternative hypothesis that the PAC score does not come from a single Gaussian cluster(352). If no p values are significant, we accept the null hypothesis, whereas if p values are significant the null hypothesis can be rejected and the alternate hypothesis accepted.

Further advantages of M3C over PAC and consensus clustering are its resistance to noisy datasets and incorporation of Monte Carlo simulations to add additional statistical robustness to the process. In their original publication the authors were able to apply M3C to both simulated and real gene expression datasets from the Cancer Genome Analysis set, demonstrating superior performance to both PAC and Monti consensus clustering approaches(352,358).

Having determined a suitable k value, the Relative Cluster Stability Index (RCSI) can be calculated using the M3C package to assess the robustness of these clusters to perturbations in the dataset. These include the addition of noise, resampling and feature adjustment. A higher RCSI value represents a more durable clustering structure to these perturbations(352)

13.1 CAUSAL INFERENCE APPROACHES TO PREDICTIVE MODELLING

Though not a primary focus of this thesis, formalised causal inference approaches to questions of predictive modelling are important to consider when assessing the potential prognostic utility of gene expression scores in Results 4. Within this section Directed Acyclic Graph (DAG) analysis was carried out using the Dagitty package in R (359). DAGs are casual diagrams that allow researchers to document their knowledge and assumptions about a causal pathway. This allows for identification of the estimand (causal effect of interest) and construction of the correct estimator (Model or approach that estimates the estimand)(360–362).

DAGs are able to depict graphically causal relationships between variables (nodes). They allow the identification of potential sources of bias within a causal pathway such as collider bias (When models condition on a node with two parent variables) and confounding bias (Where a node has a causal effect on both the exposure and outcome). DAG construction allows for identification of mediators that sit between an exposure and outcome in a causal pathway. This allows for subsequent mediation analysis where the effect of an exposure on an outcome can be decomposed into the direct effect and the indirect effect via the mediator. Many packages exist to perform this calculation including MMA and mediation in R(363,364).

The use of DAGs opens our a priori assumptions to intellectual scrutiny as well as avoiding multicollinearity in predictive models. Within this thesis DAG analysis is applied during the construction of predictive modelling based around gene expression scores to limit these effects on the output.

13.2 LINEAR MIXED EFFECTS MODELLING

Linear mixed effects models are employed within Results 4 to investigate associations between causal inference informed covariates and gene expression scores in ANA-RMDs. Mixed effects modelling can account for variability within identified covariates (Fixed effects) as well as variability between individual cases, such as between individual patients in longitudinal follow up studies such as DEFINITION Cohort C (Random effects). Fixed effects

should be conserved across experiments as they are fixed at a population level(365). Additionally residual variance, i.e. the variance not sufficiently explained by the random and fixed effects can be estimated(365). Mixed effects modelling is adept at handling missing data in comparison with other methods such as ANOVA which employ listwise deletion of NA values. Moreover, continuous and categorical covariates and outcomes can be handled effectively within mixed effects models(365). Estimation of variance components in mixed effects models can be aided by calculation of Akaike or Bayesian Information Criteria (AIC and BIC). However, with longitudinal data the restricted maximum likelihood ratio (REML) is frequently preferred(366).

13.3 Multivariable Logistic Regression Modelling

Basic multivariable modelling was utilised in Results 4 alongside the causal inference informed approaches previously discussed. Multivariable linear regression modelling is able to assess the impact of multiple independent variables on a dependent variable. Dependent variables can be continuous (Linear regression) or binary (Logistic regression). Multivariable analysis is extensively utilised within observational research, but several key considerations can significantly impact the interpretability and robustness of outputs from multivariable logistic regression models(367). Covariate selection is dependent on both overall sample size and outcome event frequency. Multicollinearity and the presence of unobserved potential confounders must also be considered along with the potential for overfitting which can limit the generalisability of model outputs.

13.4 Data Manipulation

13.4.1 Imputation of missing variables

Data missingness presents an issue for the implementation of traditional clustering techniques such as K-means clustering. Compression with VAEs can help to deal with data missingness however uncompressed data, or datasets with high levels of missingness may require a degree of imputation. Within this work we utilised multiple imputation with chained equations (MICE). This technique involves a multiple stage process outlined below(368):

1. Initial imputation of missing variables with mean, median or random variables
2. Imputed variables are removed sequentially with models created using the other complete samples.
3. Missingness is imputed into each variable in turn using the created model and the remaining variables This is repeated until the process is completed.

4. The new completed data is subtracted from the initial imputed variables to create a difference matrix which forms the basis of the next iteration
5. Sequential iterations are carried out until the initial dataset and completed datasets reach a pre-specified difference threshold.

Within R this process can be completed using the MICE package, which was utilised within this thesis(369).

13.5 DIFFICULTIES ASSOCIATED WITH DATA ANALYSIS

13.5.1 Free text fields

Analysis of the DEFINITION database free text fields is problematic with multiple terms used for single disease manifestations; For example, arthritis, synovitis, MCP inflammation etc being used to denote inflammatory arthritis. This was managed with the use of text parsing functions in R which collated these terms into single fields for analysis. Time-consuming manual review of the free text fields was required to define some features within DEFINITION, though not for the PRECISESADS cohort. As a result, fields such as the presence of arthritis were collated from any mention of it within the free text fields or any presence of arthritis specific items within the disease activity or classification fields.

13.5.2 Comparative analysis of ΔCT and RNA-Seq gene expression data

Comparative analysis of the PRECISESADS and DEFINITION cohorts required normalisation of the RNA-Seq data for comparison with reflected ΔCT values from DEFINITION. To overcome this, we utilised normalisation techniques from the Bioconductor resource in R (Limma-Voom, DNASeq2- VST and edgeR- counts per million (CPM)). Voom proved to be the most effective package, and this data was taken forward for analysis.

13.5.3 Comparative analysis of parallel datasets

There was considerable scope for difficulty in analysing two parallel cohorts based in different academic centres. Prior to the commencement of this thesis, we were able to compare and contrast CRFs from both datasets and identify variables to be taken forward for model construction and subsequent analysis. Consequently, there were no significant issues in this regard within the project.

13.5.4 COVID-19 restrictions and the impact on recruitment to DEFINITION

Covid-19 restrictions took place during the second half of recruitment to DEFINITION. During this period, no patients attended hospital for research at all. DEFINITION was embedded in routine clinical practice, so if patients had to attend hospital for routine care, then data may

have been collected. However, in these instances some procedures could not be completed, and no fresh flow cytometry was performed. We were able to reschedule some visits outside the period of COVID-19 restrictions, and to perform partial data collection for other visits. Impact of these changes was reported as an overall summary rather than individual protocol deviations in line with local guidelines in place during the pandemic.

13.5.5 Data stewardship

Study data conformed to General Data Protection Regulation (GDPR) guidelines and was under the oversight of data protection officers both within the university of Leeds and the Centre for Genomics and Oncology Research (GENYO). Formalised agreements were drafted with GENYO prior to data sharing. PRECISESADS and DEFINITION data was anonymised and shared via secure methods. Data is stored in R files and CSV files on secure university servers in an anonymised format. PRECISESADS contains approximately 2500 samples and extensive genomic data with a file size of 11GB. DEFINITION comprises a far smaller dataset of 294 patients. The PRECISESADS group at GENYO and the University of Granada retain access to the PRECISESADS dataset. This is overseen by Dr Guillermo Barturen and Professor Alarcon. The DEFINITION data is stored on University of Leeds servers and is overseen by Professor Edward Vital.

Data collection for both studies utilised case report forms reviewed and approved by local ethics and data stewardship committees. Electronic data was stored on secure servers and managed with pseudo anonymisation techniques to reduce the risk of participant identification. Informed consent and ethical review as discussed earlier in this section were implemented in both cohorts.

14 RESULTS 1: PRELIMINARY PURE FLOW CYTOMETRY BASED RECLASSIFICATION OF ANA-RMDS USING ESTABLISHED CLUSTERING TECHNIQUES

14.1 INTRODUCTION

The heterogeneity within the connective tissue disease patient cohort is not adequately reflected by current diagnostic labels(139,154). Identification of similar patient phenotypes across diagnostic groups may permit further investigation and basket therapeutic trials. The DEFINITION cohort contains extensive data on multiple different connective tissue diseases suitable for this approach.

Flow cytometry analysis of autoimmune rheumatic disease is an established technique(370). Within single disease cohorts it has been used extensively and previous work in Leeds has demonstrated the value of plasmablast enumeration as a biomarker, especially after B cell depletion and has also identified B-cell tetherin, a lipid raft protein as a marker of IFN-I expression in SLE patients(327). The advantage of tetherin over IFN-I gene expression scores is the ability to quantify IFN-I response on individual cells or cell populations. This is particularly important for B cells since B cell response to IFN-I is a critical pathway of SLE pathogenesis, but whole blood IFN-I biomarkers are dominated by monocyte ISG expression. Memory B cell tetherin was shown to correlate better with clinical features, as well as plasmablast repopulation after rituximab, as compared with monocyte tetherin or overall blood ISG expression. The IFN-I blocking medication Anifrolumab has demonstrated positive trial results in SLE and is in trials in SSc. However, given IFN-I pathway activation appears to be a conserved mechanism across ANA-RMDs it may be beneficial to characterise B-cell tetherin expression across ANA-RMDs and incorporate this within flow cytometric analysis when aiming to reclassify ANA-RMDs. Within the single disease paradigm of SLE lymphocyte subset analysis using flow cytometry on 143 patients and hierarchical clustering was able to identify 4 clusters with differential disease activity(371). However, to date there has been limited application of these techniques to multi-disease cohorts. Characterisation of the B-cell populations of these patients may be particularly useful given the availability of targeted therapies such as Belimumab, Rituximab and Obinutuzumab.

Traditional approaches to disease reclassification within rheumatology have centred on single modality clustering using single data types such as flow cytometric or gene expression data. Moreover, approaches utilised within these analyses have generally been simple approaches such as K means, hierarchical or partitioning around medoids based approaches. The core aim of this thesis is the multimodal reclassification approach presented in Results Chapter 3. However, given the importance of IFN-I pathway activation in reclassification, and the availability of flow cytometry data, I first performed a more limited analysis on these data.

Therefore, as part of introductory work in this thesis and to achieve basic literacy with R and basic clustering, I set out to apply similar approaches to a subset of the DEFINITION cohort and assess whether these approaches were able to define satisfactory clusters in isolation.

14.2 OBJECTIVES

To achieve familiarity with data manipulation and preliminary clustering analysis in R, and to perform a focused analysis on flow cytometric biomarkers, by applying PAM clustering to flow cytometry data from a mixed cohort of connective tissue disease patients.

14.3 METHODS

142 patients with extant flow cytometry data available were selected from the DEFINITION cohort and taken forward for analysis as part of this focused preliminary study.

Highly sensitive flow cytometry was used to identify PBMC subsets (described in detail below). ANA and ENA titres were assessed via the Bioplex assay. Demographic and clinical data was obtained from the DEFINITION cohort. Clustering utilised flow cytometry data and the partitioning around the medoid package in R. The flow cytometry panel was designed to enumerate major circulating cell subsets using conventional definitions. SIGLEC-1 (on monocytes) and Tetherin (on all subsets) were also included as cell surface markers of IFN-I pathway activation. Output clusters were retrospectively compared to their legacy diagnoses and serology.

Statistical analysis was carried out in R, categorical variables using a chi square test and continuous variables using ANOVA using base statistics packages.

Clustering analysis was carried out using a partitioning around medoids approach on the cell types in Table 8, expressed as percentages of total peripheral blood mononuclear cells

(PBMCs) as well as the MFI of SIGLEC-1 and Tetherin. This analysis was carried out using the cluster (v2.16) package in R(372)

14.3.1 Flow cytometry

In order to effectively isolate NK, T and B-cell subsets from peripheral blood mononuclear cells (PBMCs) a multiple gating strategy was employed as per previous work in the Vital group. This gating strategy is for NK, T and B-cell subsets is summarised as follows:

- PBMCs were gated by size forward scatter and side scatter with doublet cells eliminated.
- CD3 and CD56 were gated from single PBMCs
- CD19 positive cells were gated from single PBMCs, which were used to gate CD27 negative naïve and CD27 positive memory B cells
- Plasmablasts were gated as CD19 intermediate / high, CD27 positive, CD38 bright mononuclear cells with exclusion markers for CD3 and CD14 as well as the CD19+ naïve and memory B cells.

Isolation of monocyte subsets was as follows:

- Monocytes were gated by size forward scatter and side scatter with elimination of doublet cells.
- Monocyte single cells gated
- CD14 positive monocytes gated
- Non-classical monocytes gated (CD14+ CD16+)
- Classical monocytes gated (CD14++ CD16-)
- Intermediate monocytes gated (CD14++ CD16+)

Conventional methods were used for the definition of CD3+ T cells, NK cells, and NKT cells.

SIGLEC-1 expression was quantified as mean fluorescence intensity (MFI) on total monocytes. Tetherin expression was quantified as MFI on each cell subset described below. In keeping with previous work, Tetherin on memory B cells and classical monocytes were included in the analysis.

This flow cytometric gating and analysis was carried out by Ms Zoe Wigston prior to the start of this doctorate. The cell subsets utilised for this clustering are summarised in Table 8.

Cell population	Cell Surface Marker
NK bright cells	CD3- CD56++
NK dim cells	CD3- CD56 intermediate
NKT cells	CD3+ CD56+
T-cells	CD3+
B-cells	CD19+
Naïve B cells	CD19+ CD27-
Memory B-cells	CD19+ CD27+
Plasmablasts	CD19+/- CD27+ CD38++ CD3-CD14-
Monocytes	CD14+
Non Classical Monocytes	CD14+ CD16+
Intermediate Monocytes	CD14++ CD16+
Classical Monocytes	CD14++ CD16-)
SIGLEC-1 on Classical monocytes	CD169 MFI on CD14++CD16-
Tetherin on memory B cells	CD317 MFI on CD19+ CD27+ cells
Tetherin on classical monocytes	CD317 MFI CD14+ CD16+

Table 8: Cell subsets used in PAM clustering

14.4 RESULTS

Within 142 patients with extant flow cytometry data within DEFINITION, 3 clusters were identified with memberships of 45, 49 and 48 patients respectively. An elbow plot is shown in Figure 8. A k value of 3 was chosen as a compromise between generating clusters with sufficient membership to be clinically useful and effectively subdividing the small sample size.

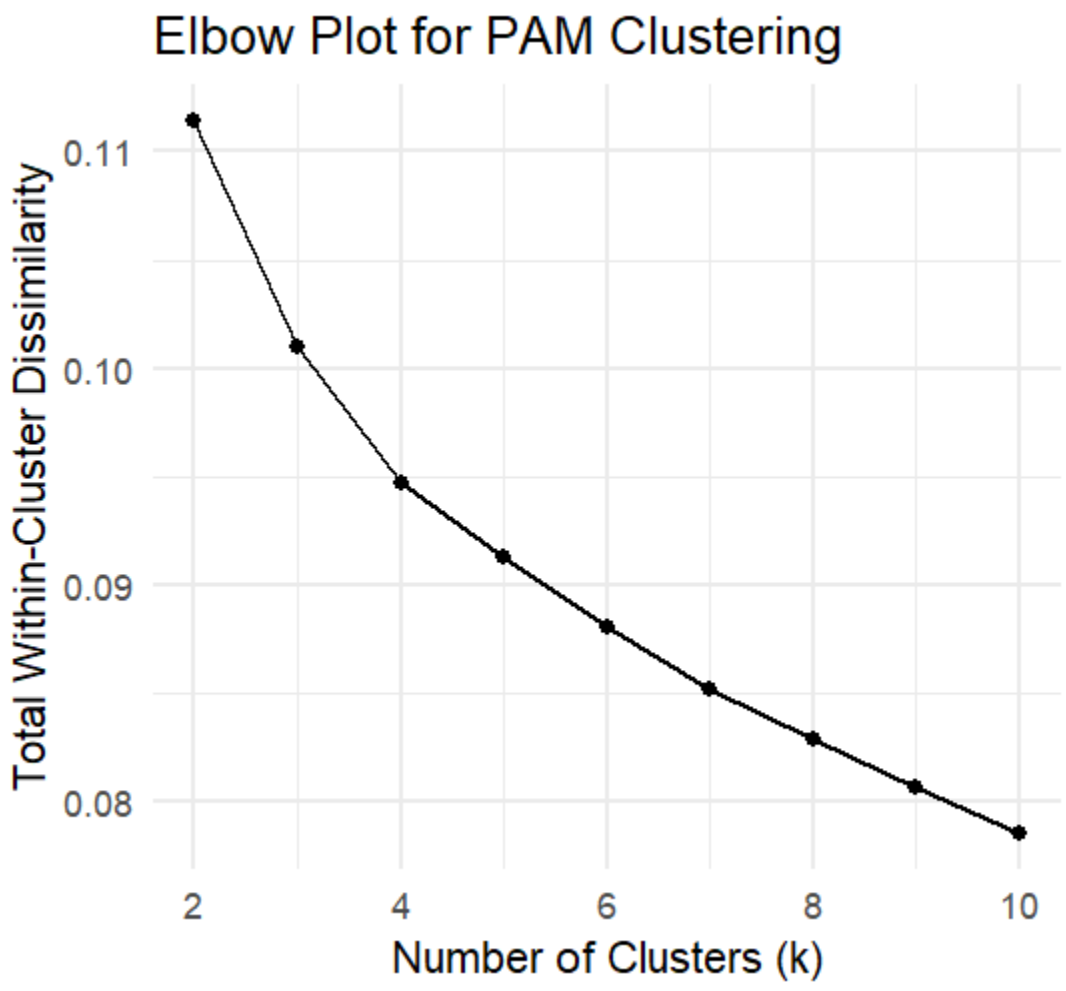


Figure 8: Elbow plot of Total Within Cluster Dissimilarity vs (k) clusters

These clusters are visualised using t-SNE in Figure 9. These are reasonably separated but with significant overlap.

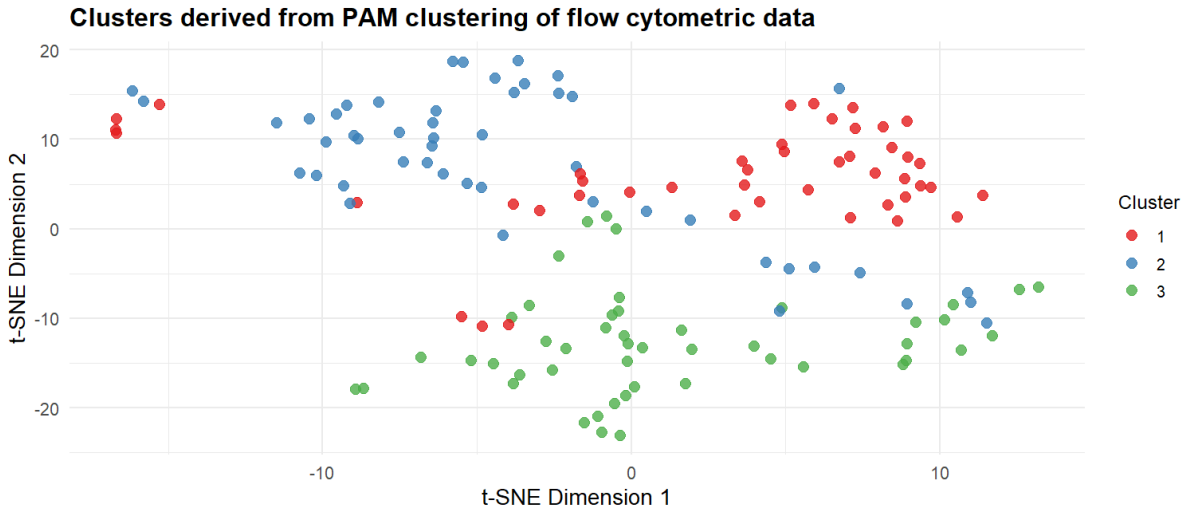


Figure 9: t-SNE plot of flow cytometry derived clusters

Cluster 1 (**UCTD / T cell dominant**) was characterised by a higher proportion of UCTD patients (27%) and larger percentages of CD3 positive T cells and NK bright cells.

Cluster 2 (**SSc dominant / Tetherin deplete**) was characterised by a higher population of systemic sclerosis patients (10%) and a lower proportion of SLE patients (20%). High percentages of classical and non-classical monocytes were demonstrated alongside low plasmablast percentages and low SIGLEC-1 / memory B cell and monocyte Tetherin expression.

Cluster 3 (**SLE / pSS, Ro / Tetherin dominant**) included a higher proportion of SLE and primary Sjogren's syndrome patients (52% and 19%). This cluster showed larger plasmablast percentages and higher SIGLEC-1 / Tetherin expression. Serologically this cluster had higher levels of Ro60, Ro52, Sm-RNP and La positivity.

Statistical analysis revealed significant differences in Ro60 ($p<0.01$), Ro52 ($p < 0.05$), La ($p<0.01$), and Sm ($p<0.05$) antibody positivity between the 3 clusters. Statistically significant differences in CD19+ CD27+ memory B cell and CD14+ CD16- classical monocyte SIGLEC expression were also observed (both $p < 0.05$). A significant difference in SLE distribution was noted between the clusters ($p < 0.01$).

Diagnosis	Total (n=142)	Cluster 1 (n=45) n/N (%)	Cluster 2 (n=49) n/N (%)	Cluster 3 (n=48) n/N (%)	P value
SLE	57	22/45 (49))	10/49 (20)	25/48 (52)	<0.01
UCTD	25	12/45 (27)	8/49 (16)	5/48 (10)	0.06
Sjogren's	15	2 /45 (4)	4/49 (8)	9/48 (19)	0.06
Systemic sclerosis	7	0 /45 (0)	5/49 (10)	2/48 (4)	0.07
MCTD	9	2/45 (4)	4/49 (8)	3/48 (6)	0.76
Myositis	8	3/45 (7)	3/49 (6)	2/48 (4)	0.86
Antibody positivity					
dsDNA	43	16/45 (36)	11/49 (22)	16/48 (33)	0.33
Ro60	45	12/45 (27)	7/49 (14)	26/48 (54)	<0.01
Ro52	36	10/45 (22)	7/49 (14)	19/48 (40)	0.01
La	18	2/45 (4)	2 /49 (4)	14/48 (29)	<0.01
Sm	14	1/45 (2)	4/49 (8)	9/48 (19)	0.02
Sm_RNP	23	7/45 (16)	5/49 (10)	11/48 (23)	0.23
RNP	14	3/45 (7)	5/49 (10)	6/48 (13)	0.64
SCL-70	1	0/45 (0)	0/49 (0)	1/48 (2)	0.37
Jo-1	4	1/45 (2)	2/49 (4)	1/48 (2)	0.80
Centromere	5	1/45 (2)	4/49 (8)	0/48 (0)	0.08
Chromatin	29	10/45 (22)	8/49 (16)	11/48 (23)	0.68
Ribosomal P	4	1/45 (2)	1/49 (2)	2/48 (4)	0.78

Table 9: Key legacy diagnoses and ENA positivity for the flow cytometry identified clusters

14.5 DISCUSSION

The chief aims of this work were (i) to rapidly develop skills in basic data manipulation and clustering within R using a subsection of the DEFNITION cohort and (ii) to perform a focussed analysis on the flow cytometry data available prior to the more complex analysis in Results Chapters 2 and 3. This preliminary study identified 3 flow cytometry clusters from the DEFINITION cohort. There was overlap between these clusters, indicating flow cytometry and PAM alone are not sufficient to fully characterise patients. The numbers of patients were also relatively limited for this method. Larger sample sizes would be required to analyse these clusters in more detail.

Cluster 1 contained many UCTD patients. These patients are known to often have borderline evidence of clinically significant inflammation. The “high” percentage of T and NK cell numbers in this cluster may therefore actually indicate relatively normal flow cytometry, since these cell populations are known to be low in SLE. These patients had high proportions of dsDNA and intermediate Ro and La positivity.

Within Cluster 2 the high proportion of systemic sclerosis patients and classical /nonclassical monocyte population was in keeping with known characteristics of these patients(373). Not all systemic sclerosis patients were segregated into this class. In the literature non-classical monocytes have been associated with an aberrant IFN-I in this patient cohort(374). This may suggest that there are both IFN-I driven and IFN-I independent disease processes driving pathology within systemic sclerosis patients. Generally, ENA positivity was proportionally less than the other clusters, with the notable exception of anti-centromere positivity.

Finally, Cluster 3 demonstrated higher memory B cell populations which along with plasmablasts have been associated with higher disease activity in SLE patients(375). This was mirrored by the high levels of Sm, SmRNP and RNP positivity which have been associated with high disease activity in SLE patients(376).

Importantly the flow cytometric analysis carried out as part of this analysis is generally only available in tertiary centres and would not be readily accessible in peripheral or resource poor environments. Therefore, classifications based around this approach will invariably have limited global utility. Moreover, this small-scale approach was insufficient to identify any meaningful clusters for further analysis and occurred prior to the processing and data-cleaning of the clinical variables which would have provided much needed clinical context to

the outputs. However, it succeeds in the sense that it provided a working approach and skillset which can be refined later within this doctorate. Moreover, it demonstrated multiple shortcomings associated with reductive approaches such as this and the additional utility of more complex approaches when approaching the central problem of reclassification across ANA-RMDs

15 RESULTS 2: UNSUPERVISED MACHINE LEARNING BASED RECLASSIFICATION OF ANA-RMDS. DEFINING ANA-ARTHRITIS, A CLINICAL AND BIOMARKER DEFINED POPULATION SUITABLE FOR CLINICAL TRIALS.

15.1 INTRODUCTION

ANA-associated RMDs (ANA-RMDs) are a spectrum of overlapping diseases characterized by autoreactivity to nuclear antigens encompassing systemic lupus erythematosus (SLE), primary Sjogren's Syndrome (pSS), idiopathic inflammatory myopathies (IIM) and systemic sclerosis (SSc). Many patients have overlap syndromes, meeting classification criteria for multiple diseases, or undifferentiated forms of ANA-RMDs (UCTD) that are not easily classified(183,185,377). Despite distinct clinicopathological manifestations, such as specific antibodies, or skin fibrosis in SSc versus exocrine gland inflammation in pSS, ANA-RMDs also share features such as arthritis and immunopathogenic signatures(378).

Despite these shared features, treatment inequity exists between SLE and other ANA-RMDs. In SLE patients with arthritis, two targeted therapies are licensed while there are none for pSS patients with arthritis(379,380). Additionally, UCTD patients lack evidence-based treatment strategies and are ineligible for clinical trials. In single diseases such as SLE, diverse clinical and immunological presentations, such as cutaneous and musculoskeletal symptoms, pose challenges in defining trial populations, measuring outcomes, and assessing treatment effectiveness(381,382). Heterogeneity within SLE may partially explain how, despite encouraging clinical responses for certain disease manifestations, several studies failed to meet multisystem primary endpoints leading to programme discontinuation(379).

Reclassifying ANA-RMD patients into alternative "baskets" may address these issues, as exemplified by approaches to autoimmune disease-associated interstitial lung disease(271). Baskets may be defined as groups of patients from different legacy diagnoses into a new grouping that is suitable for a similar therapeutic intervention. Baskets may be based on shared pathogenic mechanisms (e.g. B-cells or Type-I interferon pathway activation) or a shared clinical problem, e.g. arthritis. Conducting clinical trials in a well-defined basket cohort could address unmet clinical needs and yield evidence-based interventions for

patients across a wider spectrum of diagnoses. Furthermore, basket population trials may bolster effect sizes by utilizing more homogeneous study populations than existing trials, which cover multiple organ manifestations, biomarker subgroups, and background therapeutics. ANA-RMD arthritis may be a suitable basket for this strategy as it is common and significantly impacts quality-of-life, functional disability, work impairment, and health-economic outcomes(111,383).

The study objectives were: (i) to evaluate the prevalence of “ANA-arthritis”, defined as synovitis, tenosynovitis, enthesitis or other articular/periarticular inflammation in patients with ANA-RMD, in a multi-disease study; (ii) to test the hypothesis that ANA-arthritis is associated with a similar clinical impact across legacy diagnoses; (iii) to define new basket groupings of patients across the ANA-RMD spectrum for clinical trials; (iv) to evaluate immunological profile, and therefore suitability for therapies, of legacy diagnoses and new therapeutic baskets.

15.2 METHODS

15.2.1 The DEFINITION cohort

Patient recruitment, variables collected and use in statistical analyses are summarized in Figure 10.

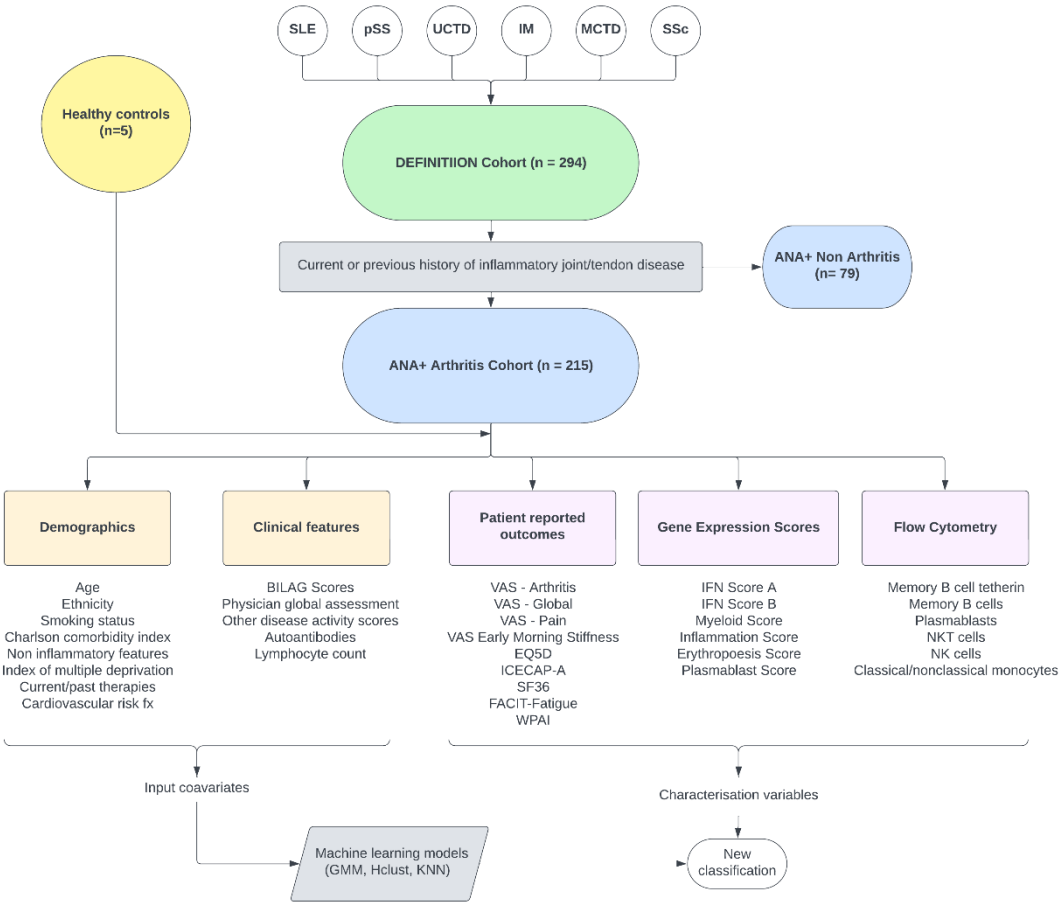


Figure 10: Study schematic

DEFINITION is a prospective, multi-disease ANA-RMD cohort study in Leeds, United Kingdom since May 2017. The primary aims were to better define the role of IFN-I and other biomarkers in ANA-RMD and refine the use of IFN-I targeted, conventional and other therapies. This analysis focused on patients with a history of inflammatory arthritis or currently active disease. Patients were identified through ultrasound documented synovitis at enrolment, the presence of inflammatory arthritis items on validated instruments (BILAG-2004 A-C articular/tendinopathy domains or ESSDAI), or any of the following terms in their medical documentation: arthropathy; arthritis; arthralgia; synovitis; tenosynovitis; joint tenderness; epicondylitis; polyarthralgia. All individuals provided informed written consent, and this research was carried out in compliance with the Declaration of Helsinki. This study

was approved by the National Health Service Health Research Authority (REC Ref: 17/YH/0166). Healthy control participants' peripheral blood was collected under the study number 04/Q1206/107. All experiments were performed in accordance with relevant guidelines and regulations. University of Leeds was contracted with administrative sponsorship.

15.2.2 Demographics and comorbidity

We collected baseline age, gender, patient-identified ancestry, smoking status, index of multiple deprivation (IMD)(384), Charlson Comorbidity index (CCI)(385), clinical features of fibromyalgia, hypermobility and osteoarthritis. Diagnoses were recorded according to consultant physician review. This was preferred over diagnostic criteria which are absent in UCTD and often unfulfilled in pSS without tissue biopsy.

15.2.3 Laboratory measures

Full blood count, complement C3 and C4 levels and ANA subtypes, including anti-double-stranded DNA, Ro-52, Ro-60, La, Sm, SM/RNP, RNP, Scl-70, Jo-1, Centromere, Chromatin and Ribosomal-P antibodies (Bioplex multiplex analyser) were measured in a routine diagnostic laboratory. PBMC subsets were analysed using 8-colour flow cytometry as a proportion of total PBMC count (T-cells (CD3+CD56-), NK-cells (CD56+), NKT-cells (CD3+CD56+), Memory B-cells (CD19+CD27+, Plasmablasts (CD19+/-CD27+CD38++), classical monocytes (CD14++CD16-) intermediate monocytes (CD14++CD16+)and non-classical monocytes (CD14+CD16+)). Tetherin (CD317) mean fluorescence intensity was quantified on each cell subset, with memory B-cell level as the primary biomarker(168).

Two validated interferon-stimulated gene expression scores (IFN Score-A and IFN Score-B) were analysed. PBMC were separated using density gradient method (Lymphoprep; Alere-Technologies, Norway) from EDTA-anticoagulated blood. Total RNA purification kit (Norgen-Biotek, Canada) was used followed by quantitative real-time reverse transcriptase-PCR (qRT-PCR) using TaqMan assays (Applied Biosystems, Invitrogen) for the selected 30 ISGs as previously described(386,387). Scores for genes annotated to plasmablast, myeloid lineage, inflammation and erythropoiesis function were included from previously described modules based on their known molecular function(388). We used untransformed dCT gene expression scores to preserve a normal distribution. For untransformed values in the figures and tables, numerically lower dCT values represent higher gene expression.

15.2.4 Clinical assessment

Disease activity was assessed at baseline using validated instruments applied to all diagnostic groups: EULAR Sjögren's syndrome disease activity index (ESSDAI); British Isles

Lupus Assessment Group (BILAG) 2004 index; SLEDAI-2K. Rodnan skin score and MITAX were collected but not analysed due to limited relevance to clinical features and patient numbers . Physician global assessment (PGA) was also assessed. The validity of the articular component of the BILAG-MSK domain (excluding myositis) across non-SLE diagnoses was explored using association with PGA.

15.2.5 Patient-reported outcomes

Patient-reported disease impact was assessed using the following: 36-item Short Form Survey (SF36) – Composite and domain scores; Functional Assessment of Chronic Illness Therapy – Fatigue (FACIT-Fatigue); EuroQol-5 Dimension 5-level Score (EQ5D-5L) – Index and domain scores; ICECAP-A; Patient-reported visual analogue scales (VAS); Arthritis-VAS; Pain-VAS; Global-VAS; Fatigue-VAS; Global health-VAS and Early Morning Stiffness-VAS.

15.2.6 Machine learning

Model covariates were selected based on background evidence (MSK-BILAG Sm/SmRNP/RNP antibody status) and principle component analysis (PCA). PCA of 40 covariates, including Age, IMD-rank, prednisolone dosage, 15 ENA values, PGA, numeric MSK-BILAG, 8 gene-expression scores, 6 flow cytometry subsets, and 6 non-inflammatory features, identified 7 covariates for GMM. The primary variance was explained by IMD-rank, prednisolone dosage, numeric MSK-BILAG score, lymphocyte count, chromatin antibody positivity, Ro52/Ro60 antibody positivity, and Sm/SmRNP/RNP antibody positivity. Selected values explained >99.99% of data variance in the first 3 principal components within the 7-covariate model through singular value decomposition.

Multiple imputation with chained equations (MICE) was utilized to address missingness, with 3.75% (n=8) of the IMD-rank data and 8.92% (n=19) of the lymphocyte count values being imputed. Hierarchical clustering, k-means clustering, and Gaussian Mixture Model (GMM) were trialled using the hclust and base-R packages.

15.2.7 Statistical analysis

Statistical analysis and data visualisation utilised the heatmap, corrplot, ggplot and tableone packages in R version 4.1.2. Multiple group comparison employed Kruskal-Wallis testing, while twin group comparisons of categorical and continuous variables utilized Chi-square and T-tests, respectively. For correlation analyses, Spearman's rank correlation coefficient was used, considering correlations ≥0.3 or ≥-0.3 as substantive. Bonferroni correction was applied to compensate for multiple-hypothesis testing. Principle component analysis and Gaussian mixture modelling utilised the Mclust v6.0.0 packages. Data imputation used the MICE v3.15.0 package. Sankey plots were generated using SankeyMatic(389).

15.2.8 Patient and public involvement (PPI)

The NIHR Leeds Biomedical Research Centre PPI group have regular insight and input into planning and conduct of local ANA-RMD research. A workshop when designing the study identified arthritis as a key problem of interest

15.3 RESULTS

15.3.1 Prevalence of inflammatory joint and tendon disease in ANA-RMDs.

Of 294 patients with ANA-RMDs recruited to DEFINITION, 213 with inflammatory articular features were included. Key baseline features are detailed in Table 10: Baseline characteristics of DEFINITION cohort

The SSc and pSS groups had a higher median age and were more comorbid than other ANA-RMDs with a higher baseline CCI (p=0.021 and 0.033 respectively). SLE and myositis groups had higher proportions on long-term prednisolone therapy. MSK inflammation was common, and most prevalent in SLE and MCTD patients (87% and 77% respectively).

No significant differences were observed in physician-defined fibromyalgia features among diagnostic groups. No significant differences were found in the prevalence of nodular osteoarthritis (on clinical examination), x-ray confirmed osteoarthritis, or hypermobility syndrome.

15.3.2 Validity of MSK-BILAG across ANA-RMDs

To compare disease activity across ANA-RMDs, we explored the concurrent validity of articular scores within the MSK component of BILAG-2004 Index in Figure 11. This demonstrates face validity across all RMDs, relying on the presence of inflammatory pain or swelling to categorize arthritis/tenosynovitis severity. Its definition mirrors the articular MSK assessment in MITAX and ESSDAI, each encompassing mild, moderate, and severe grades for MSK inflammation. BILAG-2004 articular MSK grades A-D were significantly associated with PGA across both SLE ($F=14.43$, $p<0.001$) and non-SLE patients ($F=11.62$, $p<0.001$), supporting the use of this measure in classifying arthritis patients with various ANA-RMDs, pending further validation.

Variable	SLE	UCTD	pSS	MCTD	IM	SSc	p
Total, n	104	111	33	13	19	14	
MSK inflammation, n	90	77	23	10	6	7	
MSK inflammation, %	87%	69%	70%	77%	32%	50%	
Sex = M (%)	10 (11.1)	10 (13.0)	1 (4.3)	1 (10.0)	2 (33.3)	2 (28.6)	0.33
Age, mean (SD), years	46.20 (14.22)	50.17 (12.93)	55.96 (13.93)	48.60 (12.66)	50.17 (20.44)	59.29 (15.22)	0.021
Ancestry, n (%)							
Other / unknown	8 (8.9)	10 (13.0)	3 (13.0)	1 (10.0)	1 (16.7)	0 (0.0)	
Asian	14 (15.6)	8 (10.4)	2 (8.7)	1 (10.0)	0 (0.0)	1 (14.3)	
Mixed	4 (4.4)	2 (2.6)	0 (0.0)	0 (0.0)	0 (0.0)	1 (14.3)	
European	57 (63.3)	51 (66.0)	18 (78.3)	8 (80.0)	4 (66.7)	5 (71.4)	
African	7 (7.8)	6 (7.8)	0 (0.0)	0 (0.0)	1 (16.7)	0 (0.0)	
Current Smoker (%)	10 (11.1)	8 (10.4)	0 (0.0)	2 (20.0)	1 (16.7)	0 (0.0)	0.424
IMD Rank (mean (SD))	12998 (10263)	15767 (10511)	18573 (9456)	15709 (9346)	17691 (13968)	9890 (11320)	0.14
CCI Total (mean (SD))	2.00 (1.38)	2.04 (1.34)	2.91 (1.47)	2.10 (1.97)	1.67 (1.21)	3.29 (2.87)	0.033
FMS Pain/stiffness (%)	19 (21.1)	10 (13.0)	3 (13.0)	0 (0.0)	0 (0.0)	0 (0.0)	0.221
FMS Allodynia (%)	9 (10.0)	5 (6.5)	1 (4.3)	0 (0.0)	0 (0.0)	1 (14.3)	0.698
Hypermobility syndrome (%)	3 (3.3)	3 (3.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0.883

Variable	SLE	UCTD	pSS	MCTD	IM	SSc	p
Nodal OA (%)	6 (6.7)	4 (5.2)	3 (13.0)	1 (10.0)	1 (16.7)	2 (28.6)	0.257
X-ray proven OA (%)	11 (12.2)	13 (16.9)	4 (17.4)	2 (20.0)	1 (16.7)	2 (28.6)	0.851
Lymphocyte count x 10*9/L (mean (SD))	1.27 (0.59)	1.53 (0.73)	1.43 (0.65)	1.13 (0.64)	1.15 (0.70)	1.32 (0.31)	0.153
BILAG Numeric (mean (SD))	4.96 (6.01)	3.09 (4.45)	2.83 (2.81)	5.80 (5.73)	5.50 (7.18)	4.86 (4.38)	0.14
ESSDAI Total (mean (SD))	2.46 (3.54)	1.86 (3.58)	2.52 (3.49)	3.90 (5.36)	1.67 (2.34)	0.71 (1.50)	0.429
SLEDAI Total (mean (SD))	5.60 (4.10)	3.19 (1.82)	2.83 (2.15)	6.00 (4.97)	5.33 (4.13)	2.86 (1.57)	<0.001
PGA Q2 (mean (SD))	2.84 (2.17)	2.49 (1.88)	2.86 (1.81)	4.05 (2.36)	4.37 (3.87)	3.77 (1.91)	0.087

Table 10: Baseline characteristics of DEFINITION cohort

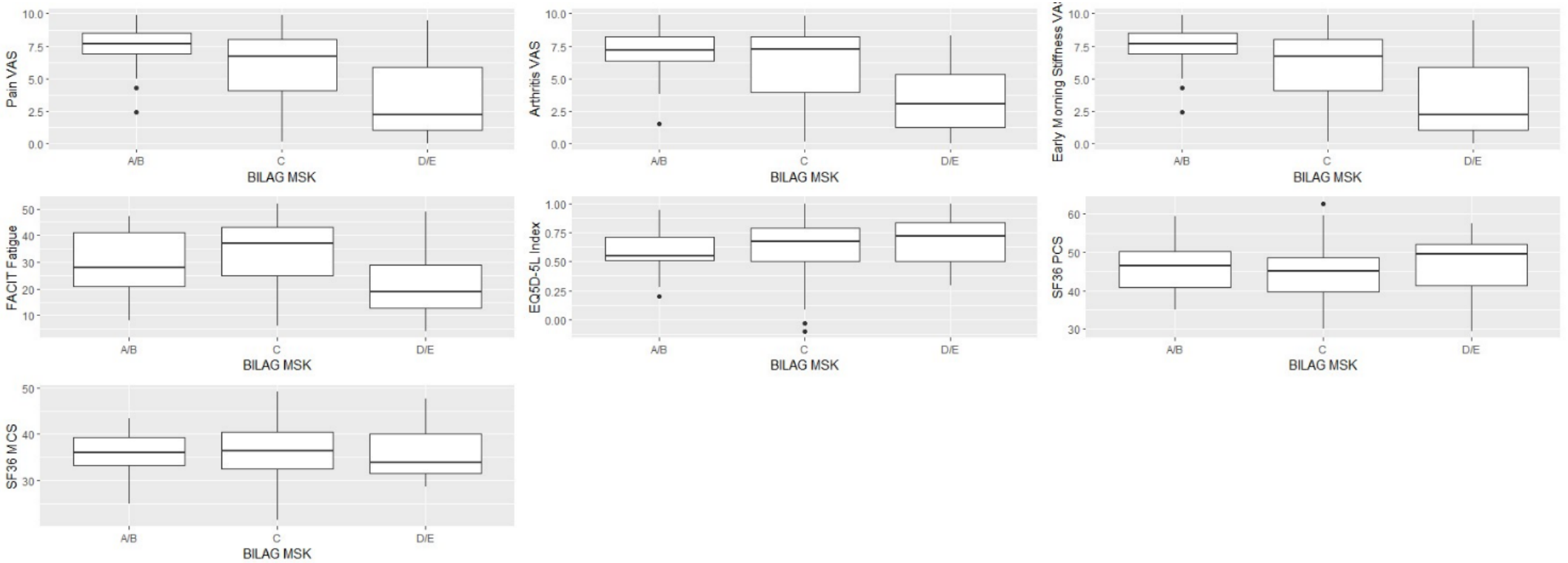


Figure 11: BILAG MSK correlation with patient reported outcomes

15.3.3 Clinical impact of joint and tendon inflammation in ANA-RMDs

To compare the clinical impact of articular symptoms across ANA-RMDs, we assessed physician-reported outcomes. Numeric BILAG-2004 values, ESSDAI total and physician global assessment did not differ significantly between diagnoses (Table 10). Overall disease activity, as per the BILAG score and individual domains, did not significantly differ across groups except for BILAG gastrointestinal domain activity, which was highest in SLE ($p>0.05$, Data not shown) SLEDAI-2K scores also differed significantly, being highest in the MCTD group ($p<0.001$, Table 10).

We compared patient-reported outcomes for symptoms (pain-VAS, EMS-VAS, arthritis-VAS, fatigue-VAS and global health-VAS) quality-of-life (SF36-MCS, SF36-PCS, EQ5D-5L), participation (ICECAP-A) and fatigue (FACIT-Fatigue) across ANA-RMDs (Figure 12).

Patients reported similar disease impact on their quality-of-life across all SF36 domains and 5 visual analogue scores (Pain, early morning stiffness, arthritis, global health and fatigue). There were numeric but non-significant differences in FACIT-Fatigue scores (Highest in MCTD patients, $F=1.767$, $p=0.12$). Significant differences were observed in EQ5D-5L index scores between RMD groups ($F=2.564$, $p=0.03$) which were lowest in MCTD patients (0.43) and the EQ5D mobility domain ($F=2.611$, $p=0.03$) which was lowest in pSS patients.

We then assessed whether patient-reported impact was associated with disease activity (Figure 11). Patient-reported VAS scores for pain, arthritis and early morning stiffness correlated well with BILAG-MSK scores when comparing BILAG A/B and D/E disease ($p<0.05$ in all). FACIT-fatigue scores also showed a significant correlation ($p<0.05$). EQ5D and SF36 domains scores were not associated as tightly with articular MSK-BILAG scores, likely due to confounding in composite scoring tools covering several domains.

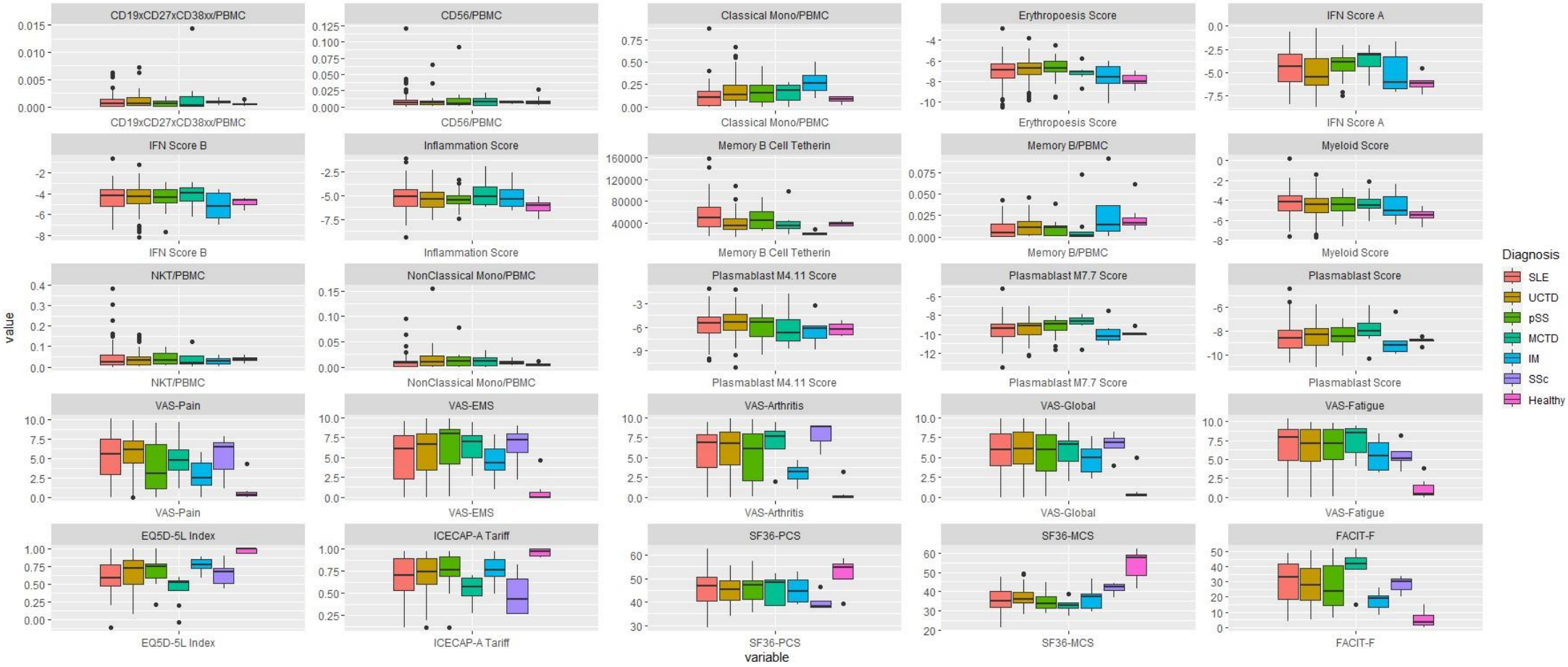


Figure 12: PRO and biomarker data by diagnosis

15.3.4 CURRENT THERAPEUTICS IN ANA-ARTHRITIS

We assessed whether the similar clinical and immunological features of ANA-arthritis across diagnoses were matched by therapeutic use. This significantly differed across diagnoses (Table 11). Current and previous biologic use was significantly associated with diagnosis (χ^2=11.933 and 12.335 respectively and p<0.05 both). Biologic use was greater in the SLE group (18/90 (20%) previously received, 12/90 (13%) currently receiving) and MCTD group (3/10 (30%) previously received, 2/10 (20%) currently receiving) compared to other diagnoses (combined, 6/113 (5%) previously received, 5/113 (4%) currently received). Previous azathioprine use was significantly greater in the SLE group and current mycophenolate use was greater in the SSc group (p<0.001 in both). Among those currently on prednisolone, doses were notably higher in the IM group (7.83mg, p<0.05). These differences may reflect current guideline impact on practice but may be unjustified given the relative homogeneity in immunological and patient-reported aspects across diagnoses.

Variable	SLE	UCTD	pSS	MCTD	IM	SSc	p
n	90	77	23	10	6	7	
Current Prednisolone (%)	33 (36.7)	10 (13.0)	3 (13.0)	4 (40.0)	4 (66.7)	1 (14.3)	0.065
Current Pred dose (mg) (mean (SD))	3.27 (5.27)	1.33 (4.35)	2.72 (8.82)	4.75 (6.71)	7.83 (8.13)	1.43 (3.78)	0.034
Current HCQ (%)	53 (58.9)	41 (53.2)	10 (43.5)	7 (70.0)	3 (50.0)	1 (14.3)	0.191
Current MTX (%)	13 (14.4)	20 (26.0)	2 (8.7)	4 (40.0)	3 (50.0)	1 (14.3)	0.047
Current MMF (%)	19 (21.1)	2 (2.6)	0 (0.0)	1 (10.0)	1 (16.7)	3 (42.9)	<0.001
Current AZA (%)	9 (10.0)	7 (9.1)	2 (8.7)	1 (10.0)	1 (16.7)	1 (14.3)	0.99
Current RTX (%)	12 (13.3)	2 (2.6)	2 (8.7)	3 (30.0)	1 (16.7)	0 (0.0)	0.036
Past AZA (%)	32 (35.6)	8 (10.4)	2 (8.7)	3 (30.0)	0 (0.0)	0 (0.0)	<0.001
Past HCQ (%)	19 (21.1)	13 (16.9)	4 (17.4)	3 (30.0)	0 (0.0)	1 (14.3)	0.725
Past MTX (%)	17 (18.9)	14 (18.2)	2 (8.7)	3 (30.0)	0 (0.0)	3 (42.9)	0.259
Past MMF (%)	18 (20.0)	6 (7.8)	1 (4.3)	2 (20.0)	1 (16.7)	0 (0.0)	0.12
Past RTX (%)	18 (20.0)	6 (7.8)	0 (0.0)	2 (20.0)	0 (0.0)	0 (0.0)	0.034
Past Cyclophosphamide (%)	15 (16.7)	6 (7.8)	0 (0.0)	1 (10.0)	1 (16.7)	1 (14.3)	0.242

Table 11: Current therapeutics in ANA+ arthritis

Regarding therapeutic confounders, among the 7 model covariates, only 3 exhibited significant associations: Chromatin antibody positivity correlated with higher hydroxychloroquine use (27.0% vs 14.3%, p=0.04), lower mean lymphocyte counts with increased azathioprine prescription (1.0 vs 1.42, p=0.01), and current mycophenolate treatment with higher previous rituximab therapy rates (34.6% vs 9.1%, p<0.01) and lower mean IMD-rank (15582.72 vs 8990.35, p<0.01).

15.3.5 Alternative predictors of disease outcomes

Statistical analysis with paired t-tests revealed several ENA subtypes linked to increased disease activity (defined by PGA). Sm, SmRNP, and RNP antibody positivity were all associated with significantly PGA scores (p-values 0.016, 0.008, and 0.005, respectively (Figure 13)

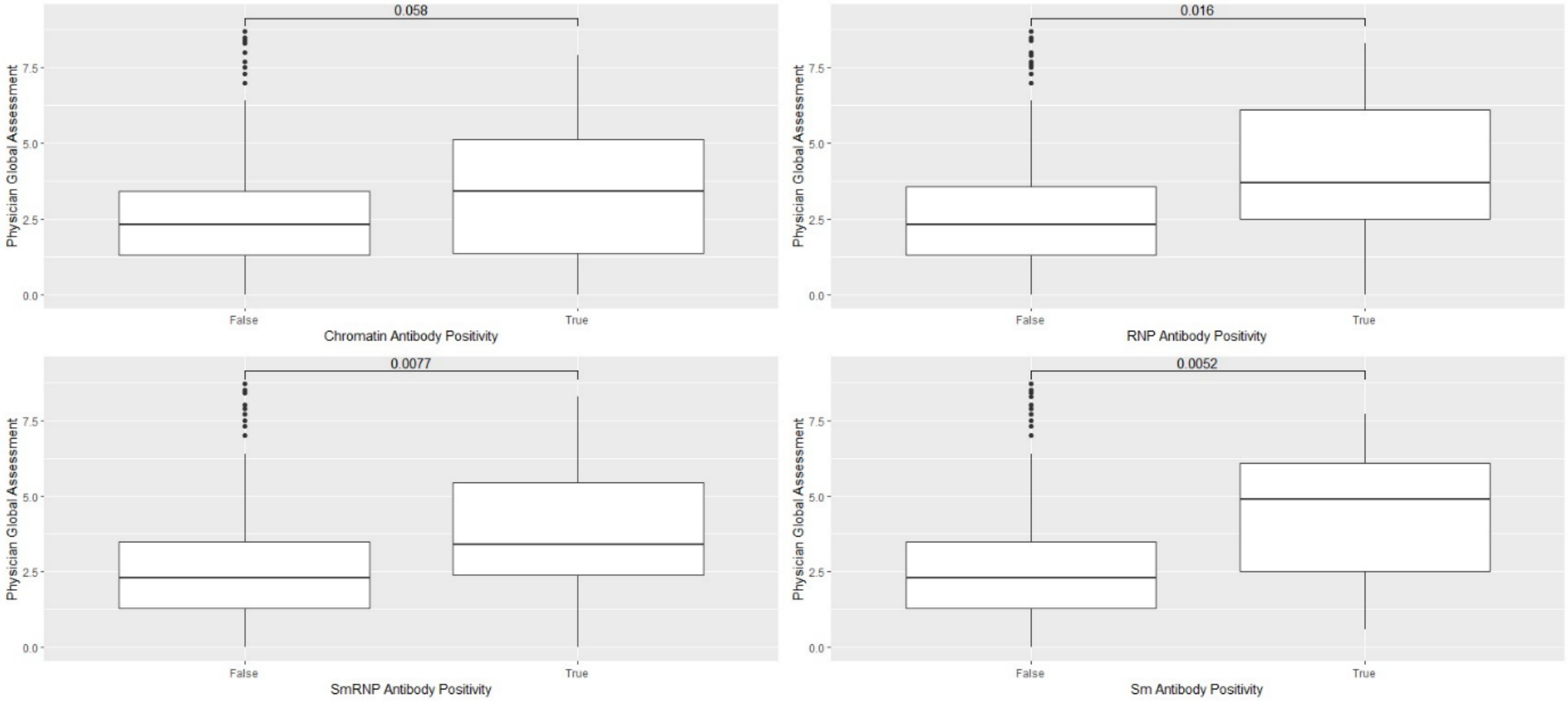


Figure 13: Diagnosis independent predictors of disease activity

15.3.6 Machine-learning reclassification of ANA-arthritis

Existing SLE trial designs recruit active disease within individual diagnoses. We estimated the proportion of patients with disease activity likely to be suitable for immunosuppressive therapy (BILAG A/B) among the 213 patients with MSK symptoms. 16 patients (7.5%) were identified with a diagnosis of SLE and BILAG-MSK A/B disease. 30 patients (13.0%) had BILAG-MSK A/B disease irrespective of their diagnosis. These values indicate prevalent baskets of patients with active arthritis across the ANA-RMD spectrum, which we explored using machine learning.

Gaussian mixture modelling identified 2 clusters (Table 12). K-means and hierarchical clustering were also trialled but were poorer identifiers of high BILAG-MSK disease activity patients than GMM, which was able to segregate all of the high BILAG-MSK disease activity patients (defined as those with A or B grade disease) into a single cluster (Appendix A). Overall, cluster 1 contained more patients with inflammatory features (High MSK Activity Cluster) and cluster 2 contained more patients with non-inflammatory causes of joint pain (Low MSK Activity Cluster). Cluster 1 comprised 89 patients (41.8%) including all patients with BILAG A/B MSK disease. Cluster 1 patients were younger with a lower proportion of UCTD and pSS. They included a higher proportion of SLE and MCTD patients with a lower mean Charlson Comorbidity Index ($p = 0.002$). Cluster 1 contained substantively lower numbers of patients with nodal OA (3.4% for cluster 1 vs. 11.3% for cluster 2, $p = 0.065$); radiographic evidence of OA (10.1% for cluster 1 vs. 19.4% for cluster 2, $p = 0.10$); fibromyalgia symptom pain and stiffness (9% for cluster 1 vs. 19.4% for cluster 2, $p = 0.058$); fibromyalgia allodynia (3.4% vs. 10.5%, $p = 0.093$). RNP/SmRNP/Sm antibody positivity was significantly greater in Cluster 1 patients. Numeric BILAG, ESSDAI, SLEDAI and physician global assessment scores were all significantly higher in Cluster 1 patients ($p < 0.001$ in all). Memory B-cell tetherin and Interferon score A expression was also significantly higher in Cluster 1 ($p = 0.018$ and $p = 0.021$, note that with untransformed gene expression scores, numerically lower values represent higher gene expression). Cluster 1 patients received more frequent and higher dose prednisolone. PCA plots formed from the 7 GMM covariates are shown in Figure 14.

Key potential confounders, including IMD rank, mucocutaneous, renal, neurological, and gastrointestinal BILAG scores, as well as Rodnan skin score, showed no significant differences between the GMM-derived clusters. Interestingly, Cluster 1 patients exhibited higher MSK disease activity despite significantly higher rates of previous treatment with rituximab (14.6% vs. 5.6%) and mycophenolate (19.1% vs. 8.9%) ($p<0.05$ for both).

The distribution of patients between legacy diagnoses and new GMM clusters is shown in Figure 15 to illustrate potential trial stratification strategies. Conventional trial designs recruit patients with SLE and swollen joints. After reclassification, the High MSK Activity Cluster (cluster 1) included all those patients as well as larger numbers from other RMDs. All patients in the High MSK Activity cluster with swollen joints would be eligible for an ANA-arthritis trial design and represent twice as many patients as a conventional SLE trial design. The remaining patients in the High MSK Activity Cluster lacked swollen joints at assessment but were similar in terms of immune biomarkers and clinical impact. These patients may be hypothesised to have a higher rate of joint inflammation when assessed over a longer time period, under different glucocorticoid or other immunosuppressive medications, or with musculoskeletal imaging, which has been shown to detect joint inflammation in a larger percentage of symptomatic populations (113). Therefore, these patients in the High MSK Activity Cluster may be additional candidates for therapy licensed for ANA-arthritis in clinical practice

	GMM Cluster 1 High MSK Activity N = 89	GMM Cluster 2 Low MSK Activity N = 124	p-value
Demographics			
n	89	124	
Sex = M (%)	13 (14.6)	13 (10.5)	
Age (mean (SD)), years	44.28 (14.46)	53.01 (12.83)	
Diagnosis (%)			<0.001
SLE (n, % of Cluster)	48 (53.9)	42 (33.9)	
UCTD	23 (25.8)	54 (43.5)	
pSS	3 (3.4)	20 (16.1)	
MCTD	9 (10.1)	1 (0.8)	
IIM	4 (4.5)	2 (1.6)	
SSc	2 (2.2)	5 (4.0)	
Other demographics			
Charlson Comorbidity Index (mean (SD))	1.78 (1.16)	2.42 (1.65)	0.002
Non Inflammatory			
FMS pain/stiffness (%)	8 (9.0)	24 (19.4)	0.058
FMS allodynia (%)	3 (3.4)	13 (10.5)	0.093
Hypermobility syndrome (%)	2 (2.2)	4 (3.2)	0.995
Nodal OA (%)	3 (3.4)	14 (11.3)	0.065
XR proven OA (%)	9 (10.1)	24 (19.4)	0.1
Current Therapies			
Current prednisolone (%)	41 (97.6)	14 (11.3)	<0.001

	GMM Cluster 1 High MSK Activity N = 89	GMM Cluster 2 Low MSK Activity N = 124	p-value
Current prednisolone dose (mean (SD))	5.53 (7.70)	0.58 (1.68)	<0.001
Current HCQ (%)	50 (56.2)	65 (52.4)	0.686
Current MTX (%)	18 (20.2)	25 (20.2)	1
Current MMF (%)	11 (12.4)	15 (12.1)	1
Current AZA (%)	9 (10.1)	12 (9.7)	1
Current RTX (%)	13 (14.6)	7 (5.6)	0.048
Previous Therapies			
Previous AZA (%)	22 (24.7)	23 (18.5)	0.359
Previous HCQ (%)	15 (16.9)	25 (20.2)	0.666
Previous MTX (%)	20 (22.5)	19 (15.3)	0.25
Previous MMF (%)	17 (19.1)	11 (8.9)	0.048
Previous RTX (%)	14 (15.7)	12 (9.7)	0.263
Previous Cyclo (%)	10 (11.2)	14 (11.3)	1
Selected other clinical features (see supplement)			
Raynauds (%)	32 (36.0)	18 (14.5)	0.001
Alopecia (%)	25 (28.1)	14 (11.3)	0.003
Immunology			
dsDNA (%)	27 (30.3)	35 (28.2)	0.856
Ro60 (%)	27 (30.3)	45 (36.3)	0.448
Ro52 (%)	23 (25.8)	27 (21.8)	0.598
La (%)	9 (10.1)	13 (10.5)	1

	GMM Cluster 1 High MSK Activity N = 89	GMM Cluster 2 Low MSK Activity N = 124	p-value
Sm (%)	20 (22.5)	0 (0.0)	<0.001
SmRNP (%)	33 (37.1)	0 (0.0)	<0.001
RNP (%)	19 (21.3)	0 (0.0)	<0.001
Chromatin (%)	45 (50.6)	0 (0.0)	<0.001
Clinical/Lab tests			
Lymphocyte count x 10*9/L (mean (SD))	1.30 (0.69)	1.44 (0.63)	0.133
BILAG Scores			
BILAG Numeric (mean (SD))	6.62 (6.64)	2.30 (2.76)	<0.001
BILAG Total (%)			<0.001
A	28 (31.5)	4 (3.2)	
B	16 (18.0)	21 (16.9)	
C	41 (46.1)	82 (66.1)	
D/E	4 (4.5)	17 (13.7)	
BILAG Mucocutaneous (%)			0.1
A	4 (4.5)	3 (2.4)	
B	15 (16.9)	16 (12.9)	
C	15 (16.9)	10 (8.1)	
D/E	55 (61.8)	95 (76.6)	
BILAG MSK (%)			<0.001
A	22 (24.7)	0 (0.0)	
B	8 (9.0)	0 (0.0)	

	GMM Cluster 1 High MSK Activity N = 89	GMM Cluster 2 Low MSK Activity N = 124	p-value
C	44 (49.4)	91 (73.4)	
D/E	15 (16.9)	33 (26.6)	
BILAG General (%)			0.015
B	6 (6.7)	2 (1.6)	
C	5 (5.6)	1 (0.8)	
D/E	78 (87.6)	121 (97.6)	
BILAG Haematological (%)			0.054
B	1 (1.1)	0 (0.0)	
C	29 (32.6)	25 (20.2)	
D/E	59 (66.3)	99 (79.8)	
BILAG Renal (%)			0.618
B	3 (3.4)	4 (3.2)	
C	2 (2.2)	6 (4.8)	
D/E	84 (94.4)	114 (91.9)	
Other physician disease activity measurements			
ESSDAI Total (mean (SD))	3.79 (4.56)	1.12 (2.07)	<0.001
SLEDAI Total (mean (SD))	5.81 (3.97)	3.31 (2.56)	<0.001
Physician global assessment (mean (SD))	3.86 (2.39)	2.13 (1.55)	<0.001
Patient reported outcome scores			
Pain VAS (mean (SD))	5.40 (2.49)	4.90 (3.07)	0.261
EMS VAS (mean (SD))	6.10 (2.65)	5.51 (3.01)	0.204

	GMM Cluster 1 High MSK Activity N = 89	GMM Cluster 2 Low MSK Activity N = 124	p-value
Arthritis VAS (mean (SD))	5.85 (2.93)	5.89 (2.91)	0.945
Global VAS (mean (SD))	5.79 (2.59)	5.61 (2.86)	0.675
Fatigue VAS (mean (SD))	6.78 (2.62)	6.38 (3.08)	0.373
EQ5D-5L Index (mean (SD))	0.63 (0.20)	0.65 (0.23)	0.509
EQ5D Self Care (mean (SD))	1.99 (1.18)	1.69 (1.06)	0.079
ICECAP Total (mean (SD))	0.70 (0.22)	0.68 (0.22)	0.432
SF36 Physical Component Score (mean (SD))	45.64 (7.32)	45.09 (6.23)	0.653
SF36 Mental Component Score (mean (SD))	36.44 (5.76)	36.15 (5.50)	0.778
FACIT Fatigue Total (mean (SD))	30.19 (12.39)	28.18 (13.85)	0.365
Biomarkers			
Memory B cell tetherin MFI (mean (SD))	52882 (31936)	41979 (20287)	0.018
Interferon Score A dCt (mean (SD))	4.33 (1.95)	4.92 (1.66)	0.021
Erythropoesis Score dCt (mean (SD))	7.21 (1.38)	6.87 (1.19)	0.064
Inflammation Score dCt (mean (SD))	5.15 (1.41)	5.29 (1.14)	0.453
Memory B cells/PBMCs (mean (SD))	0.009 (0.01)	0.013 (0.01)	0.057

Table 12: Characteristics of ANA-RMD clusters

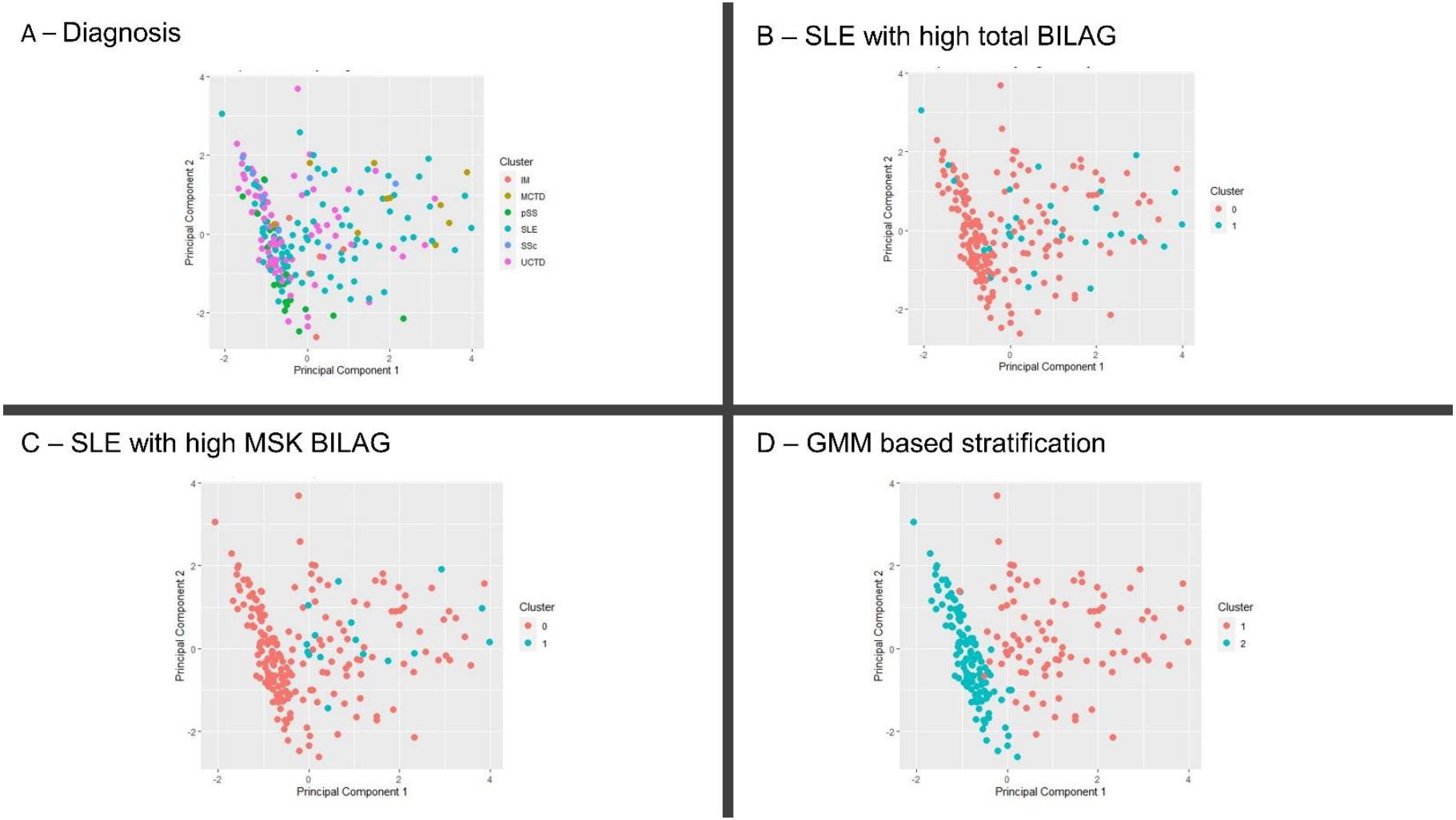


Figure 14: Collated PCA plots

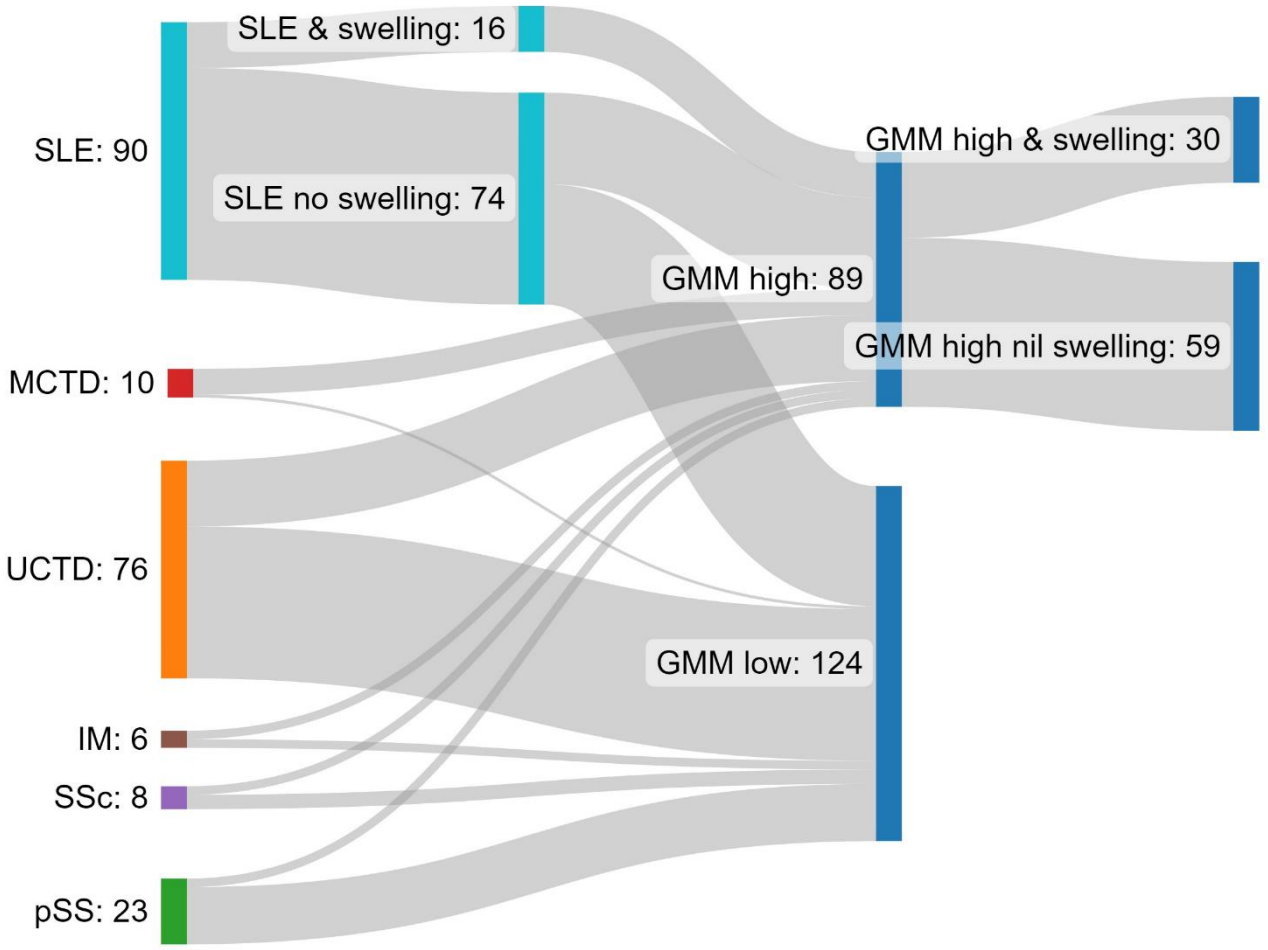


Figure 15: Distribution of patients between legacy and GMM classifications

15.4 DISCUSSION

This is the first work assessing the clinical impact and immune profile of arthritis across multiple RMDs in a systematically collected, richly phenotyped, multi-disease cohort. We demonstrate that ANA-positive RMD patients with musculoskeletal symptoms contain a High MSK Disease Activity population that is homogenous in clinical features, patient-reported impact and immune profile. The existing classification had previously distributed ANA-arthritis patients into other groups based on their additional disease features, potentially resulting in unjustified variations in therapy. Instead, we suggest and define a novel classification that consolidates all ANA-arthritis patients into a single group. This classification can facilitate basket trials, new therapy indications and inform routine clinical care guidelines.

We identified few differences between patients with musculoskeletal symptoms across RMD diagnoses, including physician and patient-reported clinical outcomes and biomarkers, with the exception that MCTD was generally worse. As expected from previous work, patients with musculoskeletal symptoms contained a mixture of patients with objective disease activity along with patients with low disease activity and non-inflammatory explanations for pain collectively, or for each legacy diagnosis.

During machine learning (ML) analysis patients with MSK symptoms were not sorted according to legacy diagnosis. Instead, the GMM approach generated High and Low MSK Disease Activity clusters, showing greater homogeneity compared to diagnoses like SLE and undifferentiated CTD. The High Disease Activity cluster contained every patient with joint swelling, and other features such as higher physician global assessment, prednisolone dose and IFN-Score. The Low Disease Activity cluster contained more patients with features of fibromyalgia and osteoarthritis. Patient-reported outcomes like pain, fatigue, and quality-of-life didn't vary between clusters, as expected, given that both inflammatory and non-inflammatory causes of pain may equally affect patient experience. The High Disease Activity cluster included many patients without joint swelling, grouped based on other features such as prednisolone dose, Sm, RNP, Sm/RNP, Chromatin and Ro antibodies, lymphocyte count and IMD-rank. Although lacking documented joint swelling on the study visit day, these patients might exhibit joint inflammation on ultrasound imaging (as we previously demonstrated with the same antibody subtypes) or would present with joint swelling if the prednisolone dose were reduced or if assessed over a longer duration (113).

The High MSK Disease Activity group may be suitable for basket clinical trials. Current SLE clinical trials involve patients with diverse organ manifestations, requiring complex disease activity instruments to compare severity and response across different presenting

complaints. These trials may also necessitate different standard-of-care therapies. These factors may have contributed to inconsistent trial results(381,382). Conversely, clinical trials in ANA-arthritis would recruit a more homogeneous population, despite different legacy diagnoses. This enables the use of robust organ-specific outcome measures, such as the LAMDA, which combines swollen joint count, patient and physician VAS, and acute phase markers. In SLE, this principle is shown by a litifilimab phase-2 trial meeting its primary endpoint of joint counts, or baricitinib trials achieving MSK-specific secondary and exploratory variables(381,390). Multisystem disease activity tools would only be required to monitor for worsening in other organs. ANA-arthritis trials could recruit more patients , and possibly incorporate musculoskeletal imaging. The resulting evidence base would be relevant to a larger patient population than SLE, addressing a healthcare inequality.

Importantly the Low MSK Disease Activity group still had a significant symptom burden, but with less evidence of active inflammation amenable to immunosuppression. These patients could potentially be offered more appropriate non-immunosuppressive therapy modalities. Our data suggests a large patient population whose needs may not be met by current research and guidelines.

SLE management guidelines have been published, covering diverse areas including diagnosis, assessment, care delivery and therapeutics(97). Many recommendations in these guidelines are not specific to MSK manifestations. However, for other patients with ANA-arthritis there are no guidelines. Further research on the described population could enhance patient outcomes in routine practice. In DEFINITION, notably, a higher proportion of non-SLE patients had ANA-arthritis compared to SLE patients in terms of pure numbers.

Biomarker analysis can help determine whether a population is immunologically homogenous and appropriate for similar therapies. The biomarker results in this study are more consistent and logical than others reported in SLE patients with arthritis. In SLE, IFN-I Scores correlate with increased skin disease activity, but not always with increased MSK disease activity—in certain studies, MSK disease activity appeared lower in IFN high patients(380,386). In our study, the High MSK Disease Activity cluster showed significantly higher tetherin and IFN Score A expression.

While our study comprised a large and extensively phenotyped cohort, certain limitations persist. Notably, the sample sizes for some diagnoses were small, thereby limiting the generalizability of findings within these groups. Consequently, validation in other cohorts along with prospective studies are essential. Additionally, the diversity of the cohort was limited by the regional population from which it was recruited. South Asian patients were better represented in our study than many other cohorts, but other groups were under-

represented. Longer follow-up and imaging data were unavailable in our study, future research should investigate these. The articular component of the BILAG MSK appears valid across these diagnoses but better instruments in development, such as the LAMDA and joint counts, should be validated.(391). Finally, although we measured a wide range of gene expression and flow cytometric biomarkers, there are others emerging in autoimmunity(378).

As patient age is included as a covariate in our model, we can explore an interesting concept regarding disease stratification. Patients with an index presentation such as lupus nephritis or interstitial lung disease are treated according to established guidelines. If these patients later develop a predominant musculoskeletal (MSK) manifestation, it may be appropriate to categorize them within the arthritis basket. Therefore, inclusion in a basket is not static for a patient throughout their disease duration; but dependent on their predominant issue at the time of assessment.

In conclusion, these data indicate that the ANA+ arthritis basket has more unifying than dividing aspects in terms of quality-of-life impact, therapeutic usage, and biomarker variables. We describe an alternative means to classify patients with arthritis across ANA-RMDs. Clinical trials in this population could generate larger effect sizes and make new guidelines and interventions available to more patients.

16 RESULTS 3: VARIATIONAL AUTOENCODER BASED RECLASSIFICATION OF ANA-RMD ACROSS A LARGE EUROPEAN COHORT, HYBRID CLINICAL AND BIOMARKER-BASED CLUSTERING.

16.1 INTRODUCTION

ANA-associated Rheumatic Musculoskeletal Diseases (ANA-RMD) are a group of heterogeneous autoimmune disorders with diverse clinical features unified by the presence of anti-nuclear antibodies. They are treated with similar therapies, primarily immunosuppression. Despite this, ANA-RMD is subclassified into distinct diseases: Systemic Lupus Erythematosus (SLE), Sjogren's Disease (SjD), idiopathic inflammatory myositis (IIM), Systemic Sclerosis (SSc), Mixed Connective Tissue Disease (MCTD) and Antiphospholipid Syndrome (APS) based on clinical symptoms, signs, and routine immunological tests.

This classification has disadvantages. Within each of these entities there is heterogeneity in presentation, response to therapy, and underlying immunopathogenesis. This leads to small effect sizes in clinical trials, although with high-responding subgroups defined by symptoms or biomarkers(128). Patients with different diagnoses may be more clinically similar than those within the same group. This leads to treatment inequity, wherein the available treatments for a common feature, such as arthritis, may differ depending on the underlying diagnosis. Trials are more difficult in less common diseases. Many patients have an overlap, meeting criteria for more than one disease, whilst a large proportion do not fulfil any specific diagnosis, often being termed Undifferentiated-CTD (UCTD).

A modern data-driven approach to reclassify larger proportions of ANA-RMD patients into more phenotypically homogenous and clinically meaningful groups may be preferable. Basket trials could group together individuals suitable for a shared therapeutic target, such as an interferon-high subset of each disease of interferon-blocking therapy, or with a shared clinical feature, such as arthritis, suitable for a similar outcome measure and standard of care. This approach has been successful in licensed treatments such as Nintedinib for pulmonary fibrosis. Trials designed in this way reduce treatment inequity, by emphasising clinical and biomarker phenotype over physician assigned disease label.

Machine learning has been used to successfully stratify several individual diseases(392,393)(394). However, to characterize the relevant features of the ANA-RMD population, many potential variables, with different metric properties and distributions are required. Traditional clustering methods, like partitioning around medoids, hierarchical, and k-means, perform poorly with high-dimensional datasets due to exponential increases in potential space, leading to overfitting and high computational costs.

Deep learning is more effective in maintaining dataset richness while reducing dimensionality to identify relevant clusters. Variational autoencoders (VAEs) are a type of deep neural network that encodes high-dimensional datasets into a lower-dimensional latent layer. The model's effectiveness is measured by its accuracy in reconstructing the original dataset from this latent space(395). Constraining the hidden layer to a smaller dimensionality forces the autoencoder to capture key features of the training set, denoise data, and preserve overall structure and complex non-linear relationships between variables.(395). Consequently, VAEs have been used to compress data for subsequent analysis with clustering techniques like k-means, hierarchical, and Gaussian mixture modelling (GMM). For example, a combined approach using a VAE and GMM has been applied to single-cell RNA sequencing data to determine the cell developmental trajectories of human preimplantation embryos(243). Other applications of this approach include detection of bleeding events in electronic health records and natural language processing of radiology reports to detect complications of ischaemic stroke in radiology(396,397).

We hypothesised that a hybrid approach, combining a VAE and traditional clustering, applied to two large datasets of richly phenotyped patients would identify more meaningful classes that each draw from multiple ANA-RMD legacy diagnoses. We further hypothesised that these ANA-RMD Classification (ARC) classes derived in this way would show better association with clinical impact and therapy response in cross-sectional and longitudinal follow up compared to the legacy diagnoses.

16.2 METHODS

The overall analysis pathway is summarised in Figure 16. Briefly, (1) We identified key covariates driving clinical decision making and the patient experience in ANA-RMDs by survey of clinical specialists and PPIE group interaction; (2) we compressed hybrid clinical, immunological and gene expression data from the European PRECISESADS ANA-RMD cohort using a deep learning approach; (3) we used this compressed data to reclassify the spectrum of ANA-RMD into more homogeneous and clinically meaningful ARC classes using advanced computational techniques; (4) we replicated this reclassification in the

DEFINITION cohort; (5) we assessed the cross-sectional impact of class membership as well as the impact on key clinical outcomes and healthcare utilisation at 5 years follow up.

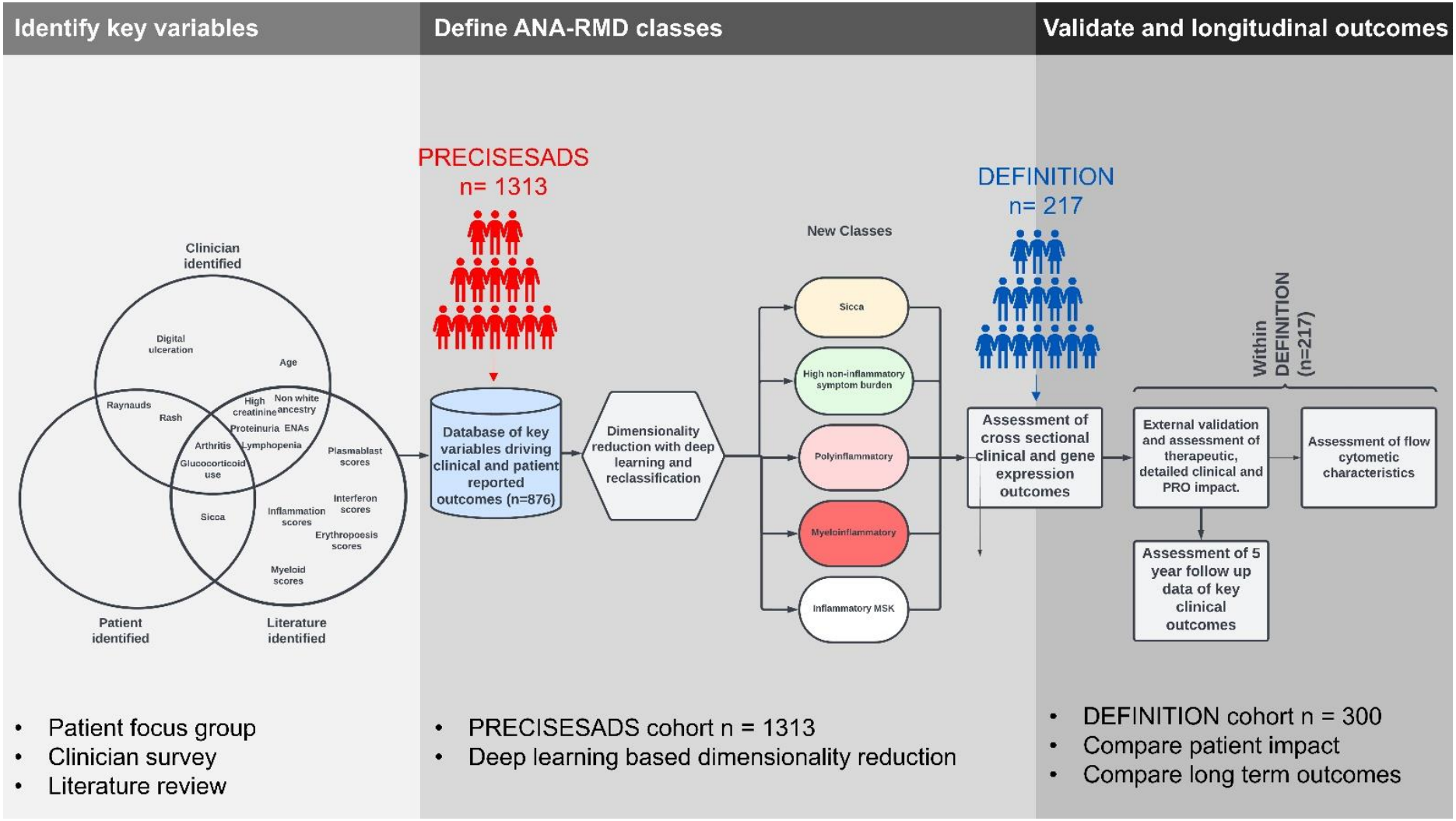


Figure 16: Analysis pathway for project

16.2.1 Prioritisation of covariates

Prioritisation of model covariates was undertaken following a two-stage process. A formalised patient and public involvement (PPI) focus group with ANA-RMD patient was carried out. Covariates identified as important within this session were combined with those identified as important by a survey of 10 specialist clinicians working in ANA-RMDs. Biomarkers and gene expression scores identified as important within the literature were also incorporated.

16.2.2 Deep learning

Deep learning models were developed using the Keras package (v2.15.0) in R using a TensorFlow backend(v2.16.0)(398,399) and the Reticulate package for running Python script(v1.38.0)(400–402). Data compression was carried out using a variational autoencoder neural network prior to unsupervised clustering using K-means, hierarchical clustering and Gaussian mixture modelling in R using base R functionality and the mclust package(v6.11)(403).

Models were trained in the PRECISESADS dataset to define new classes. Thereafter clinical characteristics of these classes was assessed by generation of summary statistics using the TableOne package in R (v0.13.2). External validation of generated models was carried out within the DEFINITION dataset using the same model parameters trained in PRECISESADS. Further evaluation of clinical characteristics of the classes in DEFINITION was carried out by generating further summary statistics. Distribution of key variables was compared between legacy diagnoses and new classes. Bonferroni correction was applied to adjust for multiple hypothesis testing.

The neural network architecture comprised a word embedding layer derived from key diagnostic terms within the case report forms (CRFs) including Raynaud's, digital ulcers, rash, sicca, telangiectasia, arthritis and nephritis. Additionally, a scaled tensor containing two validated interferon stimulated gene expression scores were included (IFN Score A and IFN Score B), along with scores for genes annotated to plasmablast, myeloid lineage, inflammation and erythropoiesis function based on gene expression scores derived from selected genes in previously described modules shown to associate with clinical features of SLE and therapy response, as described below (378,388). Prednisolone usage, age, and one-hot encoded categorical variables (Elevated WCC, elevated creatinine and extractable nuclear antibody positivity) were also included. Only variables common to both DEFINITION and PRECISESADS were considered. Following compression within a variational autoencoder neural network constructed comprised of an input dimension of 25 covariates narrowing via an intermediate layer of 16 neurons to a latent layer of 8. Data compression to

6 neuron and 4 neuron latent layers was also trialled but conferred no additional benefit. Batch sizes of 1000 were used with an epsilon value of 0.0001, LeakyReLU activation factor to avoid the issue of dead neurons in zero inflated data and the “adam” optimizer(404,405). 1000 epochs of training were performed with a mean square error loss function.

Data extracted from the 8-neuron layer was then passed to traditional clustering algorithms and assessed using the m3c package in R (v1.26.0). Stability of clustering results was analysed using the m3c package in R to incorporate Monte Carlo reference procedures as described in previous work(406). The relative cluster stability index (RCSI) and proportion of ambiguous clustering (PAC) scores were generated along with cumulative density function (CDF) plots and entropy plots to help select the optimum k values. Within this package hierarchical, K-means, partitioning around medoids (PAM) and spectral clustering techniques were assessed. Optimal cluster number was also assessed through generation of silhouette plots for the PRECISESADS and DEFINITION datasets using GMM, hierarchical and K-means derived clusters, assessment of the RSCI, entropy plots and elbow plot assessment, as well as analysis of the clinical significance of these classes.

Efficacy of deep learning reclassification techniques within the PRECISESADS dataset was compared to traditional legacy diagnosis by comparison of the silhouette score and PCA plots. The clinical significance of cluster membership was then assessed within the PRECISESADS cohort in terms of disease activity defined by physician global assessment score (PGA), prednisolone dosage, conventional immunosuppressant use and biologic therapy use.

VAE architecture is summarised in Figure 17.

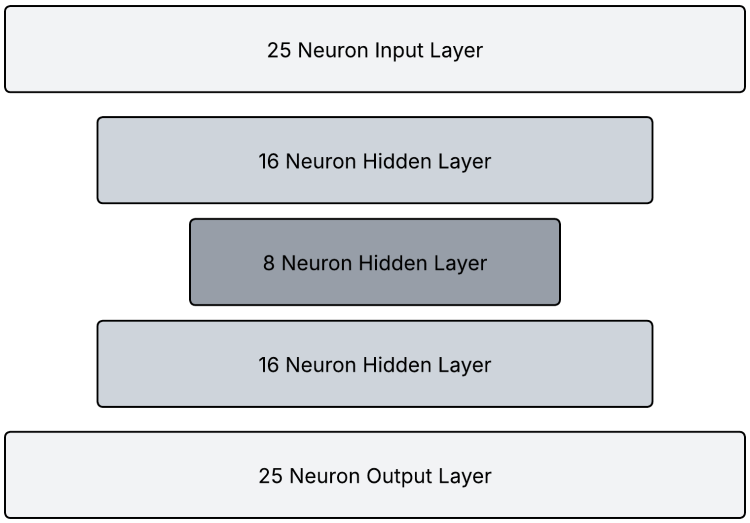


Figure 17: Variational autoencoder architecture

16.2.3 Validation and evaluation of clinical impact

Further validation was carried out by applying the same model to the 294 patient DEFINITION cohort. Significance of group membership here was assessed again in terms of conventional and biologic immunosuppressant use and prednisolone dosage. However more detailed assessment of disease activity was assessed in terms of the British Isles Lupus Assessment Group 2004 (BILAG-2004) index, the European Sjogren's syndrome disease activity index (ESSDAI) score, the Myositis Disease Activity Assessment Tool (MITAX) and the Physician Global Assessment (PGA, 0-10).

The patient reported impact of class membership was assessed in DEFINITION using patient VAS scores, the SF36 score, EQ5D-5L index and domain scores (quality of life) and ICECAP-A (perceived capability) scores.

Long term outcomes were analysed in DEFINITION up to 5 years post recruitment. Key outcomes including all-cause mortality, hospitalization, emergency department attendance, steroid escalations, outpatient rheumatology or combined clinic appointments, rituximab use and cyclophosphamide use were analysed per year of available follow up and in a time to event manner using Kaplan Meier survival plots in the ‘survival’ package in R(v3.6-4)(407).

16.2.4 Laboratory measures

16.2.4.1 Within DEFINITION

Within the DEFINITION cohort full blood count, complement C3 and C4 levels and ANA subtypes, including anti-double-stranded DNA, Ro-52, Ro-60, La, Sm, SM/RNP, RNP, Scl-70, Jo-1, Centromere, Chromatin and Ribosomal-P antibodies (Bioplex multiplex analyser) were measured in routine diagnostic laboratory. PBMC subsets were analysed using 8-colour flow cytometry as a proportion of total PBMC count (T cells (CD3+CD56-), NK cells (CD56+), NKT cells (CD3+CD56+), Memory B cells (CD19+CD27+, Plasmablasts (CD19+/-CD27+CD38++), classical monocytes (CD14++CD16-) intermediate monocytes (CD14++CD16+)and non-classical monocytes (CD14+CD16+)). Tetherin (CD317) mean fluorescence intensity was quantified on each cell subset, as previously described, with memory B cell level used as the primary biomarker (168).

Gene expression scores analysed without the knowledge of patients’ clinical status. These included two validated interferon scores (IFN Score A and IFN Score B) as well as scores for genes annotated to plasmablast, myeloid lineage, inflammation and erythropoiesis function from previously described modules based on their known molecular function(378,388,408) (386,387). These scores were calculated as a median ΔCT value from the genes listed in Table 13. Briefly, PBMC were separated using density gradient method (Lymphoprep; Alere

Technologies, Norway) from EDTA-anticoagulated blood. Total RNA purification kit (Norgen Biotek, Canada) was used followed by quantitative real time reverse transcriptase-PCR (qRT-PCR) using TaqMan assays (Applied Biosystems, Invitrogen).

Gene Expression Score	Genes
Interferon Score A	*IFI44L, CCL8, ISG15, XAF1, IFI44, GBP1, IFI27, IRF7, CXCL10, CEACAM1, RSAD2, IFIT1*
Interferon Score B	*SERPING1, IFIH1, IFI16, PHF11, LAMP3, NT5C3B, TRIM38, UNC93B1, BST2, SOCS1, SP100, TAP1, STAT1, UBE2L6*
Plasmablast M7.7 Score	*ATP5G1, CDC2.CDK1, CHPF, CIT, GLICCI1, GTSE1*
Plasmablast M4.11 Score	*TNFRSF17, TXNDC5, IGJ.JCHAIN*
Plasmablast Score	*ATP5G1, CDC2.CDK1, CHPF, CIT, GLICCI1, GTSE1, TNFRSF17, TXNDC5, IGJ.JCHAIN*
Myeloid Score	*CD55, BST1, LILRA3, CEACAM4, MKNK1, FLJ20273.RBM47., HCK, FLOT1, CKAP4, BCL6, GPR97, ITGAM, AQP9, ADAM19, HMGB2, NCF4, IL1RN, CD63, IL17R.IL17RA., MMP25, PBEF1.NAMPT, NFIL3*
Erythropoiesis Score	*DNAJA4, IFIT1L.IFIT1B, UBXD1, RAB2B, BCL2L1, GATA1, UBE2O, MAP2K3, PIP5K2A, FLCN, HPS1, MCOLN1*
Inflammation Score	*SOCS3, IRAK3, IL18R1, CLEC4D, MCEMP1, SERPINA1, IL1R2, MCTP2, FKBP5, OSM, KREMEN1, FCAR, CR1, SLPI, MAPK14, PGLYRP1, ANXA3, S100P, MMP9, S100A12, GRB10*

Table 13 Gene expression score compositions

16.2.5 Within PRECISESADS

ANA-RMD patients with SLE, pSS, SSc, UCTD and MCTD were identified within the European PRECISESADS cohort. Patients with a primary diagnosis of rheumatoid arthritis were excluded. Classification criteria were as documented in previous work(191); 1997 update of 1982 ACR SLE criteria, 2013 ACR/EULAR SSc criteria. American-European consensus group pSS criteria and Alarcon-Segovia criteria for MCTD. UCTD patients were defined as having symptoms/features of ANA-RMD and a positive ANA test but failing to fulfil any of the above criteria.

For blood samples collected between March 2016 and June 2019 all autoantibodies were measured at a single centre (University Hospital of Brest) using a chemilumescent IDS-iSYS immunoanalyser (Immunodiagnostic systems) as per previous studies(191,378). Antinuclear antibody detection was performed on Hep2 cells using an in-house technique.

Gene expression was evaluated using RNA-sequencing from whole blood as previously described(408). For the current project, gene expression scores were derived from these

RNA-Seq data. Samples with less than $7x10^6$ or greater than $1.5x10^7$ total reads were excluded during quality control as were genes with fewer than 10 counts in 100 samples. Samples with an RNA integrity number ≥7 were included for analysis. Normalisation of RNA-Seq data was carried out using 3 packages in R (Limma-Voom, DNASeq2- VST and edgeR-counts per million (CPM)). Voom proved to be the most effective package, and this data was taken forward for analysis. Gene expression scores were calculated from the pre-defined selected genes in the same way as for DEFINITION.

16.3 RESULTS

16.3.1 Baseline characteristics of PRECISESADS and DEFINITION cohorts

Key baseline characteristics for the PRECISESADS and DEFINITION cohorts are summarized in Table 14..

PRECISESADS	MCTD	PAPs	SjD	SLE	SSc	UCTD	p
n	59	46	220	262	217	72	
Demographics							
Age (mean (SD))	51.05 (14.63)	47.50 (12.42)	57.97 (13.34)	46.05 (13.81)	58.73 (13.50)	53.86 (13.96)	<0.001
Ethnicity white (%)	53 (89.8)	46 (100.0)	216 (98.2)	251 (95.8)	211 (97.2)	71 (98.6)	0.016
Medical hx							
Smoker (%)	8 (13.6)	11 (23.9)	19 (8.6)	48 (18.3)	26 (12.0)	15 (20.8)	0.011
Diabetes (%)	0 (0.0)	1 (2.2)	5 (2.3)	5 (1.9)	14 (6.5)	3 (4.2)	0.169
HTN (%)	17 (28.8)	17 (37.0)	71 (32.3)	75 (28.6)	67 (30.9)	21 (29.2)	0.849
Medical hx of Arthritis (%)	38 (64.4)	7 (15.2)	67 (30.5)	192 (73.3)	65 (30.0)	40 (55.6)	<0.001
Medical hx of Rash (%)	8 (13.6)	4 (8.7)	24 (10.9)	153 (58.4)	4 (1.8)	10 (13.9)	<0.001
Medical hx of Raynauds (%)	53 (89.8)	5 (10.9)	66 (30.0)	103 (39.3)	210 (96.8)	38 (52.8)	<0.001
Medical hx of Sicca (%)	21 (35.6)	3 (6.5)	204 (92.7)	56 (21.4)	73 (33.6)	29 (40.3)	<0.001
Medical hx of Nephritis (%)	5 (8.5)	2 (4.3)	17 (7.7)	84 (32.1)	3 (1.4)	3 (4.2)	<0.001
Therapeutics							
Current glucocorticoid therapy (%)	25 (42.4)	3 (6.5)	44 (20.0)	138 (52.7)	46 (21.2)	25 (34.7)	<0.001
Glucocorticoid dose (mean (SD))	5.36 (3.41)	10.33 (8.74)	5.93 (3.39)	6.95 (6.57)	6.02 (2.23)	6.64 (4.64)	0.455
Current antimalarial therapy (%)	27 (45.8)	7 (15.2)	81 (36.8)	187 (71.4)	16 (7.4)	32 (44.4)	<0.001
Current DMARD therapy (%)	21 (35.6)	1 (2.2)	34 (15.5)	94 (35.9)	52 (24.0)	14 (19.4)	<0.001
Current biologic therapy (%)	2 (3.4)	0 (0.0)	0 (0.0)	0 (0.0)	7 (3.2)	4 (5.6)	0.001
Laboratory							
dsDNA (%)	5 (8.6)	4 (8.7)	7 (3.2)	85 (32.7)	6 (2.8)	3 (4.2)	<0.001
Ro60 (%)	5 (9.3)	2 (4.3)	117 (68.4)	62 (27.0)	8 (3.9)	10 (14.5)	<0.001
Ro52 (%)	10 (18.5)	3 (6.5)	119 (69.2)	48 (20.9)	25 (12.3)	9 (13.0)	<0.001
La (%)	2 (3.6)	1 (2.2)	67 (35.4)	16 (6.7)	3 (1.4)	3 (4.2)	<0.001
Sm (%)	14 (25.5)	0 (0.0)	2 (1.1)	13 (5.4)	0 (0.0)	0 (0.0)	<0.001
U1-RNP (%)	41 (74.5)	3 (6.5)	14 (7.4)	50 (20.9)	13 (6.2)	14 (19.7)	<0.001
SCL-70 (%)	4 (7.3)	0 (0.0)	0 (0.0)	4 (1.7)	77 (36.5)	0 (0.0)	<0.001
Centromere (%)	0 (0.0)	0 (0.0)	8 (3.7)	4 (1.5)	100 (46.5)	4 (5.6)	<0.001
Chromatin (%)	11 (20.0)	4 (10.0)	19 (9.6)	133 (55.0)	13 (8.3)	7 (10.9)	<0.001
B2GP (%)	1 (1.7)	22 (48.9)	1 (0.5)	31 (12.0)	2 (0.9)	1 (1.4)	<0.001

DEFINITION	IM	MCTD	SjD	SLE	SSc	UCTD	p
n	15	8	27	79	12	78	
Demographics							
Age (mean (SD))	52.67 (16.85)	46.00 (11.45)	54.56 (15.04)	45.99 (12.74)	54.75 (14.10)	51.40 (13.14)	0.021
Ethnicity white (%)	13 (86.7)	5 (62.5)	24 (88.9)	48 (60.8)	9 (75.0)	60 (76.9)	0.037
Medical hx							
Smoker (%)	0 (0.0)	1 (12.5)	0 (0.0)	9 (11.4)	1 (8.3)	8 (10.3)	0.401
T2DM (%)	2 (13.3)	0 (0.0)	0 (0.0)	6 (7.6)	0 (0.0)	1 (1.3)	0.099
HTN (%)	2 (13.3)	1 (12.5)	4 (14.8)	16 (20.3)	0 (0.0)	14 (17.9)	0.629
Medical hx of Arthritis (%)	5 (33.3)	2 (25.0)	4 (14.8)	32 (40.5)	2 (16.7)	25 (32.1)	0.168
Medical hx of Rash (%)	6 (40.0)	3 (37.5)	5 (18.5)	20 (25.3)	2 (16.7)	9 (11.5)	0.074
Medical hx of Raynauds (%)	3 (20.0)	4 (50.0)	3 (11.1)	17 (21.5)	9 (75.0)	15 (19.2)	<0.001
Medical hx of Sicca (%)	1 (6.7)	2 (25.0)	18 (66.7)	14 (17.7)	1 (8.3)	12 (15.4)	<0.001
Medical hx of Nephritis (%)	0 (0.0)	0 (0.0)	1 (3.7)	12 (15.2)	0 (0.0)	1 (1.3)	0.006
Therapeutics							
Current glucocorticoid therapy (%)	7 (46.7)	4 (50.0)	4 (14.8)	28 (35.4)	3 (25.0)	11 (14.1)	NA
prednisolone dose (mean (SD))	4.10 (6.16)	4.69 (6.04)	2.50 (8.17)	3.22 (5.64)	2.08 (3.96)	1.36 (4.39)	0.22
current hydroxychloroquine (%)	5 (33.3)	5 (62.5)	10 (37.0)	46 (58.2)	2 (16.7)	34 (43.6)	0.038
current methotrexate (%)	2 (13.3)	1 (12.5)	1 (3.7)	10 (12.7)	1 (8.3)	20 (25.6)	0.082
current mycophenolate (%)	4 (26.7)	2 (25.0)	1 (3.7)	14 (17.7)	3 (25.0)	1 (1.3)	0.001
current azathioprine (%)	3 (20.0)	2 (25.0)	2 (7.4)	9 (11.4)	2 (16.7)	8 (10.3)	0.661
current rituximab (%)	2 (13.3)	3 (37.5)	2 (7.4)	10 (12.7)	0 (0.0)	1 (1.3)	0.003
current cyclophosphomide (%)	0 (0.0)	0 (0.0)	1 (3.7)	0 (0.0)	0 (0.0)	0 (0.0)	0.21
Laboratory							
dsDNA (%)	0 (0.0)	1 (12.5)	1 (3.7)	35 (44.3)	2 (16.7)	21 (26.9)	<0.001
Ro60 (%)	3 (20.0)	1 (12.5)	18 (66.7)	28 (35.4)	0 (0.0)	20 (25.6)	<0.001
R052 (%)	2 (13.3)	1 (12.5)	18 (66.7)	22 (27.8)	1 (8.3)	9 (11.5)	<0.001
La (%)	1 (6.7)	0 (0.0)	8 (29.6)	7 (8.9)	0 (0.0)	6 (7.7)	0.013
Sm (%)	0 (0.0)	4 (50.0)	0 (0.0)	11 (13.9)	0 (0.0)	2 (2.6)	<0.001
RNP (%)	0 (0.0)	6 (75.0)	0 (0.0)	8 (10.1)	0 (0.0)	5 (6.4)	<0.001
Scl70 (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (3.8)	0.358
Centromere (%)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.5)	3 (25.0)	2 (2.6)	0.001
Chromatin (%)	0 (0.0)	6 (75.0)	0 (0.0)	22 (27.8)	2 (16.7)	6 (7.7)	<0.001
B2GP (%)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.5)	0 (0.0)	0 (0.0)	0.612

Table 14: Baseline characteristics of PRECISESADS and DEFINITION cohorts

16.3.1.1 PRECISESADS

The PRECISESADS cohort comprised 1,313 ANA-associated RMD patients of which 876 contained extant clinical and biomarker data and were taken forward for analysis. Figure 4 demonstrates the methodology behind data exclusion.

PRECISESADS had a predominantly white European ancestry population. As expected, clinical and demographic characteristics differed between legacy diagnoses. SLE patients were significantly younger than other groups with a higher utilization of glucocorticoids (mean age 46.1, 52% on glucocorticoids, $p<0.001$ in both). There were significant differences in most clinical features measured (Table 14). Interferon score A and B were significantly higher in MCTD patients (1.10 and 1.01, $p<0.001$ in both). Plasmablast, Inflammation and Myeloid scores were highest in SLE patients ($p<0.001$, p,0.001 and $p = 0.001$ respectively).

16.3.1.2 DEFINITION

Within DEFINITION 219 patients with extant data were analysed. These comprised patients with IM, MCTD, pSS, SLE, SSc and UCTD. Significantly lower proportions of white European participants were identified in the MCTD and SLE cohorts (62.5% and 60.8% respectively, $p=0.04$). SSc and pSS patients were significantly older than their counterparts. There were no significant differences in Charlson comorbidity index (CCI), mean prednisolone dosage and smoking status between legacy diagnoses. Rituximab use was significantly greater in SLE patients ($p=0.003$), possibly due to commissioning and treatment guidelines. Patient reported pain VAS was significantly higher in UCTD and MCTD patients ($p=0.04$).

In terms of disease activity measurements, ESSDAI and overall BILAG-2004 index, did not differ significantly between groups. Rodnan score was highest in SSc patients (3.92, $p<0.001$) and physician global assessment was highest in IM and SSc patients (4.23 and 2.08, $p<0.001$). Composite patient-reported outcome measures did not differ significantly between legacy diagnostic groups apart from the EQ5D-5L index (highest in pSS, $p=0.02$).

Gene expression scores, non-inflammatory features, and long-term clinical outcomes adjusted for time in follow up did not differ significantly between diagnoses.

16.4 COVARIATE SELECTION

16.4.1 Key clinical covariates for ANA-RMD patients and clinicians

Key clinical features driving patient experience in ANA-RMDs were identified with the Leeds BRC PPIE group. Sicca, arthritis, high steroid requirement, and fatigue were highlighted.

However, fatigue was not well-represented in the datasets and was excluded from the final model. A survey of 10 physicians managing ANA-RMD patients identified Raynaud's, digital ulceration, skin rash, arthritis, and nephritis as key drivers of therapy escalation.

16.4.2 Key gene expression covariates from the literature

In order to characterise both innate and adaptive immune abnormality, drawing from a literature review of biomarkers and molecular reclassification, and on the use of gene expression scores to stratify response to therapy, we incorporated a set of previously validated gene expression scores(200,202,409–411). These gene expression scores are detailed in Table 6. The model included high creatinine, proteinuria, specific ENA positivity, and lymphopenia due to their impact on RMD prognosis. Non-white ancestry and age were also included.

16.5 TRAINING OF VARIATIONAL AUTOENCODER

Data compression to latent spaces of 4, 6, and 8 was tested using the described methodology. Clustering analysis and clinical review showed inferior performance in the 4 and 6 neuron groups. Entropy and relative cluster stability index plots were generated using k-means, PAM, spectral and hierarchical clustering on the 8-neuron latent layer using the M3C package (Appendix B and Appendix C). PAM, spectral and hierarchical clustering offered no greater utility over k-means.

16.6 RECLASSIFICATION VS LEGACY DIAGNOSIS

16.6.1 Heuristic determination of optimal cluster number

Heuristic analysis of uncompressed pooled data from PRECISESADS and DEFINITION using relative cluster stability index (RCSI) and entropy plots was conducted to determine the optimal clustering number. RCSI plots were indeterminate but plateaued at k values of 4-6. To select between these values, we explored the cross-sectional clinical features of the clusters derived from each k value (described below). From this we selected a k value of 5 as yielding optimal heuristic and clinical significance. RCSI, silhouette and entropy plots for the PRECISESADS dataset are shown in Figure 18 . Legacy diagnoses had significantly negative silhouettes in most groups, with an average width of -0.03. In contrast, the new deep learning derived classes showed positive values in 3 out of 5 groups, with an average silhouette width of 0.06. Given the heterogeneity of the cohort, the new classification was felt to represent a substantial improvement in stratification. Following the analysis of clinical characteristics (described below), these classes were named Sicca, Myeloinflammatory,

High non-inflammatory symptom burden, Inflammatory MSK and Polyinflammatory. We named these new classes ANA-RMD Classification Classes (ARC Classes).

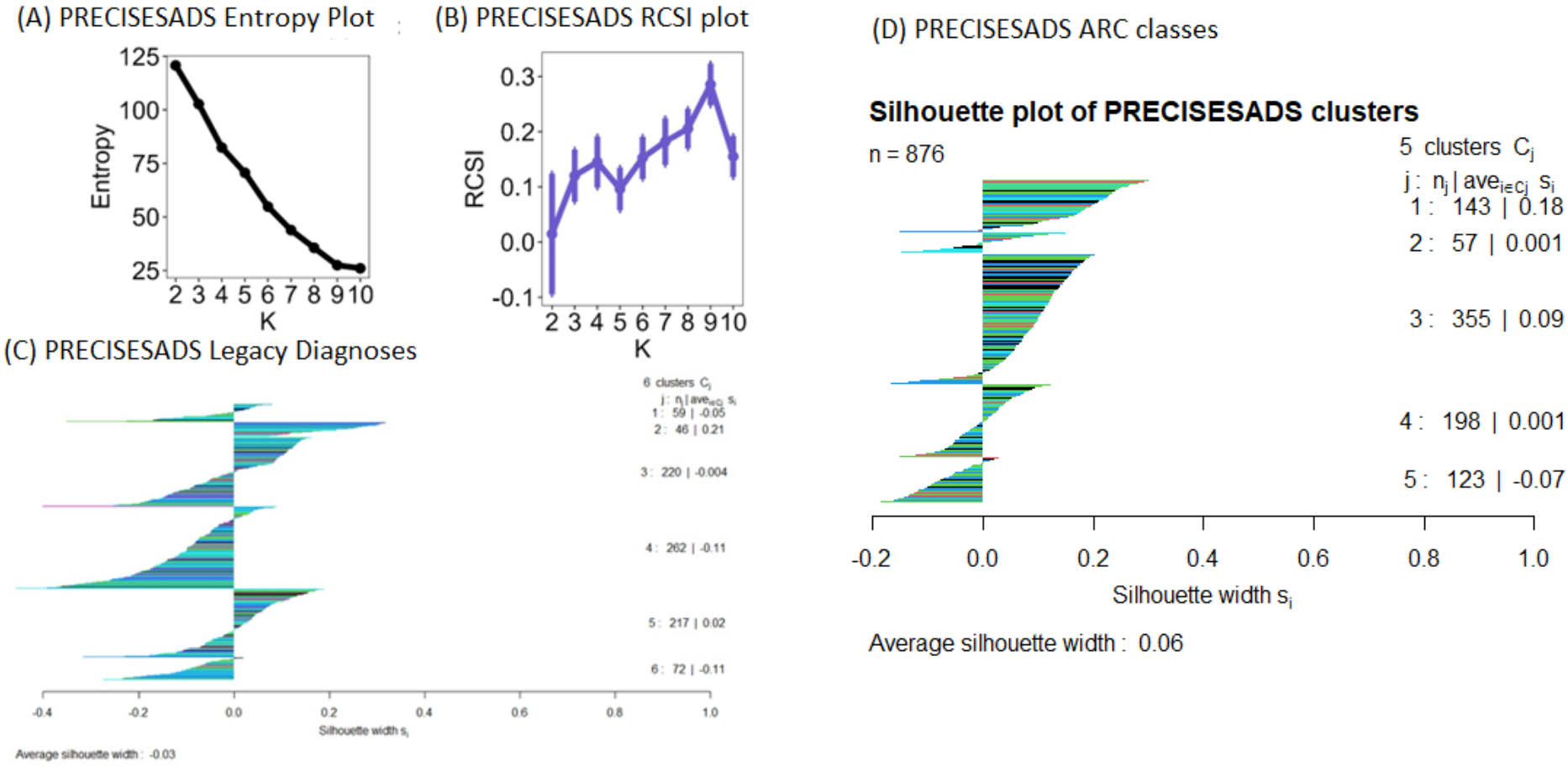


Figure 18: PRECISESADS silhouette, entropy and RCSI plots

16.6.2 Reclassification visualisation

The ability of legacy diagnoses and deep learning-based reclassification to sub-stratify covariates was visualized using t-SNE plots. As expected, more distinct clusters were observed in the deep learning-derived classes in both the PRECISESADS (Figure 19) and DEFINITION cohorts compared to the legacy diagnoses.

In order to analyse the effect of reclassification on each individual legacy diagnosis, we created PCA plots for the clustered data, stratified by individual diagnosis, overall legacy diagnoses, and ARC classes, color-coded by cluster number (Figure 20). Each legacy diagnosis contained individuals assigned to one of each ARC class except PAPs (lacking the Myeloinflammatory class) and MCTD (lacking the Sicca class). The Sicca class was most prominent in pSS, The High IFN High Disease Activity in SLE, The High non-inflammatory symptom burden class across SLE, SSc, pSS, and UCTD, while the Inflammatory MSK and Polyinflammatory classes were well represented across all legacy diagnoses. PCA plots by diagnosis and cluster show greater homogeneity by clustering, highlighting significant heterogeneity in existing diagnostic criteria

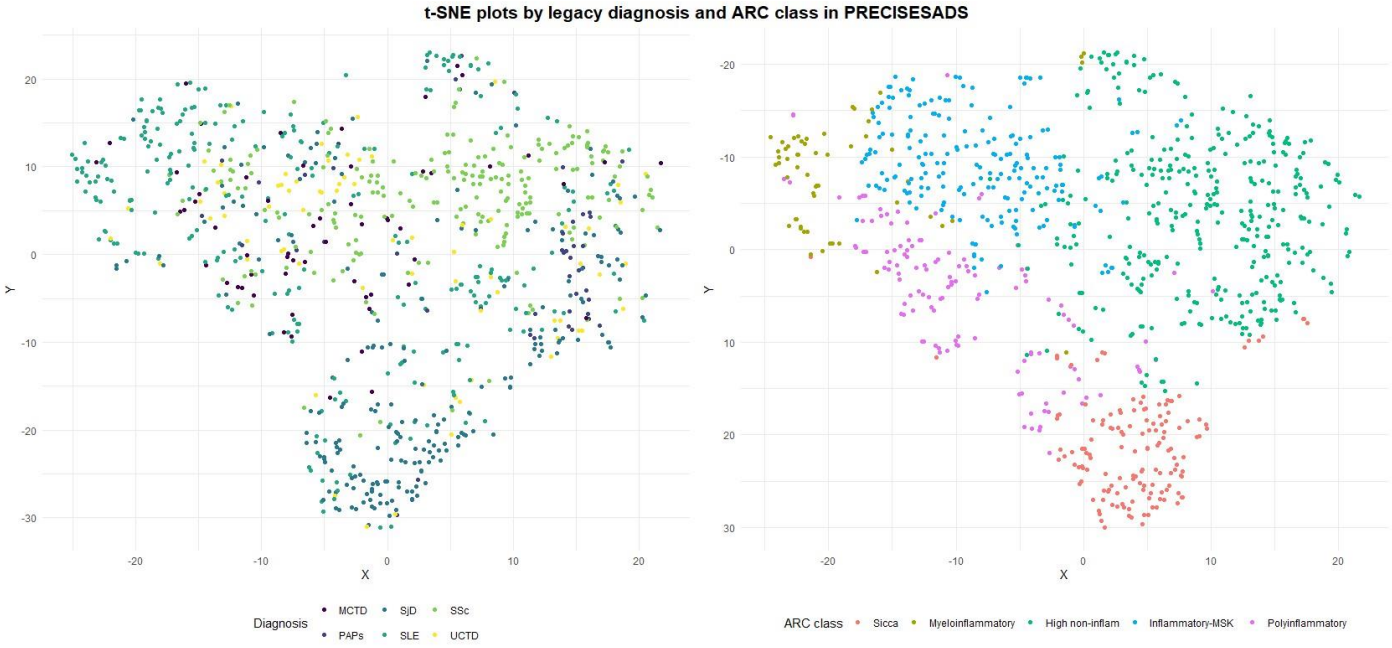


Figure 19: t-SNE plots showing stratification by deep learning classes and legacy diagnoses

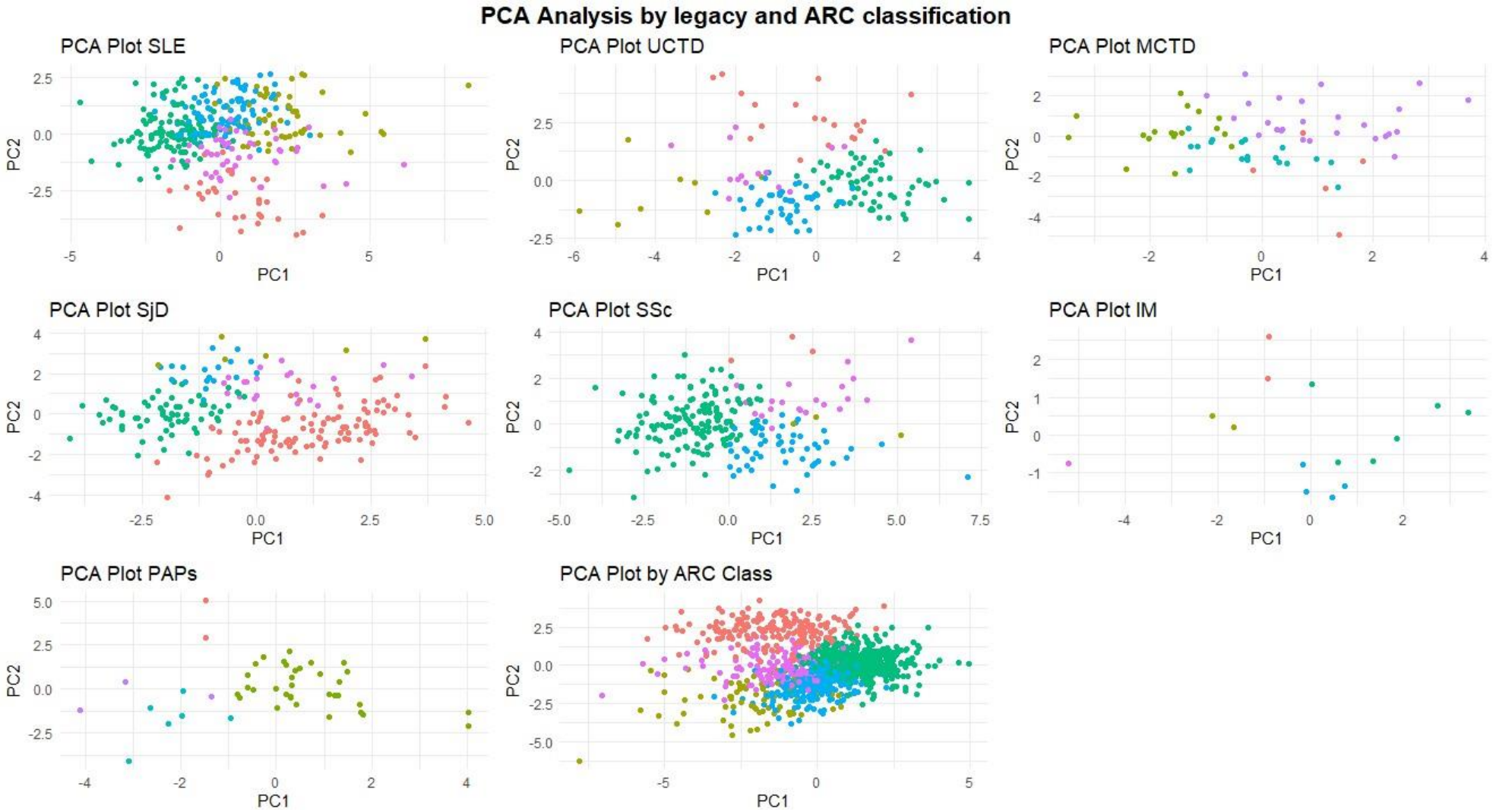


Figure 20: PCA plots stratified by individual diagnoses, and ARC class

16.6.3 Disease phenotype of ARC classes vs legacy diagnoses

ARC classes showed significant differences in patient-reported outcomes, clinical phenotypes, gene expression and flow cytometric patterns.

16.6.4 Clinical, gene expression and flow cytometric phenotype

Figure 21 presents circular bar charts comparing the clinical, biomarker, PRO, and long-term outcome phenotypes within DEFINITION for ARC and legacy diagnoses. Legacy diagnoses showed poor correlation with disease activity, except for the Rodnan skin score, which was highest in the SSc group. In contrast, the ARC classes were better descriptors of disease activity.

16.6.4.1 Sicca

This comprised 17% of the pooled cohorts and was characterised by immunological activity with a high plasmablast score and moderate to high activity across other gene expression scores. Clinically there was a high prevalence of Sicca and a high non inflammatory burden. Glucocorticoid use was low. Impacts on patients, healthcare system and society were moderate with intermediate PRO impact and low healthcare utilization.

16.6.4.2 High non-inflammatory symptom burden

This comprised 42% of the pooled cohorts and was characterised by immunological inactivity with low gene expression and biological markers of disease activity. Clinically, they demonstrated a high rate of non-inflammatory symptoms with low levels of inflammatory disease manifestations. Glucocorticoid use was low. Patients in this class reported a high burden of disease and had high rates of outpatient department attendance but low hospitalization rates.

16.6.4.3 Inflammatory-MSK

This comprised 22% of the pooled cohorts and was characterised by a high prevalence of arthritis, high patient reported pain scores with a low prevalence of fibromyalgia. Glucocorticoid use was moderate, and perceived capability scores were poor.

16.6.4.4 Myeloinflammatory

This comprised 7% of the pooled cohorts and was defined by high myeloid and inflammatory gene signatures with moderate IFN and plasmablast scores. Lymphopenia incidence was high and there was high prevalence of solid organ involvement with high physician-assessed disease activity. Long-term glucocorticoid use was greatest in this group and, despite low patient-reported symptoms, healthcare utilisation was high.

16.6.4.5 Polyinflammatory

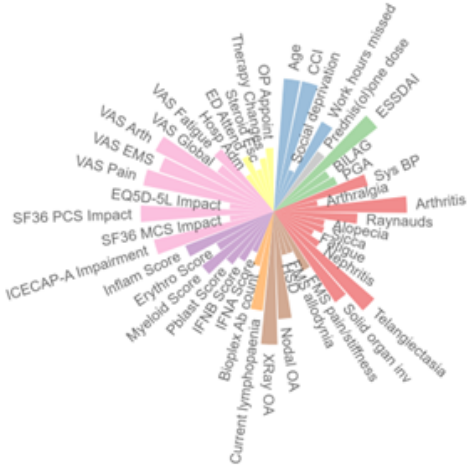
This comprised 12% of the pooled cohorts and was characterised by globally high gene expression scores and immunological activity. They also showed high clinical symptom prevalence and moderate glucocorticoid requirements. In terms of patient impact, patient reported outcomes were significantly impaired however there was relatively low healthcare resource use within this group.

Sicca

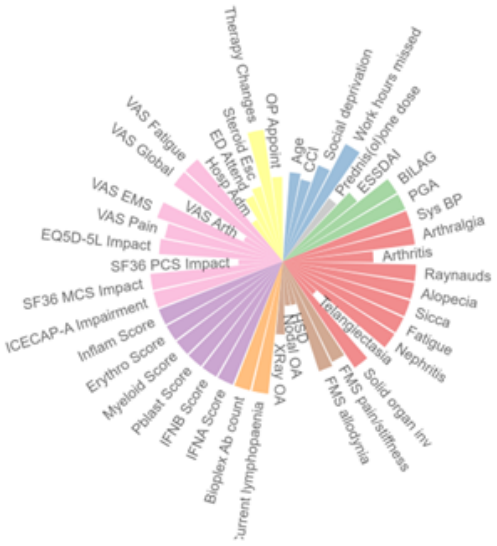
Myeloinflammatory

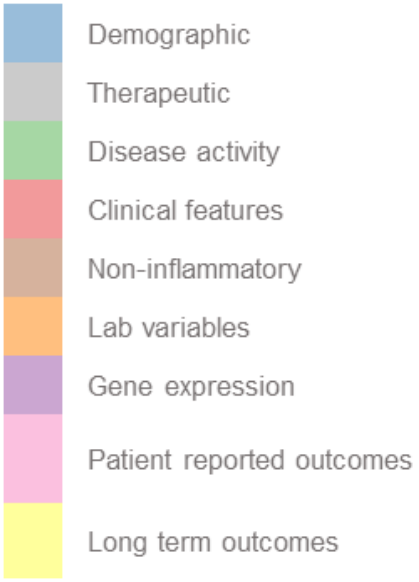
High non-inflammatory symptom burden

Inflammatory MSK



Polyinflammatory





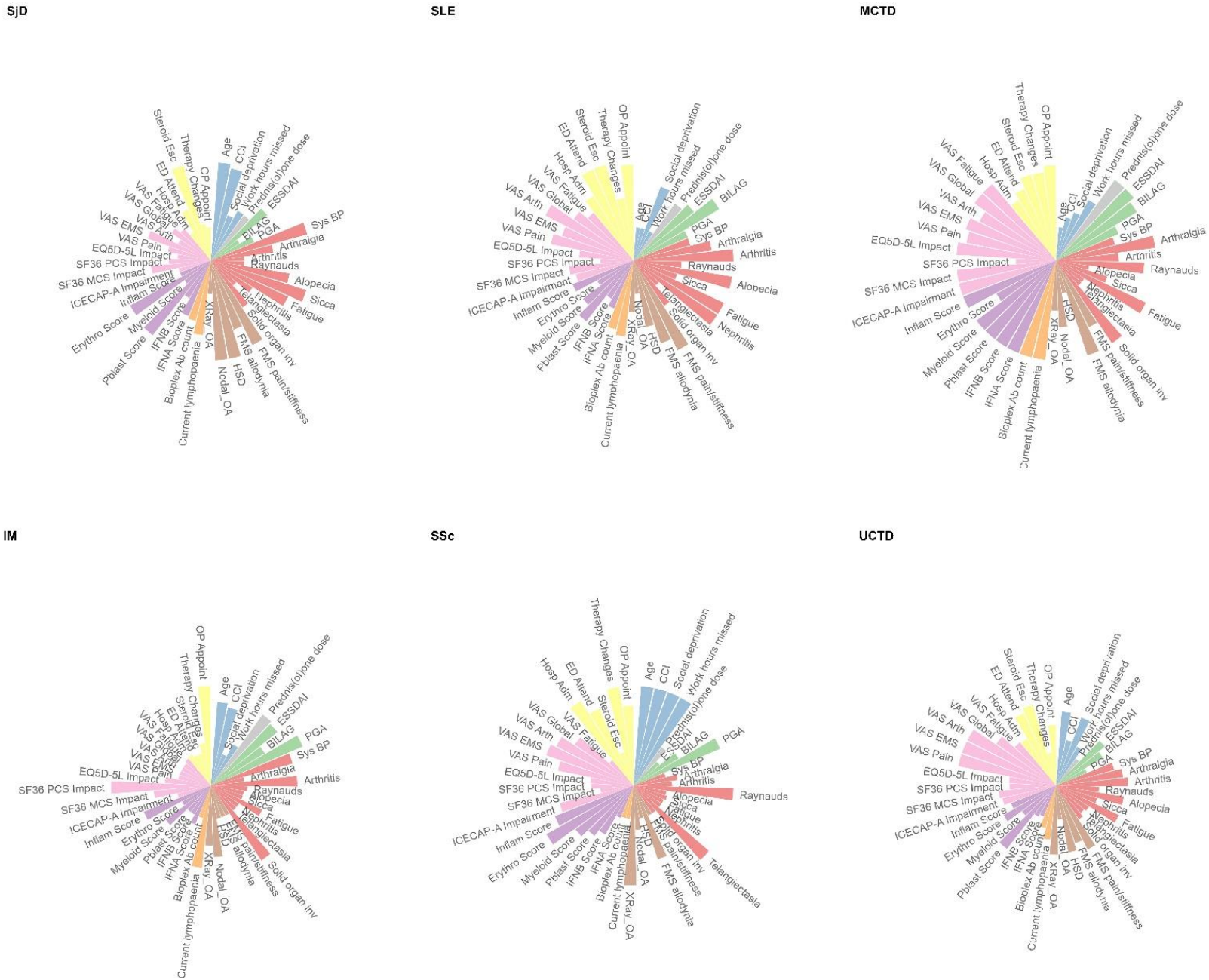


Figure 21: Summary circular bar charts of clinical, disease activity, gene expression score, PRO and long term outcome phenotypes by legacy diagnosis and ARC class within DEFINITION

16.6.5 Gene expression and flow cytometric comparisons

Figure 21 also summarises several key gene expression score patterns within the ARC classes. Gene expression activity was shown to poorly correlate with legacy diagnoses but shows stronger correlation with ARC classes.

A Polyinflammatory cluster, two interferonopathy clusters (Sicca and High IFN, High disease activity) and a low gene expression activity cluster (High non-inflammatory symptom burden) were observed independently in both the PRECISESADS and DEFINITION cohorts. Summary statistics showed significant differences between gene expression scores, classified by ARC clusters in both the PRECISESADS and DEFINITION cohorts ($p<0.05$ in all scores).

A low gene expression activity class was noted in both PRECISESADS and DEFINITION characterized by lower IFN, myeloid, erythroid and inflammation gene expression scores.

Flow cytometric analysis showed significantly higher classical monocyte populations in the SSc and IM groups, lowest in the SLE group ($p=0.016$). Tetherin expression on memory B-cells was highest in the SLE group (PEA 54317.08, $p=0.008$). In the ARC classes, classical monocyte populations were highest, though not significantly, in the MSK pain class ($p=0.094$). Tetherin expression on memory B-cells was greatest in the Sicca and Polyinflammatory clusters, aligning with their high interferon gene score expression.

16.6.6 Cross sectional and long-term clinical outcomes

Legacy diagnoses poorly predicted long-term outcomes in the DEFINITION dataset. There were no significant associations between legacy diagnoses and pro-rata hospital admissions, emergency department visits, escalation of steroid therapy, or all-cause mortality.

In contrast classification by ARC class was more predictive of long-term outcomes. Myeloinflammatory patients had significantly higher rates of hospital admission and emergency department attendance than other classes (0.87 and 1.05 per year of follow up, $p= 0.006$ and 0.003, respectively). Polyinflammatory patients exhibited the second-highest rate of emergency department attendance (0.49 per year) and the highest rate of therapy change (0.5, $p=0.327$).

16.7 INCLUSIVITY OF CLASSIFICATION

A Sankey plot illustrating how the combined patients from DEFINITION and PRECISESADS were stratified according to the new classification system is shown in Figure 22 Legacy diagnostic criteria leave 14% of patients unclassified (UCTD).

Importantly SLE and UCTD patients were divided across all ARC classes. Primary Sjogren's syndrome patients contributed to the Sicca group but a large proportion were classified into other ARC classes. PAPs patients were largely classified into the High non-inflammatory symptom burden class, possibly due to the distinct nature of their clinical features. Many SSc patients were categorised in the High non-inflammatory symptom burden class, which may reflect their individual state of disease rather than an overall mild phenotype Finally MCTD was noted to contribute the highest proportion of people to the Polyinflammatory group and the lowest proportion to the High non-inflammatory symptom burden class.

Finally, ARC classification was able to provide a diagnostic label to 100% of sampled patients. Most legacy diagnoses were represented within each ARC class. PAPs patients were largely categorised into one ARC class and given their distinct pathological features it may be preferable to retain their original classification, though patients with secondary APLS may benefit from reclassification

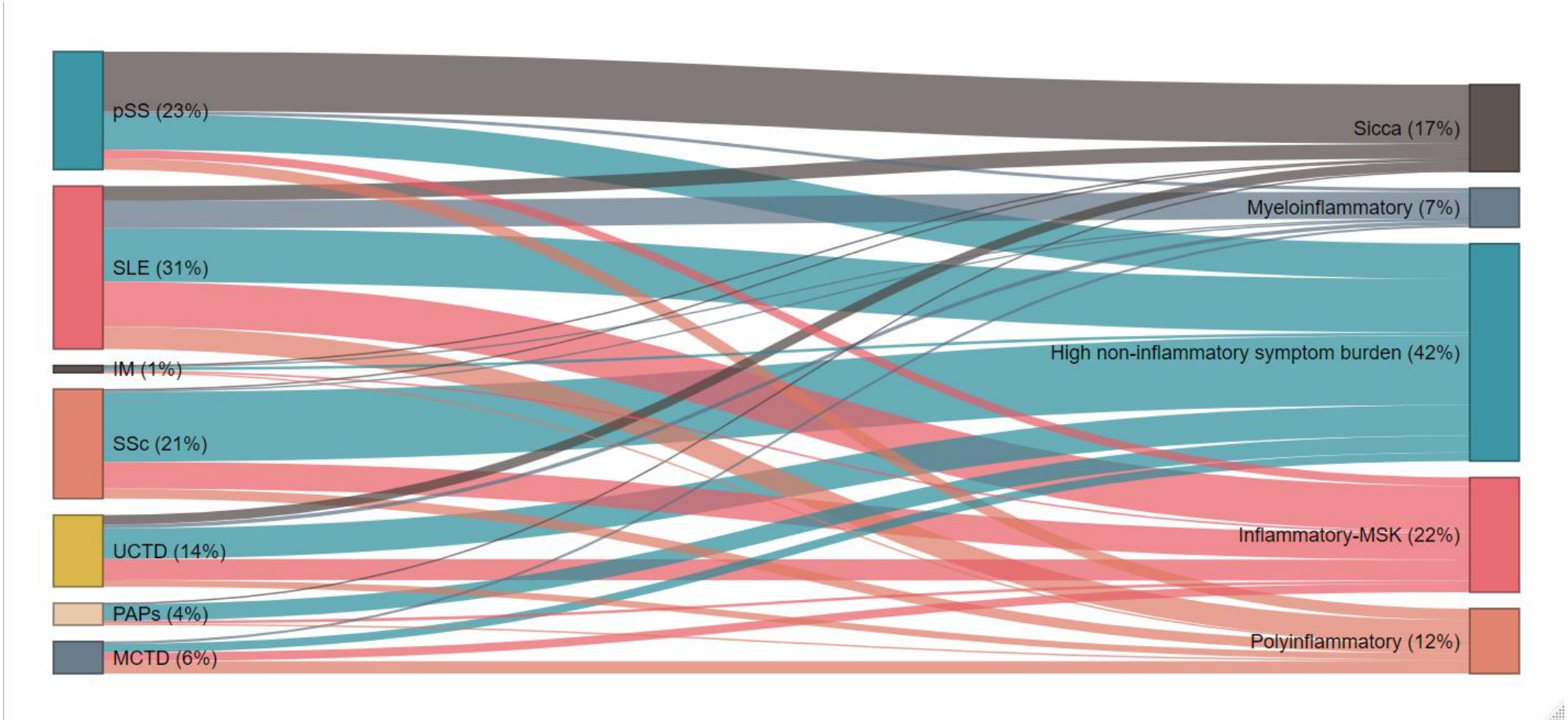


Figure 22: Sankey plot showing sub-stratification of ANA-RMD patients by ML methodology

16.7.1 Comparison with traditional ANA-RMD trial design

ARC classes were compared to legacy criteria in the context of trial populations. The EXPLORER trial in SLE recruited patients on conventional immunosuppression with moderate to severe active SLE characterized by 1 BILAG-A or 2 BILAG-B domains(272) SLE BRAVE II recruited patients with SLE, ANA/dsDNA/Sm positivity, SLEDAI>6 and 1 BILAG-A or 2 BILAG-B domains(379). Both SLE-BRAVE II and EXPLORER failed to meet their primary endpoints. The CVAY736A2201 study of Ianalumab in pSS recruited patients with pSS, Ro or SSA positivity and an ESSDAI ≥6 and a recent study of Abatacept therapy in primary Sjogren's syndrome(412). Table 15 summarises the suitability of legacy diagnosis patients within DEFINITION for clinical trials and the distribution of these patients across ARC classes.

Trial	Legacy criteria	% of legacy criteria eligible in DEFINITION	ARC Class				
			Sicca	High non inflammatory symptom burden	Inflammatory MSK	Polyinflammatory	Myeloinflammatory
EXPLORER	SLE	19/79 (24.1%)	1 (5.3%)	4 (21.1%)	8 (42.1%)	2 (10.5%)	4 (21.1%)
SLE BRAVE II	SLE	15/79 (19.0%)	1 (6.7%)	3 (20%)	7 (46.7%)	2 (13.3%)	2 (13.3%
Ianalumab in pSS (CVAY736A2201)	pSS	4/27 (14.8%)	2/4 (50%)	1/4 (25%)	0	0	1 (25%)
Abatacept in pSS (NCT02915159)	pSS	4/27 (14.8%)	2/4 (50%)	1/4 (25%)	0	0	1 (25%)

Table 15: Evaluation of trial suitability by legacy diagnosis and ARC class

16.8 DISCUSSION

Here we present a comprehensive reclassification of ANA-RMDs that incorporates demographic, clinical, antibody, therapeutic and molecular data with characterisation of the new classes using patient-reported outcomes, long term health outcomes and flow cytometric data. Our work benefits from two large multidisease cohorts with rich clinical and immunological data and deep learning computational methods. This reclassification may enable more effective classification and prognostication, better clinical trial design and licensing of new therapies and greater availability of appropriate treatments to larger group of patients.

The persistence of certain disease phenotypes, such as the Sicca class, through reclassification suggests that completely discarding existing diagnostic criteria isn't desirable. This is reinforced by the dominant classification of PAPs patients into the High non-inflammatory symptom burden class, indicating their distinct pathology from the rest of the ANA-RMD spectrum despite low disease activity. However, many non-Sjogren's patients were included in the Sicca class, indicating significant cross-diagnostic overlap in gene expression, lab, and clinical features. In contrast, SLE, IM, UCTD, and non-glandular pSS showed significant heterogeneity and were represented in multiple ARC classes, as shown in Figure 22.

A crucial problem in ANA-RMD trial design is recruitment, particularly for rarer ANA-RMDs like IM and MCTD. Composite endpoints based on multisystem disease measures have led to the failure of several high-profile SLE trials. In EXPLORER, this resulted in a negative trial for Rituximab, despite its known effectiveness in refractory SLE. Similarly, in SLE-BRAVE II, baricitinib failed to produce a positive result despite a promising phase 1 trial.

Our comparison of trial inclusion by ARC criteria to legacy diagnosis is important because SLE patients, for example, are represented in all five ARC classes. About 24% of DEFINITION participants with SLE would have qualified for EXPLORER, yet this cohort includes phenotypically distinct new classes. Similar patterns are observed with the inclusion criteria for SLE-BRAVE II, Abatacept, and Ianalumab in pSS. Targeted biologic therapies in the UK are currently available only to SLE (Rituximab and belimumab) and inflammatory myositis (Abatacept) patients, comprising 32% of this cohort. Therefore given phenotypical and pathophysiological overlap across ANA-RMDs it is desirable to open up access to the remaining 68% of patients.

This suggests that hybrid reclassification may more effectively provide homogeneous cohorts for clinical trials, unlike legacy diagnosis, which recruits a heterogeneous population.

Notably, the inclusion of a large proportion of patients from our 'High non-inflammatory symptom burden' cluster (20-25% of recruitment across all four trials) may drive high placebo response rates. These patients show biological inactivity, with low gene expression scores, high lymphocyte counts, and higher rates of physician-assessed allodynia. Using this data-driven agnostic approach we define two classes (Polyinflammatory and high disease and clinical activity interferonopathy) that display disproportionately high healthcare utilization and have common pathophysiology at a gene expression level. These clusters should be evaluated as possible baskets for future clinical trials. The incorporation of gene expression scores into these ARC classes adds objectivity to clinical trial selection. Work within the MASTERPLANS consortium has previously showed an association between certain gene expression patterns and differential response to biological therapies. Indeed, these patterns were also shown to be heavily influenced by ancestry-specific factors(203).

We also identify a High Non-Inflammatory Symptom Burden group, representing 42% of the combined cohorts, who are not adequately addressed by existing treatment options. These patients might experience better clinical outcomes with a focus on rehabilitative and non-immunosuppressive therapies, which could have a more significant impact on their condition. Furthermore, considering the non-inflammatory characteristics of their disease, avoiding immunosuppression both within and outside of clinical trials could prevent harmful side effects and potentially lower placebo response rates in clinical trials. This novel method of reclassification can be applied across other medical specialties and is entirely agnostic in its approach. Reclassification in this way avoids nominative discrimination associated with legacy diagnostic criteria. Our work demonstrates a substantial contingent of UCTD patients within these high disease impact groups, who currently have no licensed treatments, but high clinical activity.

Further work could apply this data-compression and reclassification approach to similar questions in other medical specialties. The Polyinflammatory and Myeloinflammatory classes should be evaluated for basket trials for more equitable trial design. Limitations of this work are its inability to use variables not conserved across datasets, particularly PRECISESADS extensive gene expression dataset and the DEFINITION patient-reported outcome library and the potential impact of the COVID-19 pandemic on the 5 year follow up in this trial. Inflammatory myositis patients were also underrepresented within DEFINITION and absent in PRECISESADS.

A clear limitation of this approach are the relatively low silhouette scores for the ARC classes. While these scores outperform those of legacy diagnoses, they remain modest compared to many other data science applications. This is expected, as ANA-RMDs

represent a complex and overlapping disease family, making the definition of clear clusters challenging. However, the new ARC classes offer potential therapeutic targets based on their gene expression scores, shared clinical features, and, in one class, their non-inflammatory pathology. Additionally, this approach successfully classifies 100% of ANA-RMD patients, a substantial improvement over legacy diagnoses. In summary, deep-learning approaches are agnostic and able to accommodate multiple different data types and results in more inclusive classification criteria with differential long-term outcomes that are replicable across cohorts with different ethnic makeups.

17 RESULTS 4: PRELIMINARY EXPLORATION OF CAUSAL AND PREDICTIVE MODELS USING GENE EXPRESSION FOR LONGITUDINAL DISEASE ACTIVITY ACROSS ANA-RMDS

17.1 INTRODUCTION

The previous chapters in this thesis have all analysed cross-sectional data. Here, I make use of the longitudinal data in DEFINTION Cohort C.

From the previous chapters, as well as published literature, it is clear that within the spectrum of ANA-associated rheumatic musculoskeletal diseases (ANA-RMDs) there are shared gene expression signatures that span multiple legacy diagnoses(154,371). These are mirrored by overlapping clinical presentations and shared therapeutic targets. The efficacy of medications such as the interferon blocking drug Anifrolumab is established in SLE and is under investigation in primary Sjogren's syndrome and Systemic Sclerosis due to a shared overactivation of the type 1 interferon (IFN-I) pathway(413,414). The utility of IFN-I assays in clinical practice was recently assessed by a EULAR taskforce which concluded that they have high potential for implementation in the clinical management of RMDs both within and beyond rheumatology(203) Therefore, given their shared biomarker profiles it appears logical to explore the possibility of gene signature-based baskets as predictors of therapeutic response and longer-term disease outcomes. Ancestry associated gene expression patterns have already been identified in the MASTERPLANS cohort which correlated with differential responses to rituximab in SLE(203).

Globally there is a move toward a molecular taxonomy of disease, both within rheumatology and medicine as a whole(154,378). This approach is well established within oncology, and in the UK, biomarker driven treatment strategies have been in use for some time. The centre for Genomics and Oncological research (GENYO) at the University of Granada has been instrumental in defining a molecular taxonomy of disease within rheumatology and oncology(262,415). Their PRECISESADS cohort comprises 2500 ANA-RMD and rheumatoid arthritis patients with healthy controls. Previous work in this cohort identified three active gene expression signature clusters (Interferon, lymphoid, inflammatory) and a fourth characterised by globally low gene expression and low associated disease activity which included multiple healthy controls(262). The scope of this approach has been expanded further with the 3TR project which aims to combine and analyse multi-omic data from autoimmune, allergic and inflammatory conditions to look for shared molecular and microbiome profiles spanning existing diagnoses. Prospective work as part of the 3TR

project has expanded this molecular taxonomy of disease approach beyond a single medical specialty, performing cross-diagnostic analysis of 7 key diseases (Chronic obstructive pulmonary disease, asthma, Crohn's disease, ulcerative colitis, rheumatoid arthritis, multiple sclerosis and systemic lupus erythematosus)(416,417).

There is interest in understanding the relationship between gene signatures such as IFN-I pathway activation from two perspectives. From the point of view of prediction, identifying patients at high risk of flare can allow appropriate increased therapy, or glucocorticoid tapering. This is particularly important in severe diseases that require more intensive therapy such as lupus nephritis. This can be achieved using a predictive multivariable model.

Another perspective is to understand the causal relationship between gene signatures and longitudinal disease activity. This is important to understand the mechanism of inflammation and whether certain therapeutic targets (e.g. blocking IFN-I) would be appropriate. Historically, multivariable model outputs have been used to attempt such inferences by analysing the odds ratios of individual variables. However, it is now recognised that this introduces the possibility of "table two fallacy" whereby the effects of individual variables in a multivariable model are compared to look for independent predictors and infer relationships between them. This is problematic as it leads to the incorrect attribution of causation to these odds ratios. Including multiple variables without understanding the relationship between them is acceptable for predictive modelling but may be very misleading if causality is then inferred. For example, if an exposure is adjusted for a second variable that lies in the causal pathway to the outcome of interest, then it's true effect may be negated or even inverted. This was famously seen in the OpenSAFELY study, published in *Nature*, which included both smoking and chronic lung disease as variables in a multivariable model to predict severe Covid outcomes(418,419). The authors presented a paper with multiple adjusted predictors from a multivariate model. Unfortunately, many readers misconstrued this presentation of the data and came to the false conclusion that smoking was protective against severe Covid when in fact the reverse is true. Causal inference approaches such as Directed Acyclic Graphs (DAGs) in the construction of multivariable models is one approach to both reduce the incidence of this error and open covariate selection within these models to academic scrutiny(420).

Within Leeds the multi disease DEFINITION cohort C contains 124 of the 300 ANA-RMD patients with rich clinical and biomarker phenotyping were followed up sequentially at baseline, 3 months and 6 months. This longitudinal follow up data, which is currently lacking in this field, may demonstrate how dynamic the IFN-I and other pathways are, as well as their relationship to disease activity across ANA-RMDs.

In this chapter, I explore both approaches within two subchapters; first the potential value of gene expression scores in a predictive model, and second explore a causal model for their role in causing flares. These have potential value in identifying when therapy should be used, and what class of therapy might be appropriate respectively. These analyses are preliminary and exploratory. Robust predictive modelling requires validation processes and causal modelling for these questions may be complex including mediation effects. We considered such complex work beyond the scope of this thesis. Lastly, I explore changes in gene expression scores over time in patients receiving conventional therapy, which may help to further understand whether relevant disease processes are effectively inhibited.

For clarity the results of this chapter are presented as overall cohort findings then as two results subchapters detailing the results of the causal inference informed approach and the multivariable predictive modelling respectively.

17.2 Hypotheses, Aims & Objectives

17.2.1 Hypothesis

In ANA-RMD patients:

1. Specific gene expression signatures are positively correlated with increased disease flare rates in longitudinal data.
2. Baseline gene expression signatures are predictive of subsequent disease flare at 3 and 6 month follow up after accounting for other known predictors
3. Baseline gene expression signatures have a causal role in flare
4. Initiation of immunosuppression is associated with gene expression changes

17.2.2 Aims

1. To assess the causal relationship between IFN Score and follow up disease activity across ANA-RMDs
2. To assess predictive models for follow up disease activity including IFN Score and other gene expression scores across ANA-RMDs
3. To assess the impact of conventional immunosuppression and glucocorticoid use on longitudinal IFN Score and other gene expression scores.

17.2.3 Objectives

1. To use DAG analysis to identify measured and unmeasured factors governing the impact of key gene expression scores on disease activity and flare then construct a causal model
2. Evaluate the explanatory power of this model using linear mixed effects modelling

3. To construct a predictive model including demographic and clinical variables, IFN and other gene expression scores to predict 3 and 6 month disease activity
4. To assess the responsiveness of specific gene expression scores to changes in therapeutics and glucocorticoid exposure using LMER.

17.3 METHODS

17.3.1 The DEFINITION cohort

DEFINITION is a prospective, multi-disease ANA-RMD cohort study in Leeds, United Kingdom since May 2017. The primary aim was to better define the role of IFN-I and other biomarkers in ANA-RMD so that the use of IFN-I targeted as well as conventional or other targeted therapies can be refined. DEFINITION comprises the extensive dataset discussed previously within this thesis including several validated gene expression scores as shown in Table 6 and the demographic, patient-reported and clinical variables detailed in the general methods section..

17.3.2 Demographics and comorbidity

We collected baseline age, gender, patient-identified ancestry, smoking status, index of multiple deprivation (IMD)(384), Charlson Comorbidity index (CCI)(385), clinical features of fibromyalgia, hypermobility and osteoarthritis. Diagnoses were recorded according to consultant physician review. This was preferred over diagnostic criteria which are absent in UCTD and often unfulfilled in pSS without tissue biopsy.

17.3.3 Laboratory measures

Full blood count, complement C3 and C4 levels and ANA subtypes, including anti-double-stranded DNA, Ro-52, Ro-60, La, Sm, SM/RNP, RNP, Scl-70, Jo-1, Centromere, Chromatin and Ribosomal-P antibodies (Bioplex multiplex analyser) were measured in a routine diagnostic laboratory. PBMC subsets were analysed using 8-colour flow cytometry as a proportion of total PBMC count (T-cells (CD3+CD56-), NK-cells (CD56+), NKT-cells (CD3+CD56+), Memory B-cells (CD19+CD27+, Plasmablasts (CD19+/-CD27+CD38++), classical monocytes (CD14++CD16-) intermediate monocytes (CD14++CD16+)and non-classical monocytes (CD14+CD16+)). Tetherin (CD317) mean fluorescence intensity was quantified on each cell subset, with memory B-cell level as the primary biomarker(168).

Two validated interferon-stimulated gene expression scores (IFN Score-A and IFN Score-B) were analysed. PBMC were separated using density gradient method (Lymphoprep; Alere-Technologies, Norway) from EDTA-anticoagulated blood. Total RNA purification kit (Norgen-Biotek, Canada) was used followed by quantitative real-time reverse transcriptase-PCR

(qRT-PCR) using TaqMan assays (Applied Biosystems, Invitrogen) for the selected 30 ISGs as previously described(386,387). Scores for genes annotated to plasmablast, myeloid lineage, inflammation and erythropoiesis function were included from previously described modules, defined based on the known molecular function of the constituent genes(381).

17.3.4 Clinical assessment

Disease activity was assessed at baseline and follow up using validated instruments applied to all diagnostic groups: EULAR Sjögren's syndrome disease activity index (ESSDAI); British Isles Lupus Assessment Group (BILAG) 2004 index; SLEDAI-2K.Rodnan skin score and MITAX were collected but not analysed due to limited relevance to clinical features and patient numbers.

Patients who were in clinical disease flare at the time of their appointment were defined by the presence of BILAG A/B category disease, MITAX A/B disease or ESSDAI moderate to severe disease activity. Information on therapeutic changes including new oral immunosuppressive agents, dose escalation, antimalarial and glucocorticoid therapy was collected at each visit.

17.3.5 Patient-reported outcomes

Patient-reported disease impact was assessed at each appointment using the following: 36-item Short Form Survey (SF36) – Composite and domain scores; EuroQol-5 Dimension 5-level Score (EQ5D-5L) – Index and domain scores; ICECAP-A; Patient-reported visual analogue scales (VAS); Arthritis-VAS; Pain-VAS; Global-VAS; Fatigue-VAS; Global health-VAS and Early Morning Stiffness-VAS.

17.3.6 Statistical analysis and mixed effects modelling

Statistical analysis and data visualisation utilised the heatmap, corrplot(v0.92), ggplot2(v3.5.1) and tableone(v0.13.2) packages.in R version 4.1.2(421–424). Initial multivariable analysis was carried out using the glm function from the stats package in R(425). Multiple group comparison employed Kruskal-Wallis testing, while twin group comparisons of categorical and continuous variables utilized Chi-square and T-tests, respectively. Directed acyclic graphs were created using the Dagitty package in R (v3.1)(359). Mediation analysis was carried out using the Mediation package in R(v4.5.0)(363).

Linear mixed effects modelling incorporating fixed and random effects was carried out using the lme4 package in R(v1.1-35.5)(426). Variance inflation factor (VIF) calculations used the car R(v3.1-2) package (427). 6 gene expression scores (IFN Score A and B, Plasmablast, Myeloid, Erythropoiesis and Inflammation scores), age and Charlson Comorbidity Index

(CCI) were modelled using a linear effect. Anti Ro antibody positivity (Ro52 or Ro60), Anti-RNP/Sm/SmRNP positivity, baseline glucocorticoid dose escalation, new hydroxychloroquine therapy at baseline and ethnicity (Characterised as European yes/no due to low sample sizes) were modelled using binary outcomes.

17.4 SHARED RESULTS

17.4.1 Baseline variables

Key baseline variables of DEFINTION Cohort C are summarized in Table 16 . The overall sample of 124 ANA-RMD patients contained: SLE n=46, UCTD n=49, pSS n=16, IM n=7, SSc n=4, MTCD n=2. Charlson comorbidity index (CCI) was highest in the SSc patients (4.0, $p<0.02$) and Physician global assessment (PGA) was highest in IM patients (6.21, $p<0.02$). There were no significant differences in gene expression scores between diagnoses. At baseline we identified 56 who were in clinical disease flare at the time of their appointment as defined by the presence of BILAG A/B category disease, MITAX A/B disease or ESSDAI moderate to severe disease activity.

Follow up visits were conducted at 3 (n=58) and 6 months (n=68). During these visits clinical and therapeutic use information was collected along with laboratory, clinical and disease activity variables. Follow up was incomplete due to patients not attending their appointments. This was particularly impacted by the Covid pandemic at the time of this study. For this reason, numbers of patients with each diagnosis differs across time points. At 3 months, 31 (53.4%) and 6 months, 25 (36.8%) of patients were in disease flare. PGA scores were significantly lower at 3 and 6 month follow up with respect to baseline (BL 3.38, 3mth 2.28 and 6mth 2.77, $p<0.02$) consistent with the benefit of therapies prescribed at baseline to flaring patients.

Characteristic	IM	MCTD	pSS	SLE	SSc	UCTD	p
n	7	2	16	46	4	49	
Demographics							
Age (mean (SD))	50.0 (17.2)	59.0 (14.1)	48.6 (14.1)	43.5 (12.1)	58.5 (4.9)	48.3 (13.1)	0.259
Male gender(%)	4 (57.1)	0 (0.0)	1 (6.2)	8 (17.4)	1 (25.0)	7 (14.3)	0.070
Ethnicity (%)							0.813
Asian	0 (0.0)	0 (0.0)	2 (12.5)	9 (20.5)	1 (25.0)	4 (8.9)	
Black	0 (0.0)	0 (0.0)	0 (0.0)	4 (9.1)	0 (0.0)	7 (15.6)	
Other/Mixed	0 (0.0)	0 (0.0)	1 (6.2)	3 (6.8)	0 (0.0)	4 (8.9)	
Unknown	0 (0.0)	0 (0.0)	2 (12.5)	2 (4.5)	0 (0.0)	4 (8.9)	
White	7 (100.0)	2 (100.0)	11 (68.8)	26 (59.1)	3 (75.0)	26 (57.8)	
Index of Multiple Deprivation Rank (mean (SD))	13786 (12974)	16650 (NA)	18177 (10871)	13092 (10096)	6437 (NA)	12790 (9582)	NA
Charlson Comorbidity Index (mean (SD))	2.71 (2.36)	4.00 (4.24)	2.38 (1.15)	1.80 (1.22)	4.00 (3.46)	1.88 (1.33)	0.017
Laboratory							
dsDNA (%)	0 (0.0)	0 (0.0)	1 (6.2)	19 (41.3)	1 (25.0)	14 (36.8)	0.041

Characteristic	IM	MCTD	pSS	SLE	SSc	UCTD	p
Ro60 (%)	1 (14.3)	0 (0.0)	13 (81.2)	21 (45.7)	0 (0.0)	10 (26.3)	0.001
Ro52 (%)	2 (28.6)	0 (0.0)	10 (62.5)	13 (28.3)	0 (0.0)	2 (5.3)	<0.001
La (%)	0 (0.0)	0 (0.0)	5 (31.2)	7 (15.2)	0 (0.0)	0 (0.0)	0.014
Sm (%)	0 (0.0)	0 (0.0)	0 (0.0)	11 (23.9)	0 (0.0)	1 (2.6)	0.012
SmRNP(%)	0 (0.0)	1 (50.0)	0 (0.0)	14 (30.4)	0 (0.0)	5 (13.2)	0.024
RNP (%)	0 (0.0)	1 (50.0)	0 (0.0)	7 (15.2)	0 (0.0)	1 (2.6)	0.036
Jo1 (%)	2 (28.6)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	<0.001
Centromere (%)	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.2)	1 (25.0)	2 (5.3)	0.229
Chromatin (%)	0 (0.0)	1 (50.0)	0 (0.0)	17 (37.0)	0 (0.0)	5 (13.2)	0.004
Ribosomal P (%)	0 (0.0)	0 (0.0)	0 (0.0)	4 (8.7)	0 (0.0)	1 (2.6)	0.606
Lupus anticoagulant (%)	0 (0.0)	0 (0.0)	1 (6.2)	0 (0.0)	0 (0.0)	0 (0.0)	0.295
Lymphocyte count (mean (SD))	1.20 (0.53)	2.05 (NA)	1.19 (0.51)	1.07 (0.47)	NaN (NA)	1.34 (0.69)	NA
Gene expression							
IFN Score A (mean (SD))	4.79 (2.06)	6.25 (0.18)	4.08 (1.50)	4.33 (2.07)	4.65 (1.10)	5.15 (1.82)	0.184
IFN Score B (mean (SD))	4.28 (1.12)	5.52 (0.94)	4.56 (1.11)	4.60 (1.46)	4.37 (0.51)	4.83 (1.46)	0.794

Characteristic	IM	MCTD	pSS	SLE	SSc	UCTD	p
Plasmablast Score (mean (SD))	8.54 (1.41)	9.31 (1.41)	8.45 (0.98)	8.36 (1.20)	8.66 (0.72)	8.54 (1.20)	0.885
Myeloid Score (mean (SD))	4.19 (1.10)	5.41 (0.98)	4.69 (1.27)	4.55 (1.39)	4.29 (0.61)	4.75 (1.48)	0.831
Erythropoiesis Score (mean (SD))	7.03 (0.80)	7.82 (1.25)	7.04 (1.13)	7.29 (1.48)	7.00 (0.42)	7.17 (1.64)	0.969
Inflammation Score (mean (SD))	5.05 (0.97)	5.51 (0.72)	5.57 (1.34)	5.30 (1.52)	5.03 (0.82)	5.41 (1.45)	0.958
Disease activity							
PGA (mean (SD))	6.21 (2.34)	2.30 (0.71)	3.22 (1.91)	3.56 (2.63)	3.25 (0.92)	2.76 (2.05)	0.020
Overall disease flare(%)	5 (71.4)	0 (0.0)	9 (56.2)	22 (47.8)	2 (50.0)	18 (36.7)	0.310
Patient reported outcomes							
VAS Arthritis (mean (SD))	5.12 (2.50)	2.00 (NA)	4.28 (4.38)	5.51 (2.97)	5.30 (NA)	6.18 (2.09)	NA
VAS Global (mean (SD))	3.97 (1.51)	4.40 (3.39)	3.84 (3.19)	5.55 (2.86)	6.15 (1.06)	5.83 (2.56)	0.183
VAS Pain (mean (SD))	5.35 (2.03)	3.65 (3.61)	4.29 (3.36)	5.07 (2.87)	5.85 (0.92)	5.77 (2.16)	0.596
VAS Fatigue (mean (SD))	5.17 (2.24)	8.75 (0.64)	4.71 (3.38)	6.19 (2.99)	5.10 (0.00)	6.32 (2.75)	0.326
VAS Early Morning Stiffness (mean (SD))	4.18 (2.07)	6.65 (1.91)	4.14 (4.09)	5.28 (2.88)	5.70 (0.14)	6.41 (2.66)	0.232

Table 16: Baseline characteristics of DEFINITION cohort C

Characteristic	Baseline	3 months	6 months	p
n	124	58	68	
Demographics				
Age (mean (SD))	47.1 (13.2)	46.4 (13.8)	46.8 (12.6)	0.960
Gender = M (%)	21 (16.9)	11 (19.0)	15 (22.1)	0.685
Ethnicity (%)				0.823
Asian	16 (13.6)	10 (17.5)	10 (15.4)	
Black	11 (9.3)	1 (1.8)	5 (7.7)	
Other/Mixed	8 (6.8)	5 (8.8)	6 (9.2)	
Unknown	8 (6.8)	4 (7.0)	3 (4.6)	
White	75 (63.6)	37 (64.9)	41 (63.1)	
Index of Multiple Deprivation Rank (mean (SD))	13864 (10086)	14936 (9320)	14477 (9605)	0.833
Charlson Comorbidity Index (mean (SD))	2.06 (1.54)	2.04 (1.37)	1.92 (1.29)	0.811
Diagnosis (%)				0.848

Characteristic	Baseline	3 months	6 months	p
IM	7 (5.6)	2 (3.7)	3 (4.6)	
MCTD	2 (1.6)	1 (1.9)	1 (1.5)	
pSS	16 (12.9)	8 (14.8)	6 (9.2)	
SLE	46 (37.1)	22 (40.7)	29 (44.6)	
SSc	4 (3.2)	0 (0.0)	0 (0.0)	
UCTD	49 (39.5)	21 (38.9)	26 (40.0)	
Laboratory				
dsDNA (%)	35 (31.0)	20 (40.8)	23 (36.5)	0.451
Ro60 (%)	45 (39.8)	20 (40.8)	23 (36.5)	0.877
Ro52 (%)	27 (23.9)	13 (26.5)	12 (19.0)	0.623
La (%)	12 (10.6)	5 (10.2)	6 (9.5)	0.974
Sm (%)	12 (10.6)	5 (10.2)	5 (7.9)	0.842
SmRNP (%)	20 (17.7)	12 (24.5)	12 (19.0)	0.602
RNP (%)	9 (8.0)	6 (12.2)	4 (6.3)	0.520
Jo1 (%)	2 (1.8)	1 (2.0)	1 (1.6)	0.984

Characteristic	Baseline	3 months	6 months	p
Centromere (%)	4 (3.5)	0 (0.0)	1 (1.6)	0.344
Chromatin (%)	23 (20.4)	13 (26.5)	10 (15.9)	0.382
Ribosomal P (%)	5 (4.4)	2 (4.1)	2 (3.2)	0.920
Lupus anticoagulant (%)	1 (0.9)	0 (0.0)	0 (0.0)	0.608
Lymphocyte count (mean (SD))	1.19 (0.56)	1.51 (1.14)	1.26 (0.56)	0.067
Gene expression				
IFN Score A (mean (SD))*	4.69 (1.90)	5.26 (1.87)	5.35 (1.62)	0.028
IFN Score B (mean (SD))*	4.67 (1.37)	5.34 (1.30)	5.55 (1.14)	<0.001
Plasmablast Score (mean (SD))*	8.48 (1.17)	8.58 (1.02)	8.83 (1.01)	0.108
Myeloid Score (mean (SD))*	4.63 (1.37)	5.36 (1.39)	5.29 (1.38)	<0.001
Erythropoesis Score (mean (SD))*	7.19 (1.44)	7.58 (1.38)	8.05 (1.41)	<0.001
Inflammation Score (mean (SD))*	5.36 (1.40)	5.93 (1.38)	5.76 (1.40)	0.022
Disease activity				
PGA (mean (SD))	3.38 (2.39)	2.28 (1.93)	2.77 (2.28)	0.011
Overall disease flare (%)	56 (45.2)	31 (53.4)	25 (36.8)	0.171

Characteristic	Baseline	3 months	6 months	p
Patient reported outcomes				
VAS Arthritis (mean (SD))	5.54 (2.72)	4.50 (2.94)	4.47 (3.03)	0.146
VAS Global (mean (SD))	5.28 (2.76)	4.65 (2.77)	4.42 (2.61)	0.123
VAS Pain (mean (SD))	5.23 (2.63)	4.13 (2.73)	3.92 (2.49)	0.007
VAS Fatigue (mean (SD))	5.99 (2.91)	5.36 (3.03)	4.91 (3.00)	0.081
VAS Early Morning Stiffness (mean (SD))	5.59 (2.88)	4.53 (2.75)	4.51 (2.62)	0.032

*Note that gene expression scores are untransformed delta-CT, so numerically higher values indicate lower expression

Table 17: Summary variables at baseline, 3 and 6 month follow up

17.4.2 Therapeutic changes over follow up period

Therapeutic management over follow up is summarised in **Error! Reference source not found.**. Glucocorticoid dose escalations were notably more frequent at baseline and 3 month follow up with respect to 6 months. At every appointment >10% of patients had escalations in dose or started a new oral DMARD/IS. Hydroxychloroquine initiation was common with a >10% frequency over the first 2 follow up visits. There was also a general trend toward greater therapeutic change at baseline and 3 months relative to 6 month follow up.

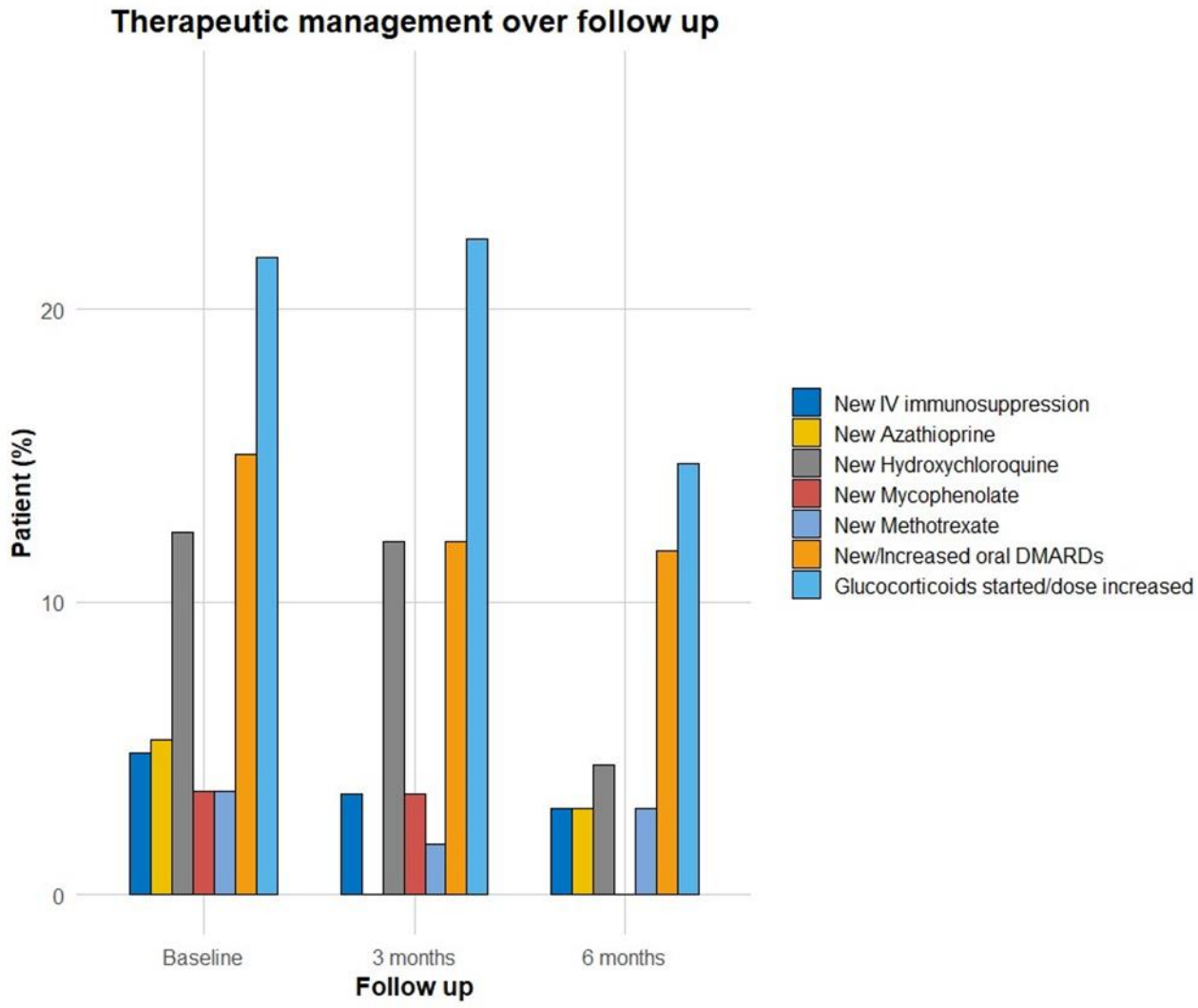


Figure 23: Therapeutic management at baseline, 3 and 6 months follow up

17.4.3 Disease activity over follow up

Proportion of patients in disease flare at each follow up visit are summarised in **Error! Reference source not found.**. There was a non-significant reduction in disease flare over time (p = 0.66, f = 0.34) with a peak prevalence at 3 months follow up. This mirrors the highest incidence of glucocorticoid escalations in **Error! Reference source not found.**.

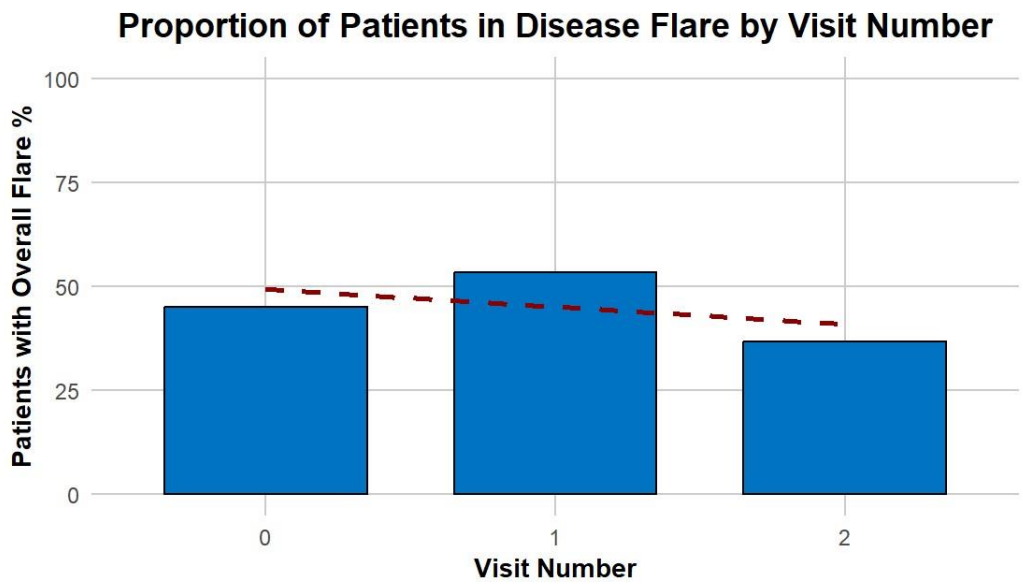


Figure 24: Proportion of patients in disease flare at follow up

17.5 RESULTS SUBCHAPTER 1: CAUSAL MODEL TO IDENTIFY FACTORS FOR LINEAR MIXED EFFECTS MODEL ANALYSIS

Within this results subchapter I present the results of a causal inference informed approach to modelling within the DEFINITION C cohort.

17.5.1 Directed acyclic graph analysis

To determine key variables and causal inference pathways prior to multivariate analysis directed acyclic graph analysis was carried out. This pathway was synthesized following formalized consultation with experts in ANA-RMD and interferon-associated disease. This took the form of a focus group with 2 experts in autoimmune disease and interferonopathy and a statistician. Directed acyclic graphs were drafted using the Dagitty package in R.

The resulting DAG pathway is summarised in Figure 25 DAG analysis identified non-white ethnicity and ENA Ro or RNP antibody positivity, as important covariates for correction when synthesizing a causal pathway.. Logistic regression models were created using these 4 baseline covariates as predictors of subsequent disease flare at 3 and 6 month follow up. These are summarised in Table 18, none were found to be statistically significant predictors of subsequent disease flare at follow up.

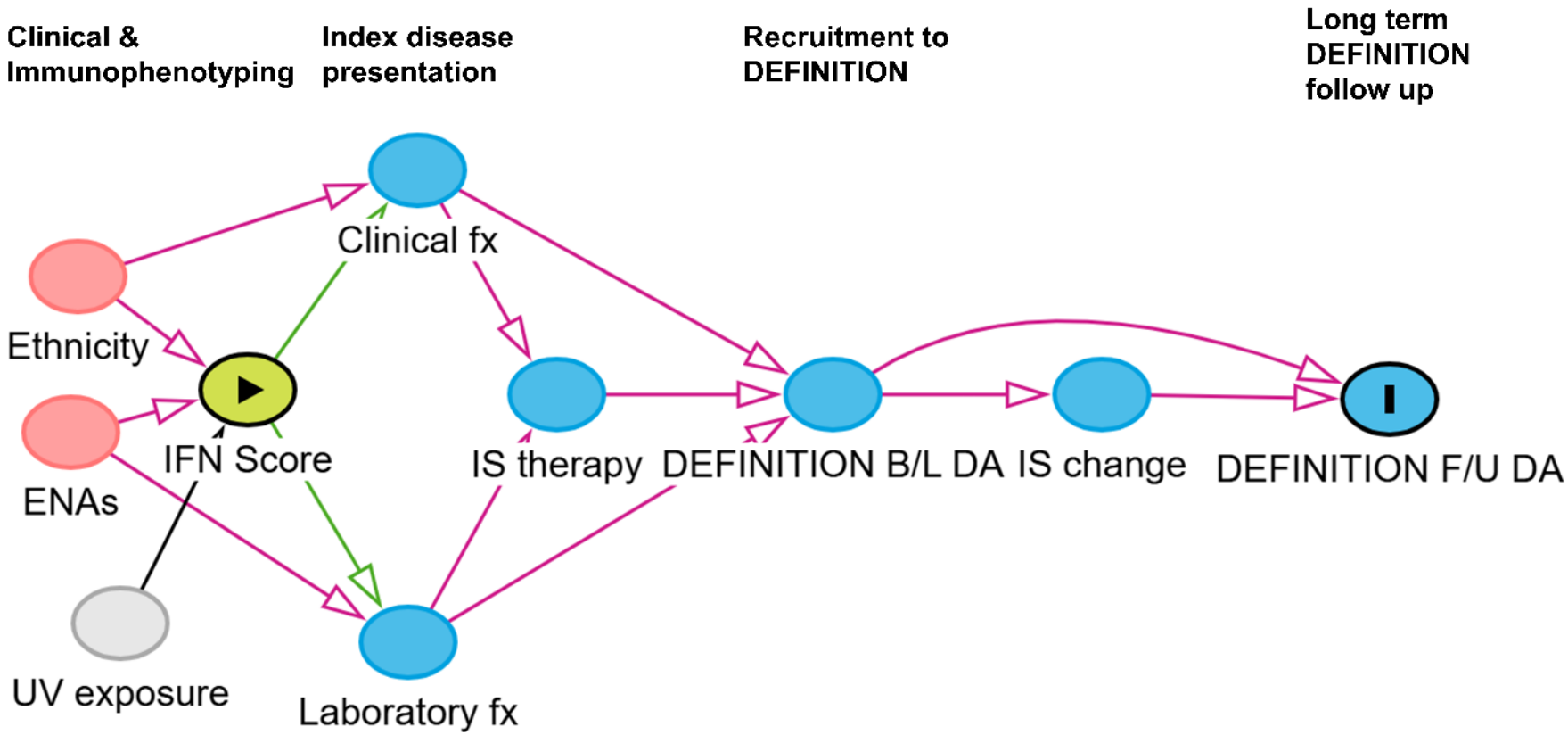


Figure 25: Impact of IFN-I expression on long term outcomes. B/L = baseline, F/U = follow up, DA = disease activity, IS = immunosuppression

Covariate		estimate	std.error	statistic	p value
IFN Score A as a baseline predictor of flare at 3 months					
1	(Intercept)	-0.79	1.05	-0.75	0.45
2	IFN Score A	0.09	0.17	0.52	0.61
3	Anti-Ro positivity	0.27	0.60	0.45	0.66
4	RNP/SmRNP or Sm positivity	0.64	0.72	0.89	0.37
5	European ancestry	0.31	0.62	0.50	0.61
IFN Score B as a baseline predictor of flare at 3 months					
1	(Intercept)	-1.55	1.40	-1.10	0.27
2	IFN Score B	0.22	0.23	0.95	0.34
3	Anti-Ro positivity	0.24	0.60	0.41	0.68
4	RNP/SmRNP or Sm positivity	0.70	0.73	0.96	0.34
5	European ancestry	0.37	0.62	0.59	0.56
IFN Score A as a baseline predictor of flare at 6 months					
1	(Intercept)	-2.28	1.12	-2.04	0.04
2	IFN Score A	0.20	0.17	1.17	0.24
3	Anti-Ro positivity	0.78	0.60	1.29	0.20
4	RNP/SmRNP or Sm positivity	0.08	0.70	0.11	0.91
5	European ancestry	0.72	0.60	1.20	0.23
IFN Score B as a baseline predictor of flare at 6 months					
1	(Intercept)	-2.29	1.35	-1.70	0.09
2	IFN Score B	0.20	0.22	0.92	0.36
3	Anti-Ro positivity	0.66	0.58	1.14	0.25
4	RNP/SmRNP or Sm positivity	-0.04	0.68	-0.05	0.96
5	European ancestry	0.82	0.59	1.38	0.17

Table 18: Causal inference informed logistic regression models for baseline IFN scores as predictors of 3- and 6-month flare

17.5.2 Linear mixed effects modelling

Linear mixed effect models were constructed as explained in the methods in section 17.3.6. Patients were analysed as an overall cohort to identify general trends in the data. then subset by those who received new oral immunosuppressants at baseline and by those in disease flare at baseline. The rationale for these sub analyses was to explore whether gene expression scores in different patient subgroups were affected differentially by disease activity and therapeutic changes.

Overall linear mixed effects modelling of Cohort C patient using all n=250 patient visits showed small but significant improvements in IFN Score B and Erythropoiesis scores at 6 months characterised by increases in ΔCT values (Estimate 0 51, p=0.04 for IFN Score B) and (Estimate = 0.64, p=0.01 for Erythropoiesis score)(**Error! Reference source not found.**). The IFN Score A model demonstrated borderline significant increases in gene expression (Characterised by reduction in ΔCT values) associated with RNP (Estimate -0.93, p = 0.09) and Ro antibody (Estimate -0.88, p=0.06) positivity(**Error! Reference source not found.**).

Other covariates demonstrated small borderline effects on overall gene expression scores with wide 95% confidence intervals as depicted in **Error! Reference source not found.**. Random effects sizes were substantial for all gene expression models particularly IFN Score A which showed substantial dispersion in comparison to IFN Score B with a random effects variance of 1.883 (**Error! Reference source not found.**). Although random effects variance was reduced in IFN Score B compared to A, residual variance remained high (Variance =1.28, SD = 1.13).

17.5.2.1 Subset by therapeutic initialisation

Given the known modifying effects of oral IS/DMARD therapy on gene expression a subset analysis was carried out on 17 patients who started on oral DMARD at their baseline appointment. Given the small sample size limited conclusions can be drawn from this data. IFN Score B and Erythropoiesis scores improved (characterised by increases in ΔCT values) with increasing visit number (Estimates 0.43 and 0.51, p values 0.01 and 0.16, respectively)(**Error! Reference source not found.**) Random effects variance was notably higher in IFN Score A, Myeloid and Inflammatory scores relative to the others (**Error! Reference source not found.**). Random effects and residual variance were lowest in the Plasmablast score model (**Error! Reference source not found.**).

17.5.2.2 Subset by baseline flare

Subset analysis was also carried out on 56 patients noted to be in disease flare at baseline. This demonstrated higher IFN Score A and B gene expression (Represented by a

borderline lower ΔCT values) associated with Ro antibody positivity in keeping with its known biological function (Estimate -0.41, p = 0.09) (**Error! Reference source not found.**). Plasmablast signatures improved significantly with commencement of new oral DMARDs at baseline (Estimate 0.75, p = 0.024) (**Error! Reference source not found.**). Plasmablast (Variance =0.42) and IFN Score B models (Variance = 0.57) had the lowest random effects variance of all the models(**Error! Reference source not found.Error! Reference source not found.**). Residual variance was again lowest in the plasmablast score model(**Error! Reference source not found.**).

17.5.2.3 Preliminary mediation analysis of baseline flare subset

Preliminary mediation analysis of the baseline flare subset of patients with IFN Score B as the exposure, flare at 3 months as the outcome and commencement of new hydroxychloroquine, new glucocorticoid therapy or a new DMARD as the 3 mediators demonstrated no significant direct or mediated effect of IFN-Score B on 3-month disease flare. Mediation analysis was not performed within the oral DMARD starter subset as sample sizes were too small to draw significant conclusions.

17.5.2.4 Variance Inflation Factor (VIF) analysis

VIF analysis of the models synthesised with the covariates above showed no significant multicollinearity (VIF<2) associated with any of the models for each of the gene expression scores.

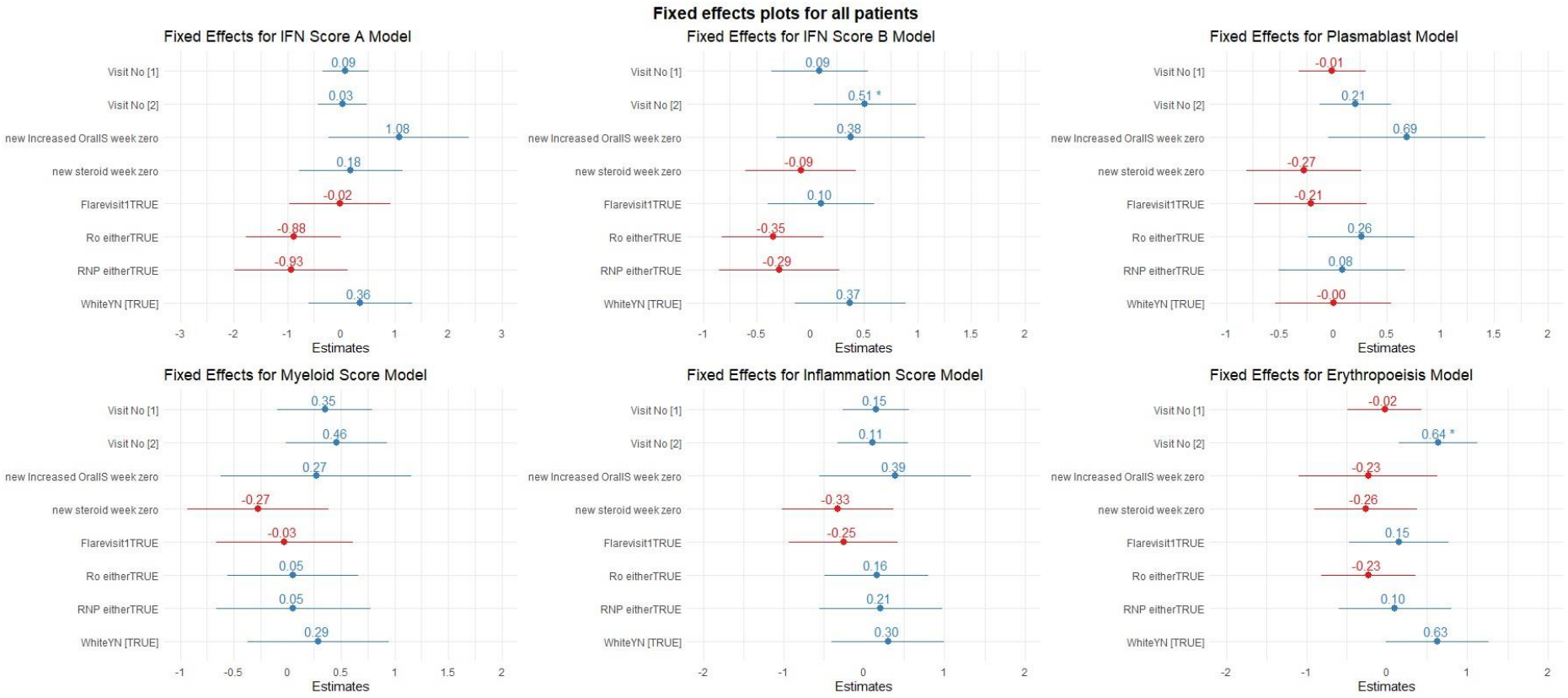


Figure 26: Fixed effects for all Cohort C patients with follow up. Note that gene expression scores are untransformed delta-CT, so numerically higher values indicate lower expression

Point estimates greater than zero (blue) indicate that the factor is associated with increased gene expression during follow up compared to baseline. Point estimates less than zero (red) indicate that the factor is associated with reduced gene expression during follow up. Error bars represent 95% confidence intervals.

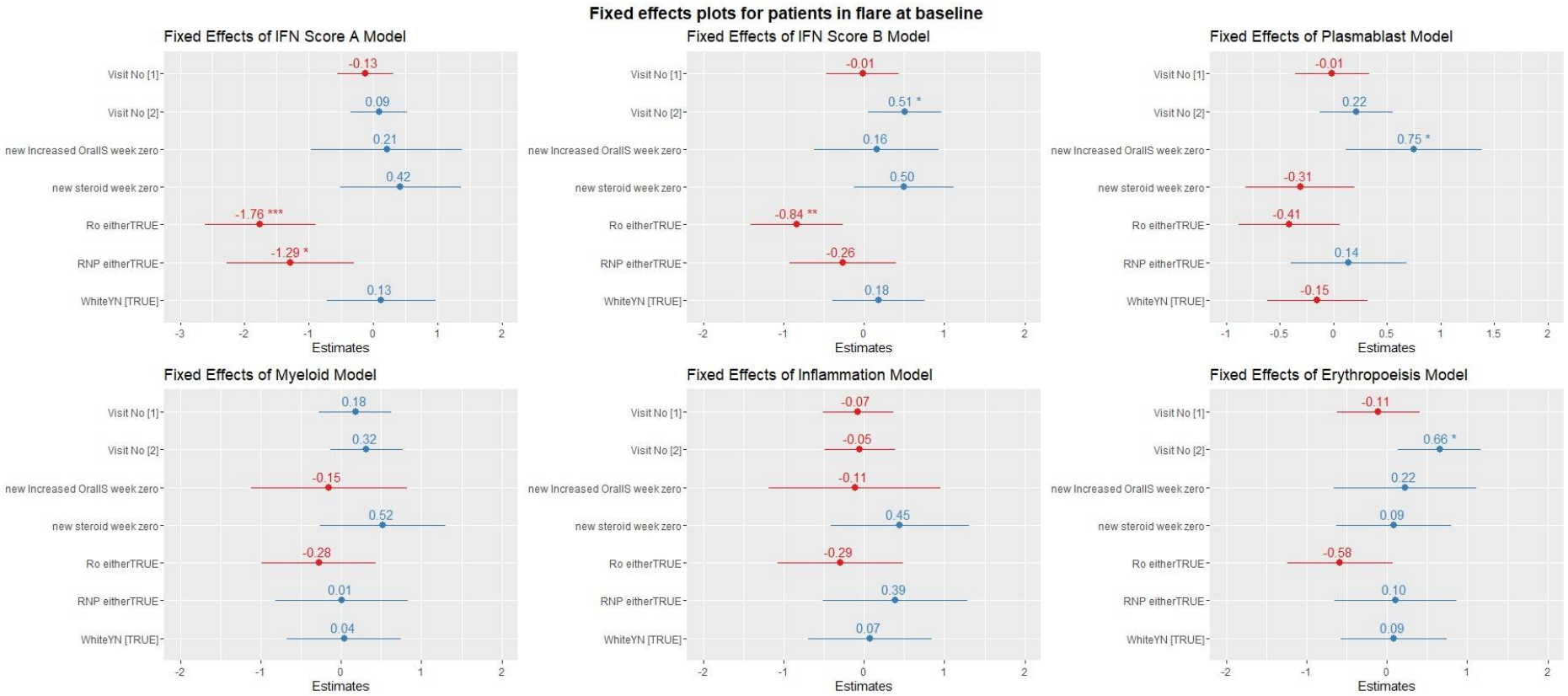


Figure 27: Fixed effects for all Cohort C patients in disease flare at baseline. Note that gene expression scores are untransformed delta-CT, so numerically higher values indicate lower expression

Point estimates greater than zero (blue) indicate that the factor is associated with increased gene expression during follow up compared to baseline. Point estimates less than zero (red) indicate that the factor is associated with reduced gene expression during follow up. Error bars represent 95% confidence intervals.

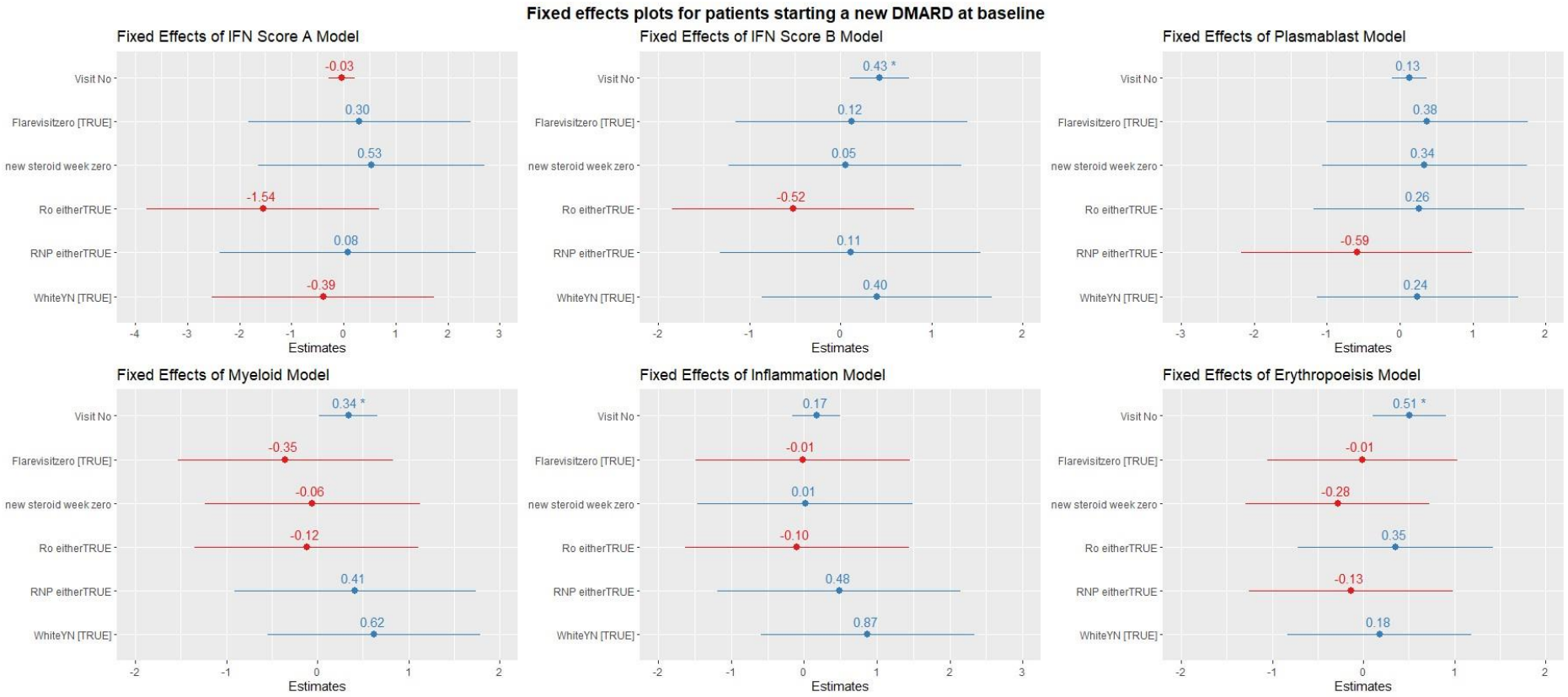


Figure 28: Fixed effects for all Cohort C patients starting a new DMARD at baseline. Note that gene expression scores are untransformed delta-CT, so numerically higher values indicate lower expression

Point estimates greater than zero (blue) indicate that the factor is associated with increased gene expression during follow up compared to baseline. Point estimates less than zero (red) indicate that the factor is associated with reduced gene expression during follow up. Error bars represent 95% confidence intervals.

Model	Population	Random effects variance (SD)	Residual variance (SD)
IFN Score A	All Cohort C	1.883 (1.372)	1.149 (1.072)
IFN Score B	All Cohort C	0.1996 (0.4467)	1.2829 (1.1327)
Plasmablast score	All Cohort C	0.4964 (0.7046)	0.6093 (0.7806)
Myeloid score	All Cohort C	0.6356 (0.7972)	1.2219 (1.1054)
Inflammation score	All Cohort C	0.8209 (0.906)	1.0539 (1.027)
Erythropoiesis score	All Cohort C	0.5363 (0.7323)	1.318 (1.1481)
IFN Score A	Flare patients at baseline	1.8962 (1.3770)	0.8183 (0.9046)
IFN Score B	Flare patients at baseline	0.5651 (0.7517)	0.9314 (0.9651)
Plasmablast score	Flare patients at baseline	0.4244 (0.6515)	0.5289 (0.7273)
Myeloid score	Flare patients at baseline	1.138 (1.067)	0.895 (0.946)
Inflammation score	Flare patients at baseline	1.4855 (1.2188)	0.8493 (0.9216)
Erythropoiesis score	Flare patients at baseline	0.7611 (0.8724)	1.2104 (1.1002)
IFN Score A	New DMARD at baseline	3.3798 (1.8384)	0.2618 (0.5117)
IFN Score B	New DMARD at baseline	0.9358 (0.9766)	0.4550 (0.6746)
Plasmablast score	New DMARD at baseline	1.3367 (1.156)	0.2333 (0.483)
Myeloid score	New DMARD at baseline	0.7864 (0.8868)	0.4451 (0.6672)
Inflammation score	New DMARD at baseline	1.3768 (1.1734)	0.4501 (0.6709)
Erythropoiesis score	New DMARD at baseline	0.3654 (0.6044)	0.7324 (0.8558)

Table 19: REML, random effects and residual variance for mixed effects mode

17.6 RESULTS SUBCHAPTER 2: CAUSAL MODEL TO IDENTIFY FACTORS FOR LINEAR MIXED EFFECTS MODEL ANALYSIS

Following a causal inference defined approach multivariable modelling was used to assess the collective predictive utility of baseline clinical, demographic, laboratory, and gene expression variables for disease flare prediction at 3 and 6 months, irrespective of the causal pathway. Given the complex pathophysiological mechanisms underlying ANA-RMD disease flare, this approach was employed to account for potential confounders and interactions between covariates.

17.6.1 Model composition

Multivariable logistic regression (MVA) was carried out using baseline demographic, clinical and laboratory factors as predictors of disease flare at 3 and 6 months. For these models 6 gene expression scores were incorporated alongside Age, European/non-European ancestry, Charlson Comorbidity Index, new baseline treatment with Hydroxychloroquine (HCQ) or glucocorticoids and the presence of either Ro52/60 or RNP/SmRNP antibody positivity. Unfortunately, due to small sample sizes, more detailed analysis of ancestry using more specific subsets was not possible. Individual ANA values were amalgamated due to significant multicollinearity within a small sample size model. Further therapeutic subsets could not be incorporated again due to small sample sizes available for analysis.

17.6.2 Multivariable predictive model (Baseline predictors of flare at 3 months)

The multivariable predictive model for baseline predictors of flare is shown in Table 20. Standard error values for several covariates were high in keeping with the small sample sizes available for modelling. Nagelkerke's R^2 value was calculated as 0.410, suggesting that the model was able to explain approximately 41% of the total variance. The Hosmer-Lemeshow test value gave a p value of 0.139 suggesting no significant lack of fit within the model.

A receiver operating curve plot for the 3-month predictive model is shown in Figure 29. The AUROC value of 0.843 suggested acceptable discrimination, within the context of an overall small sample size model. Summary statistics for this model and the 6-month predictive model are shown in Table 22.

Covariate	estimate	std.error	statistic	p value
(Intercept)	5.83	5.13	1.14	0.26
IFN Score A	-0.23	0.46	-0.5	0.62
IFN Score B	1.04	1.6	0.65	0.52
Plasmablast Score	-0.46	0.56	-0.82	0.41
Myeloid Score	2.1	1.67	1.26	0.21
Erythropoesis Score	-0.47	0.96	-0.49	0.62
Inflammation Score	-2.25	1.33	-1.69	0.09
Age	-0.06	0.05	-1.14	0.25
Charlson Comorbidity Index	0.26	0.46	0.58	0.57
Ro 52 or 60 positivity	0.64	1	0.64	0.52
RNP or SmRNP positivity	1.62	1.13	1.43	0.15
New HCQ at baseline	1.42	1.5	0.94	0.35
New glucocorticoid at baseline	0.72	0.81	0.89	0.38
White Ethnicity	1.45	1.06	1.37	0.17

Table 20: Multivariable analysis of baseline predictors of flare at 3 month follow up

17.6.3 Multivariable predictive model (Baseline predictors of flare at 6 months)

The multivariable predictive model for baseline predictors of flare is shown in Table 21. Standard error values for several covariates were high in keeping with the small sample sizes available for modelling. Nagelkerke's R^2 value was calculated as 0.302, suggesting that the model was only able to explain approximately 30% of the total variance. The Hosmer-Lemeshow test value gave a p value of 0.438 suggesting no significant lack of fit within the model.

A receiver operating curve plot for the 6-month predictive model is shown in Figure 29. The AUROC value of 0.786 suggested acceptable discrimination, within the context of an overall small sample size model. Summary statistics for both multivariable models are shown in Table 22.

Covariate	estimate	std.error	statistic	p value
(Intercept)	-6.16	3.70	-1.67	0.10
IFN Score A	0.26	0.33	0.80	0.43
IFN Score B	0.24	0.96	0.26	0.80
Plasmablast Score	0.63	0.47	1.35	0.18
Myeloid Score	-0.09	1.14	-0.08	0.94
Erythropoesis Score	0.10	0.61	0.16	0.87
Inflammation Score	-0.38	0.89	-0.43	0.67
Age	-0.07	0.05	-1.56	0.12
Charlson Comorbidity Index	0.53	0.37	1.43	0.15
Ro 52 or 60 positivity	0.85	0.89	0.95	0.34
RNP or SmRNP positivity	0.10	0.88	0.12	0.91
New HCQ at baseline	0.49	1.11	0.44	0.66
New glucocorticoid at baseline	0.48	0.84	0.57	0.57
White Ethnicity	1.18	0.90	1.32	0.19

Table 21: Multivariable analysis of baseline predictors of flare at 6 month follow up

Multivariable analysis - 3mth flare prediction	
Balanced Accuracy	0.76
AUROC	0.84
Specificity	0.70
Sensitivity	0.83
Positive Predictive Value	0.76
Negative Predictive Value	0.78
Multivariable analysis - 6mth flare prediction	
Balanced Accuracy	0.69
AUROC	0.79
Specificity	0.74
Sensitivity	0.63
Positive Predictive Value	0.60
Negative Predictive Value	0.77

Table 22: Summary statistics for 3- and 6-month multivariable predictive models

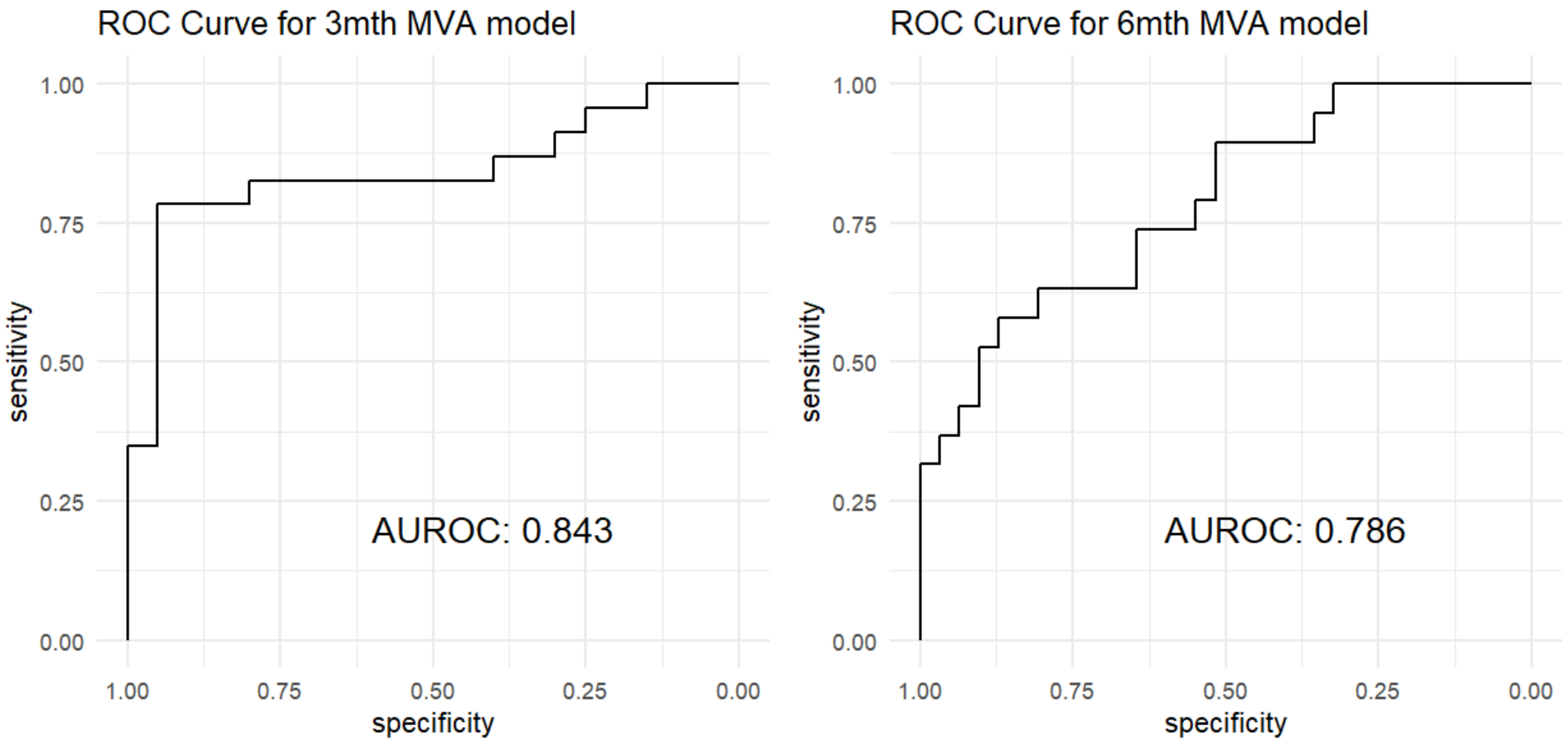


Figure 29: Receiver Operating Curves for 3 month and 6-month multivariable predictive models

17.1 DISCUSSION

17.1.1 Causal Modelling Discussion

Unfortunately, within the framework of the current study gene expression scores in isolation were poorly predictive of 3- and 6-month disease flare in the proposed causal model. This suggests either a lack of association, or that the generated models had omitted key covariates in the causal pathway linking gene expression scores such as the baseline IFN score to the measured outcomes of flare at 3 and 6 months. The complex interweaving causal pathways involved linking these gene expression scores to 6 month follow up, via differential therapeutic utilisation affected by disease label, disease activity and individual factors would necessitate extensive mediation analyses which were felt to be beyond the scope of this doctorate.

Causal inference informed logistic regression modelling of IFN scores as predictors of subsequent disease flare failed to identify a robust causal association. As mentioned above this could suggest an absence of causal association between gene expression scores or alternatively significant flaws within the drafted DAG and causal pathway. In truth the proposed DAG is likely overly naïve in the context of the complex interwoven causal pathway linking baseline gene expression scores to subsequent disease flare. These assumptions would need to be refined and reassessed in future work. DAG analysis allows for these assumptions to be documented formally, allowing empirical improvement as we refine our understanding of the causal pathway.

Secondary analysis of subgroups of patients included those starting new immunosuppression and those in flare at study baseline Sample sizes for subgroup analysis of oral DMARD samples were small with only 17 patients starting an oral immunosuppressant at baseline. Because of this drawable conclusions are limited and further sub-analysis looking at individual DMARD choices was not possible. This was felt to be in large part due to the impact of the COVID-19 pandemic on the recruitment of the Cohort C study arm, as non-essential visits were heavily restricted during this recruitment phase. Consequentially sample sizes available for modelling were small, with key covariates underrepresented. The exclusion of non-essential visits may have also created further selection pressures with only high disease activity or high symptom burden patients attending for clinical review. Moreover, those patients on more immunosuppressive treatments may also have been disincentivised from attending hospital follow up during this time due to the risk of nosocomial COVID-19 infection.

Large residual effect sizes particularly within IFN Score B, Myeloid and Inflammation score based linear mixed effects models in both unselected Cohort C patients and the flare subset suggest a high degree of unexplained variability not accounted for within this modelling approach. This implies the need for further work in defining input covariates and causal inference pathways for these models. This contrasts with the lower observed variance in the Plasmablast score models, suggesting that this analysis performed better for certain gene expression scores than others. Effect sizes may also be very small and would require a larger sample size, something severely hampered by poor recruitment during the COVID-19 pandemic. Further work in this area may need to utilise a smaller number of covariates within predictive modelling because of this under recruitment.

17.1.2 Multivariable Modelling Discussion

This approach was broadened in response to this with the addition of multivariable predictive analysis which was able to define an effective predictive model using a composite of clinical, demographic and biomarker data. The applicability of using a gene expression panel in routine clinical practice is questionable at present, however these modelling approaches could be refined and may have some utility in the clinic. This should be expanded upon and explored in further work.

17.1.3 Overall Discussion

Several issues may in part explain the difficulties when utilising this longitudinal data. DEFINITION Cohort C was recruited as a subset of the overall Cohort B with either active disease on presentation or perceived disease lability over the preceding months. However, in practice this may have introduced a significant selection bias component into the observational study. In practice many of these patients were recruited with active disease at baseline, or conceptually due to other clinician defined poor prognostic markers which would predispose to flare. This may have introduced a Berkson's paradox whereby low disease activity and low biomarker patients are systematically excluded from the cohort causing a

falsely negative correlation or abrogating any possible effect. This is depicted in Figure 30.

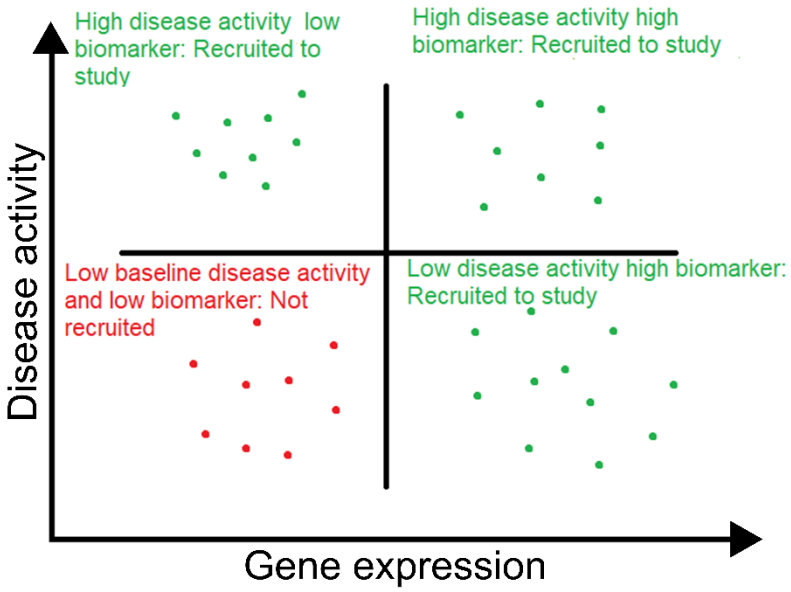


Figure 30: Berkson's paradox as applied to DEFINITION Cohort C

These issues are commonly encountered throughout the literature, in fact higher IFN-I expression has been associated with both increased disease activity and paradoxical reductions in fatigue scores in a previous study(428). Within this study the relationship was thought to be paradoxical, however it may itself be due to similar selection bias issues to those experienced here. Particularly in the context of mounting evidence for the role of IFN-I in ANA-RMD pathogenesis. Alternatively molecular approaches using single cell sequencing to characterise transcriptional and epigenetic signatures associated with disease flare in ANA-RMDs could be applied, similar to those underway in SLE(429). This may permit greater fidelity when identifying common predictive signatures.

In summary this work demonstrates the utility and pitfalls of applying causal inference methodologies both within the study design and analysis phases when answering these clinical questions. Gene expression scores appear to have some utility and responsiveness to standard of care DMARD therapies, however high residual variance within this modelling appears to suggest that as yet unidentified extraneous factors significantly impact their expression. Future work in this area should extend the directed acyclic graph approach to causal models and incorporate it in the planning stages of future data collection, rather than apply it retrospectively as was the case here. Unfortunately, the COVID-19 pandemic significantly impacted the Cohort C arm of the DEFINITION study and therefore limited the potential conclusions we can draw from this work.

18 DISCUSSION

18.1 GENERAL DISCUSSION OF RESULTS

Within this thesis I explored a variety of novel approaches to the central problem of reclassification and disease stratification within the field of ANA-RMDs. Each of these approaches had different strengths and weaknesses, culminating in my hybrid approach utilising the PRECISESADS and DEFINITION cohorts. Throughout this process I have been able to develop skills ranging from basic clustering techniques in pure biomarker datasets, identifying basket populations of interest through my work in ANA-arthritis to applying complex deep learning methodology, causal inference approaches and mixed effects modelling to address different clinical questions.

Deep learning techniques proved effective in addressing the key problems of high dimensionality and sparsely populated datasets which have been problematic for previous researchers working in the field. This approach was also able to amalgamate data from clinical and biomarker data and its replication in two large ANA-RMD cohorts represents the primary output of this doctorate. The progression from flow cytometry-based biomarker-defined clusters, through clinically defined ANA arthritis, to a hybrid approach culminating in the evaluation of key gene expression scores highlights an evolving methodology for reclassification. This pathway illustrates the increasing sophistication and precision in categorizing disease states, utilising cellular biomarkers and clinical phenotypes, ultimately integrating gene expression and therapeutic insights to enhance classification accuracy. This process also demonstrated key weaknesses within the approaches outlined, particularly Results 1 and Results 4 as discussed in detail below.

Processing of text fields using 1-hot encoding proved sufficient to generate clinically significant ARC classes, however future work should consider employing text parsing techniques such as word2vec or GLoVe. These were briefly trialed in this work but not employed due to significant input spelling errors of medical terms during data entry into the DEFINITION dataset which were not correctable using automated functions. This necessitated manual review and correction of free text fields which was ultimately time consuming. Additionally, the PRECISESADS dataset did not contain the same narrative field format, so replication of more advanced text tokenization approaches such as these would have been problematic.

Within this thesis there is a defined progression from naïve, low complexity preliminary work

using simple clustering on flow cytometry data, through GMM in clinically derived clusters and on to deep learning-based approaches, mixed effects modelling, mediation analysis and causal inference methodology as the doctorate has progressed. This expansion of my skillset I believe is evidenced by this mounting complexity and represents in many ways the primary output of this work.

18.1.1 Discussion of Results 1

Within Results 1 I applied the simple K-means clustering technique to a cohort of patients with extant data from the DEFINITION cohort. Using this approach, I was able to identify a Tetherin dominant and Tetherin deplete cluster along with a UCTD dominant cluster. Overall cohort numbers were low, and results were of limited clinical and scientific utility. It does however suggest that although IFN-I activation is a common pathogenic process across ANA-RMDs, there is a sizeable differentiated and undifferentiated ANA-RMD population that has non-IFN mediated disease. This has been reflected in the TULIP-2 trial of the interferon blocking drug Anifrolumab, which showed high efficacy in skin disease and glucocorticoid dose reduction but failed to significantly reduce swollen and tender joint counts(196). A fact supported by Results 3 which derived an Inflammatory-MSK class, characterised by MSK disease and lower IFN score A and B expression than the Polyinflammatory and Myeloinflammatory ARC classes.

18.1.2 Discussion of Results 2

Within Results 2 I identified the clinical basket, ANA-Arthritis, from the DEFINITION cohort. I was able to characterise the similarity of the impact of ANA-arthritis across a range of legacy diagnoses. In this way I was able to demonstrate similarities in gene expression, clinical and patient reported outcomes between diagnostic groups, united by this clinical manifestation.

This was expanded upon further through the utilisation of GMM to sub stratify these patients into a high and low MSK activity cluster. The High MSK disease cluster contained all patients with active joint swelling and had significantly higher prednisolone usage, joint swelling prevalence, physician global assessment and Sm/RNP/SmRNP/Chromatin positivity. Tetherin and Interferon Score-A activity were significantly greater in this group, fibromyalgia and osteoarthritis prevalence were numerically lower. This sub stratification expanded upon simply taking all patients with active disease, being able to identify high risk patients with currently inactive disease using a data driven approach.

This work represents the first clinical basket identified within ANA-RMD and has garnered interest from industry partners such as Novartis who are considering a clinical trial in this patient group. This would potentially allow for targeted therapeutics to be made available to large numbers of patients, including many UCTD patients for the first time in rheumatology.

Potentially this would provide the impetus for follow on basket trials in ANA-RMDs and a sea change in how trials are carried out in rheumatology.

18.1.3 Discussion of Results 3

The two previous chapters emphasise why a hybrid of clinical and molecular is desirable to solve this problem. Molecular data is best to check whether the same therapeutic target is appropriate. But molecular data alone such as the work within PRECISESADS previously does not adequately describe individuals and clinical problems(262). Clinical data is necessary to ensure that potential basket approaches are adequately contextualised for clinical trial evaluation. The failure of pure biomarker approaches to deliver the restratification paradigm shift in ANA-RMDs suggests that current techniques may be insufficient. Moreover, the use of large data gene expression profiling may limit the utility and generalisability of these approaches.

Within Results 3 I combine clinical and standard of care laboratory data with a specific 96 gene panel Taqman array within this reclassification effort. This array is more easily undertaken in tertiary centres than either genome wide or flow cytometry-based approaches, though remains out of reach for resource poor or more rural settings. In combining these data, I define new classes which incorporate patients from each of the legacy diagnoses and are associated with differential long-term outcomes.

Two of the identified classes are characterised by gene expression markers, the Myeloinflammatory and Polyinflammatory groups. These groups comprise the patients with the most significant healthcare utilisation and long-term outcome impacts. These gene expression signatures can be evaluated within future work as markers of poor prognostic outcome or as molecular targets for inhibition in keeping with approaches such as tyrosine kinase inhibition in oncology. Basket trials based around these concepts would provide targeted therapies to the most needful of patients, removing healthcare inequalities secondary to naming conventions.

Two ARC classes are characterised by their clinical presentations, Sicca and Inflammatory MSK were identified. Sicca was noted to comprise a large proportion, but not all, of the pSS patient cohort. This suggests that existing diagnoses do partially capture the heterogeneity within the ANA-RMD cohort, with glandular pSS patients seemingly distinct from other groups. However, patients with systemic features of pSS appear to be distributed across the remaining ARC classes. Inflammatory MSK class patients drew from multiple diagnoses, similarly to the high activity ANA arthritis cohort identified in Results 2. As previously discussed, these patients are likely to benefit from targeted trials with MSK specific outcome measures.

Finally, the identification a high non inflammatory symptom burden cluster emphasises the importance of broadening our therapeutic armamentarium beyond immunosuppressive medications. These patients demonstrate biological markers of disease quiescence but continue to be disabled by troublesome fatigue and pain symptoms. Recruiting these patients into trials of immunosuppressive medications risks diluting eventual effect sizes. More problematically though, persistently failing to identify the non-inflammatory nature of these symptoms prevents these patients from accessing targeted rehabilitative and analgesic treatments, leaving them with persistent disabling symptoms. Therefore, current disease classifications do a double disservice to these people, and identification of them using VAE based approaches would permit more effective treatments, Basket trials in the same vein as the proposed ANA arthritis trial could also be completed, with rehabilitative therapies or other non-pharmacological approaches as potential interventions.

18.1.4 Discussion of Results 4

Results 3 defined 2 ARC cases characterised by distinct gene expression score signatures. To use these in clinical trials we need to understand the factors that moderate these scores, and how they change with therapy. Observational data such as the DEFINITION Cohort C is useful in the evaluation of this question. As part of a multivariable model, gene expression scores proved to be predictive of subsequent disease activity. Unfortunately, the application of causal inference, random effects modelling and mediation analysis techniques proved less effective. This illustrated the difficulties of prognostication using scores profoundly affected by sequential changes in therapeutics and the necessity of a causal inference-based approach when planning longitudinal studies such as these, as previous work within single disease cohorts such as SLE has suggested prognostic utility for interferon scores as an example(430).

Causal inference methodologies here were shown to be pivotal when designing and analysing these studies. In the case of this thesis their application in a post-hoc manner proved insufficient to account for potential selection bias within the cohort.

18.2 THIS THESIS IN CONTEXT

Within ANA-RMDs, rheumatology and autoimmunity in general the concept of basket trials and reclassification is widely discussed. However few studies employ a basket approach to autoimmunity, and there has been limited progression within the literature towards disease reclassification in ANA-RMDs. In this respect we lag far behind our oncology colleagues who have been employing this approach since the first trials of Imatinib in susceptible malignancies in 2008(431,432). Because of this work, demonstrated efficacy in several

subgroups and the associated low prevalence of dermatofibrosarcoma protuberans, systemic mastocytosis, hypereosinophilic syndrome and myeloproliferative disorders, Imatinib has been licenced by the FDA for these indications. Similar results have been observed in a trial demonstrating the efficacy of Larotrectinib in TRK fusion-positive adult and paediatric malignancies(431,433). Basket approaches are therefore not new, but despite similarities between AIDs and oncology, have yet to migrate effectively to these applications. Across autoimmunity in general a basket trial recently enrolled 81 patients with a combination of ANA-RMDs, inflammatory arthritides, autoimmune vasculitides, sclerosing cholangitis, autoimmune hepatitis and inflammatory bowel disease to receive treatment with low doses interleukin-2. The were able to demonstrate significant regulatory T-cell expansion and activation with concomitant improvements in disease activity scores for five out of six disease groups(263). This remains to date the only basket trial approach within the field, occurring 16 years behind early work in oncology.

With respect to the question of reclassification in ANA-RMDs this thesis is the first work to apply deep learning-based approaches in hybrid datasets in large multi disease cohorts. Sub stratification as noted within the introduction has been performed within individual diseases and biomarker-based reclassification has been trialled by the PRECISESADS group themselves(262,417,434–436). However, to date the outputs of these approaches have not yielded any tangible shift in the current paradigm.

18.3 Future Directions

In order to progress this research beyond the scope of this doctorate there are multiple avenues which could be explored. These are outlined below.

18.3.1 Investigation of high impact gene expression scores

Several gene expression scores were identified as being associated with high disease impact in ANA-RMD patients. Unfortunately, longitudinal assessment of these patients over 6 months did not demonstrate significant prognostic utility due to difficulty with observational datasets. In future formalised assessment as secondary explorative endpoints within a clinical trials framework, or indeed reappraisal of previous failed clinical trial data could provide a more precise environment in which to assess their prognostic utility. This is being explored with industry partners currently and could form part of future trial design stratification.

18.3.2 Basket trials in ANA-RMD arthritis

Basket trials could be developed in the ANA+ Arthritis subgroup which I have shown to be homogenous in terms of its biomarker and patient reported impact. Trials in this group could

eschew compound disease activity measures such as the SLEDAI-2K which can be problematic in describing partial disease responses. Instead, these could be replaced with a targeted measurement of musculoskeletal disease activity, akin to a DAS-28 score in rheumatoid arthritis.

Basket indications such as ANA+ Arthritis are likely to open up targeted therapies to patients with lower prevalence diseases where sufficient recruitment to clinical trials is problematic. This will address a widening inequity between the “haves” and “have nots” within ANA-RMDs. It will also incentivise industry partners to seek cross diagnostic indications when developing trial strategies, instead of focusing on the most prevalent ANA-RMDs

This will however require a validated disease activity measurement for arthritis which spans multiple diagnoses. Within this thesis we showed that the MSK domain of the BILAG-2004 index correlated well with physician assessed disease activity score the PGA. This will need to be expanded upon and improved, work has been developed previously in SLE arthritis with the LAMDA score, and work is on-going to assess the validity of this score across multiple ANA-RMDs within this lab(437).

Additional concerns are how regulators such as the FDA will regulate and licence these newer basket indications for therapy. There is precedent for this with the licencing of Nintedinib for pulmonary fibrosis of any cause and the multiple basket indications within the field of oncology previously mentioned. Progression of this basket trial approach within ANA-RMDs will require extensive collaboration with regulatory stakeholders and industry partners, to ensure trials are feasible and likely to produce tangible licences for targeted therapies.

18.3.3 Reappraisal of previously negative ANA-RMD trials

As discussed previously, there have been multiple high profile failures within trials of promising therapeutics in ANA-RMDs such as EXPLORER, LUNAR and SLE-BRAVE II(438–440). Reappraisal of these trials by sub stratifying into my ARC classes could permit better trial design and stratification in the future, avoiding costly disappointments. More problematic though is the discontinuation of promising therapies with biologically plausible methods of action and encouraging early results, let down by imperfect disease classification and poor trial design. This approach has been trialled previously using hierarchical clustering in the TRACTISS trial patients, however the scale of reclassification and cross-diagnostic cohorts used to derive this data greatly expand the utility of these findings(128).

18.3.4 Utilisation of VAE and text parsing approaches in native clinical record data

The efficacy of this hybrid approach in combining text and biomarker data, compressing to a low dimensionality space and deriving clusters could be expanded upon with collaborations

using primary care data. Text tokenisation approaches such as GloVE and Word2Vec can parse native text into tensors for subsequent analysis. Therefore, referrals from primary care to tertiary centres could be analysed using such an approach using standard of care laboratory and written clinical data as part of an initial triage process. This could highlight patients likely to have negative longer-term outcomes prior to the first visit and potentially flag these to clinicians for more expedited review. Shallow neural network approaches similar to this have been trialled to identify SLE patients from clinical note data(441). However, to date there has been no hybridised approach using text based and biomarker data.

18.3.5 Beyond the ANA-RMD paradigm

VAE based data compression of hybrid datasets has utility outside of the ANA-RMD paradigm. it could be applied to other complex datasets within medicine.

The strength of this approach is its versatility, VAEs have been applied to address problems ranging from medical image data compression to forecasting of road traffic data(442,443). Their role within medicine in compression of radiological images and analysis is perhaps the most developed. However, as we expand our clinical and biomarker datasets with high dimensionality such as the 3TR project, we will need to employ approaches such as VAEs to avoid overfitting, high computational cost and poor clustering performance.

It seems probable given the aims of the 3TR project that traditional dividing lines in autoimmune disease will erode as our understanding of the pathological processes increases. Clinical and biomarker data is a foundation which could be built upon with the inclusion of imaging and other data subsets as potential training datasets expand.

In summary this work illustrates the utility of this approach and derives several populations of interest which can guide early basket trials. Future application of this methodology in expanding datasets should result in greater precision when labelling patients with ANA-RMDs and other complex medical conditions

19 APPENDIX A

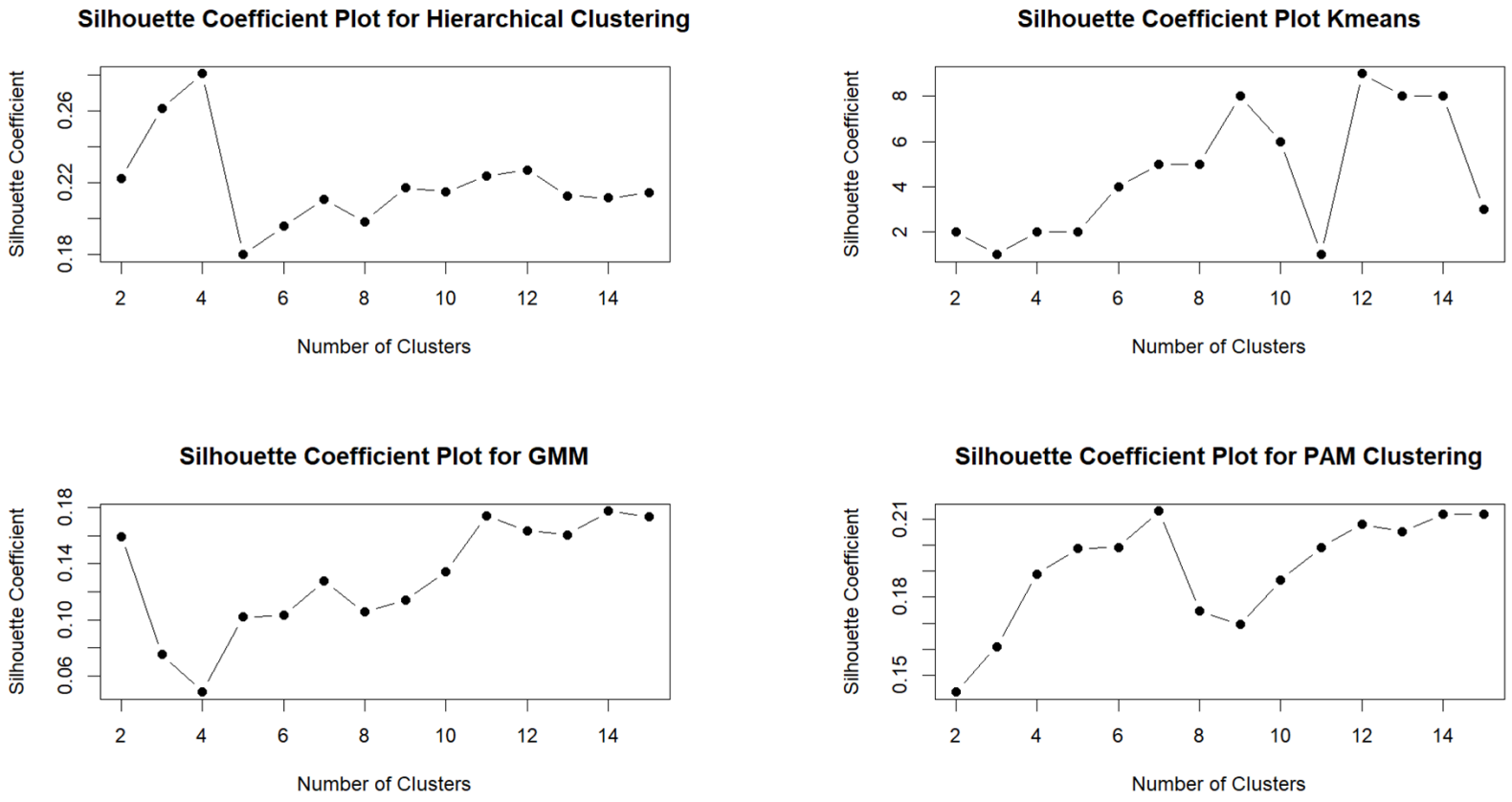


Figure 31: Average silhouette width plots of the ANA-Arthritis cohort in Results Chapter 2 clustered using hierarchical clustering, Kmeans, Gaussian mixture modelling and Partitioning around Medoids .

20 APPENDIX B

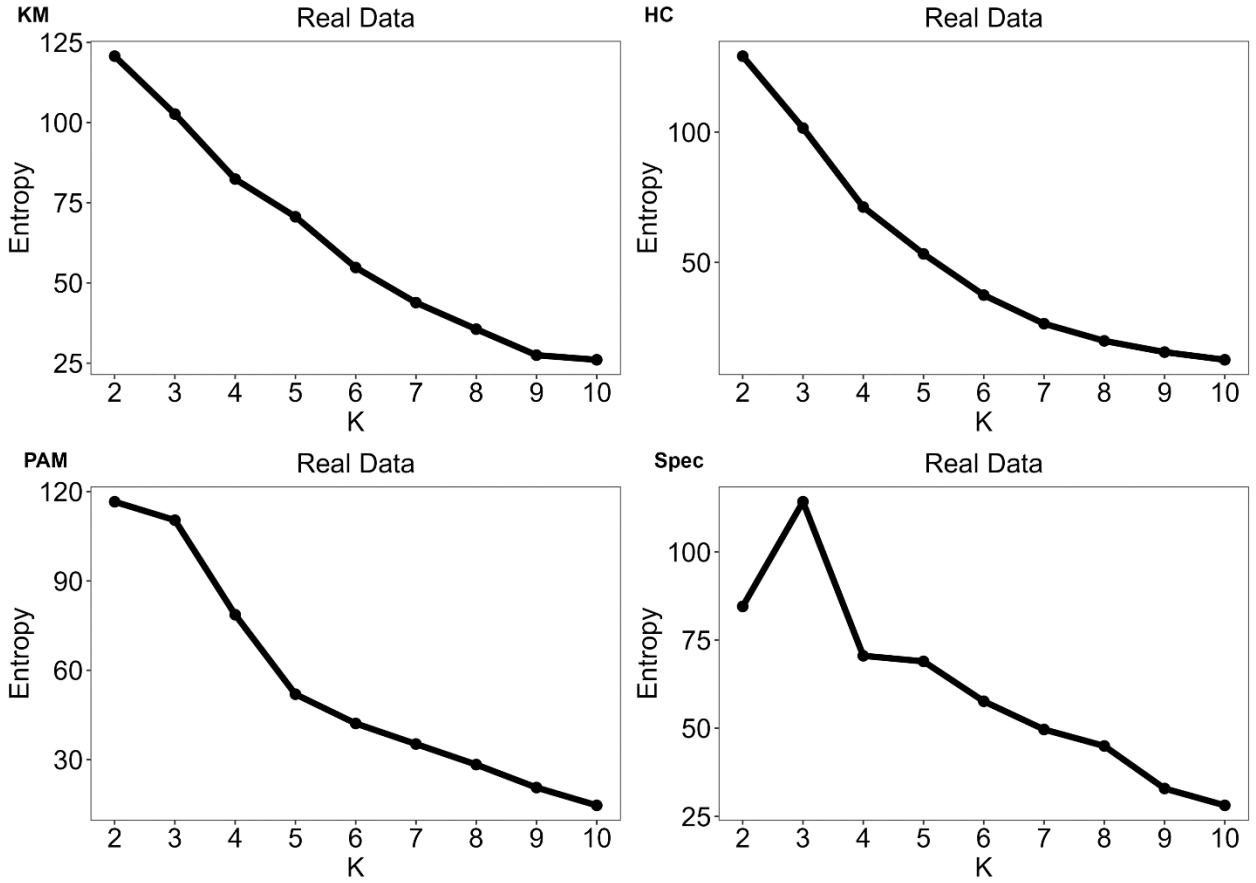


Figure 32: Entropy plots of the PRECISESADS data from Results Chapter 3 clustered using different techniques and k values. KM = Kmeans, HC = Hierarchical clustering, PAM = Partitioning around medoids, Spec = Spectral clustering.

21 APPENDIX C

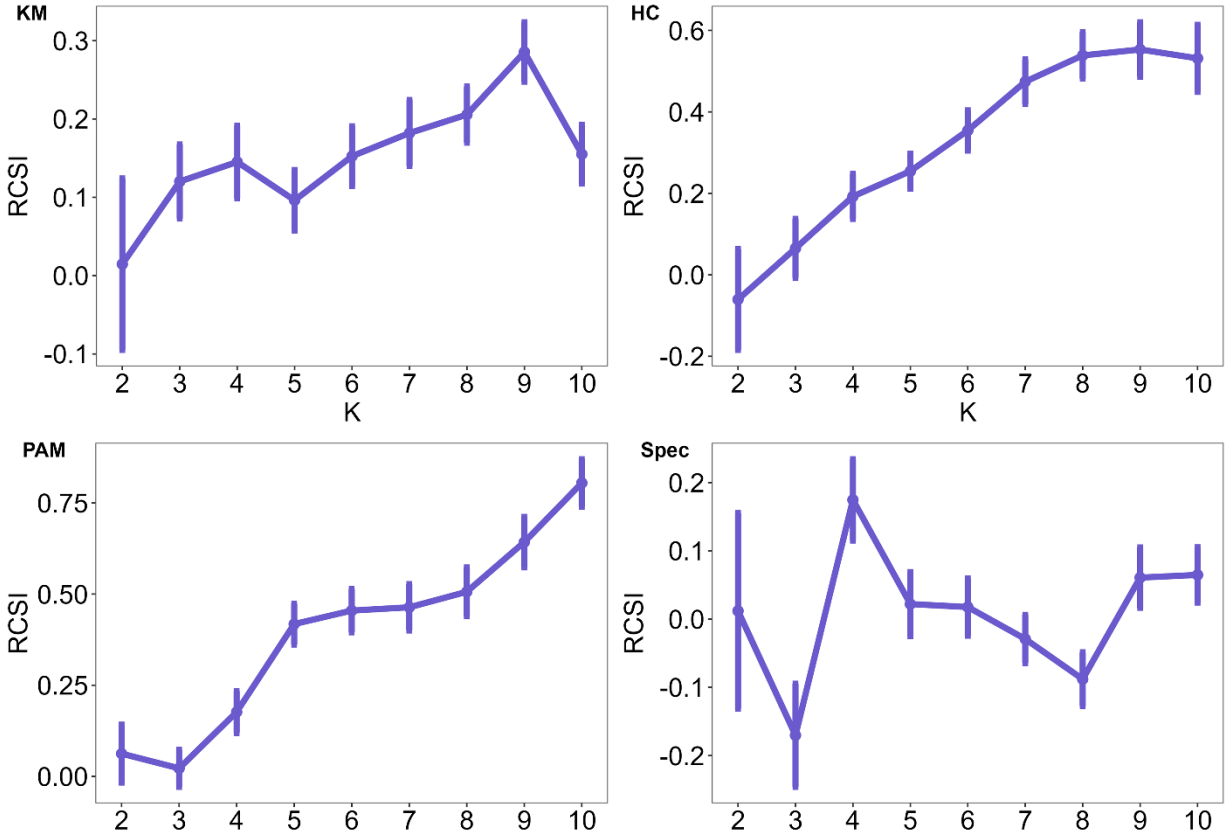


Figure 33: Relative cluster stability index plots of the PRECISESADS data from Results Chapter 3 clustered using different techniques and k values. RCSI = Relative cluster stability index, KM = Kmeans, HC = Hierarchical clustering, PAM = Partitioning around medoids, Spec = Spectral clustering.

22 BIBLIOGRAPHY

1. Reeves WH, Zhuang H, Han S. Autoantibodies in Systemic Lupus Erythematosus. In: Hochberg MC, Gravallese EM, Smolen JS, van der Heijde D, Weinblatt ME, Weisman MH, editors. Rheumatology. 8th ed. Philadelphia: Elsevier; 2023. p. 1198–204.

.

2. Irure-Ventura J, López-Hoyos M. The Past, Present, and Future in Antinuclear Antibodies (ANA). Diagnostics [Internet]. 2022;12(3). Available from: https://www.mdpi.com/2075-4418/12/3/647

3. Bossuyt X, De Langhe E, Borghi MO, Meroni PL. Understanding and interpreting antinuclear antibody tests in systemic rheumatic diseases. Nature Reviews Rheumatology 2020 16:12 [Internet]. 2020 Nov 5 [cited 2024 Sep 27];16(12):715–26. Available from: https://www.nature.com/articles/s41584-020-00522-w

4. Barber MRW, Drenkard C, Falasinnu T, Hoi A, Mak A, Kow NY, et al. Global epidemiology of systemic lupus erythematosus. Nature Reviews Rheumatology 2021 17:9 [Internet]. 2021 Aug 3 [cited 2024 Sep 27];17(9):515–32. Available from: https://www.nature.com/articles/s41584-021-00668-1

5. Dall'Era M, Cisternas MG, Snipes K, Herrinton LJ, Gordon C, Helmick CG. The Incidence and Prevalence of Systemic Lupus Erythematosus in San Francisco County, California: The California Lupus Surveillance Project. Arthritis & Rheumatology [Internet]. 2017 Oct 1 [cited 2024 Sep 27];69(10):1996–2005. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.40191

6. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Ann Rheum Dis. 2019 Sep;78(9):1151–9.

7. Crow MK, Mary Crow PK. Pathogenesis of systemic lupus erythematosus: risks, mechanisms and therapeutic targets. Ann Rheum Dis [Internet]. 2023 Aug 1 [cited 2024 Sep 27];82(8):999–1014. Available from: https://ard.bmj.com/content/82/8/999

8. Tanaka Y, Kusuda M, Yamaguchi Y. Interferons and systemic lupus erythematosus: Pathogenesis, clinical features, and treatments in interferon-driven disease. Mod Rheumatol [Internet]. 2023 [cited 2024 Sep 30];33:857–67. Available from: https://doi.org/10.1093/mr/roac140

9. Hile GA, Michelle Kahlenberg J. Immunopathogenesis of skin injury in SLE. Curr Opin Rheumatol [Internet]. 2021 Mar 3 [cited 2024 Sep 30];33(2):173. Available from: https://pubmed.ncbi.nlm.nih.gov/33315653/

10. Iwamoto T, Dorschner JM, Selvaraj S, Mezzano V, Jensen MA, Vsetecka D, et al. High Systemic Type I Interferon Activity Is Associated With Active Class III/IV Lupus Nephritis. J Rheumatol [Internet]. 2022 Apr 1 [cited 2024 Sep 30];49(4):388–97. Available from: https://pubmed.ncbi.nlm.nih.gov/34782453/

11. Shiozawa S, Kuroki Y, Kim M, Hirohata S, Ogino T. Interferon-alpha in lupus psychosis. Arthritis Rheum [Internet]. 1992 [cited 2024 Sep 30];35(4):417–22. Available from: https://pubmed.ncbi.nlm.nih.gov/1373622/

12. Crow YJ, Manel N. Aicardi–Goutières syndrome and the type I interferonopathies. Nature Reviews Immunology 2015 15:7 [Internet]. 2015 Jun 5 [cited 2024 Sep 30];15(7):429–40. Available from: https://www.nature.com/articles/nri3850

13. Panousis nikolaos, Bertsias GK, ongen H, Gergianaki irini, Tektonidou MG, Trachana M, et al. Combined genetic and transcriptome analysis of patients with SLE: distinct, targetable signatures for susceptibility and severity. Ann Rheum Dis [Internet]. 2019 [cited 2024 Oct 23];78:1079–89. Available from: http://dx.doi.org/10.1136/annrheumdis-2018-214379

14. Zen M, Salmaso L, Barbiellini Amidei C, Fedeli U, Bellio S, Iaccarino L, et al. Mortality and causes of death in systemic lupus erythematosus over the last decade: Data from a large population-based study. Eur J Intern Med [Internet]. 2023 Jun 1 [cited 2024 Sep 30];112:45–51. Available from: http://www.ejinme.com/article/S0953620523000377/fulltext

15. Azizoddin DR, Jolly M, Arora S, Yelin E, Katz P. Patient-Reported Outcomes Predict Mortality in Lupus. Arthritis Care Res (Hoboken) [Internet]. 2019 Aug 1 [cited 2024 Sep 30];71(8):1028–35. Available from: https://pubmed.ncbi.nlm.nih.gov/30144293/

16. Nikoloudaki M, Repa A, Pitsigavdaki S, Molla Ismail Sali A, Sidiropoulos P, Lionis C, et al. Persistence of Depression and Anxiety despite Short-Term Disease Activity Improvement in Patients with Systemic Lupus Erythematosus: A Single-Centre, Prospective Study. J Clin Med [Internet]. 2022 Aug 1 [cited 2024 Oct 23];11(15):4316. Available from: https://www.mdpi.com/2077-0383/11/15/4316/htm

17. Ghafoor M. Sjögren's Before Sjögren: Did Henrik Sjögren (1899–1986) Really Discover Sjögren's Disease? J Maxillofac Oral Surg [Internet]. 2012 Sep 1 [cited 2024 Sep 26];11(3):373. Available from: https://pubmed.ncbi.nlm.nih.gov/23997498/

18. Anderson JR, Gray KG, Beck JS, Kinnear WF. Precipitating autoantibodies in Sjogren's disease. Lancet [Internet]. 1961 Aug 26 [cited 2024 Sep 26];2(7200):456–60. Available from: https://pubmed.ncbi.nlm.nih.gov/13683222/

19. Moutsopoulos HM. Sjögren's syndrome: a forty-year scientific journey. J Autoimmun [Internet]. 2014 [cited 2024 Sep 26];51:1–9. Available from: https://pubmed.ncbi.nlm.nih.gov/24485155/

20. Vitali C, Del Papa N. Pain in primary Sjögren's syndrome. Best Pract Res Clin Rheumatol. 2015 Feb;29(1):63–70.

21. Marketos N, Cinoku I, Rapti A, Mavragani CP. Type I interferon signature in Sjögren's syndrome: Pathophysiological and clinical implications. Clin Exp Rheumatol [Internet]. 2019 [cited 2024 Sep 26];37:S185–91. Available from: https://www.clinexprheumatol.org/abstract.asp?a=14416

22. Hooks JJ, Moutsopoulos HM, Geis SA, Stahl NI, Decker JL, Notkins AL. Immune interferon in the circulation of patients with autoimmune disease. N Engl J Med [Internet]. 1979 Jul 5 [cited 2024 Sep 26];301(1):5–8. Available from: https://pubmed.ncbi.nlm.nih.gov/449915/

23. Gottenberg JE, Cagnard N, Lucchesi C, Letourneur F, Mistou S, Lazure T, et al. Activation of IFN pathways and plasmacytoid dendritic cell recruitment in target organs of primary Sjögren's syndrome. Proc Natl Acad Sci U S A [Internet]. 2006 Feb 21 [cited 2024 Sep 26];103(8):2770–5. Available from: https://pubmed.ncbi.nlm.nih.gov/16477017/

24. Kimoto O, Sawada J, Shimoyama K, Suzuki D, Nakamura S, Hayashi H, et al. Activation of the interferon pathway in peripheral blood of patients with Sjogren's syndrome. J Rheumatol [Internet]. 2011 Feb 1 [cited 2024 Sep 26];38(2):310–6. Available from: https://pubmed.ncbi.nlm.nih.gov/21078725/

25. Emamian ES, Leon JM, Lessard CJ, Grandits M, Baechler EC, Gaffney PM, et al. Peripheral blood gene expression profiling in Sjögren's syndrome. Genes Immun [Internet]. 2009 [cited 2024 Sep 26];10(4):285–96. Available from: https://pubmed.ncbi.nlm.nih.gov/19404300/

26. Bodewes ILA, Al-Ali S, van Helden-Meeuwsen CG, Maria NI, Tarn J, Lendrem DW, et al. Systemic interferon type I and type II signatures in primary Sjögren's syndrome reveal differences in biological disease activity. Rheumatology (Oxford) [Internet]. 2018 [cited 2024 Sep 26];57(5):921–30. Available from: https://pubmed.ncbi.nlm.nih.gov/29474655/

27. Yang HT, Hong XP, Guo JW, Zhong XL, Liao R, Liu CL, et al. Clinical Significance of Different Profiles of anti-Ro Antibodies in Connective Tissue Diseases. J Immunol Res [Internet]. 2023 [cited 2024 Sep 26];2023. Available from: https://pubmed.ncbi.nlm.nih.gov/36741231/

28. Wangkaew S, Kasitanon N, Sivasomboon C, Wichainun R, Sukitawut W, Louthrenoo W. Sicca Symptoms in Thai Patients with Rheumatoid Arthritis, Systemic Lupus Erythematosus and Scleroderma: A Comparison with Age-Matched Controls and Correlation with Disease Variables. Asian Pac J Allergy Immunol. 2006;24:213–21.

29. Luciano N, Baldini C, Tarantini G, Ferro F, Sernissi F, Varanini V, et al. Ultrasonography of major salivary glands: a highly specific tool for distinguishing primary Sjögren's syndrome from undifferentiated connective tissue diseases. Rheumatology (Oxford) [Internet]. 2015 [cited 2024 Sep 26];54(12):2198–204. Available from: https://pubmed.ncbi.nlm.nih.gov/26206346/

30. Ramos-Casals M, Brito-Zerón P, Bombardieri S, Bootsma H, De Vita S, Dörner T, et al. EULAR recommendations for the management of Sjögren's syndrome with topical and systemic therapies. Ann Rheum Dis [Internet]. 2020 Jan 1 [cited 2024 Sep 26];79(1):3–18. Available from: https://ard.bmj.com/content/79/1/3

31. Abrol E, González-Pulido C, Praena-Fernández JM, Isenberg DA. A retrospective study of long-term outcomes in 152 patients with primary Sjögren's syndrome: 25-year experience. Clinical Medicine [Internet]. 2014 Mar 1 [cited 2024 Sep 26];14(2):157. Available from: https://pubmed.ncbi.nlm.nih.gov/24715127//

32. West SG, Kolfenbach J. Rheumatology Secrets. 4th ed. Philadelphia: Elsevier; 2020.

33. De Vooght J, Vulsteke JB, De Haes P, Bossuyt X, Lories R, De Langhe E. Anti-TIF1-γ autoantibodies: warning lights of a tumour autoantigen. Rheumatology (Oxford) [Internet]. 2020 Mar 1 [cited 2024 Sep 26];59(3):469–77. Available from: https://pubmed.ncbi.nlm.nih.gov/31883334/

34. Lundberg IE, Tjärnlund A, Bottai M, Werth VP, Pilkington C, de Visser M, et al. 2017 European League Against Rheumatism/American College of Rheumatology Classification Criteria for Adult and Juvenile Idiopathic Inflammatory Myopathies and Their Major Subgroups. Arthritis and Rheumatology [Internet]. 2017 Dec 1 [cited 2024 Sep 26];69(12):2271–82. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.40320

35. Lilleker JB, Chinoy H. Can machine learning unravel the complex IIM spectrum? Nature Reviews Rheumatology 2020 16:6 [Internet]. 2020 Mar 27 [cited 2024 Sep 26];16(6):299–300. Available from: https://www.nature.com/articles/s41584-020-0412-6

36. Oldroyd A, Lilleker J, Chinoy H. Idiopathic inflammatory myopathies – a guide to subtypes, diagnostic approach and treatment. Clinical Medicine. 2017 Aug 1;17(4):322–8.

37. Bohan Anthony. Polymyositis and dermatomyositis (second of two parts). N Engl J Med [Internet]. 1975 [cited 2024 Sep 26];292:403–7. Available from: https://cir.nii.ac.jp/crid/1572261550479025792

38. Horiki T, Fuyuno G, Ishii M, Sasao T, Shibuya M, Yanagimachi N, et al. Polymyositis and dermatomyositis (first of twoparts). N Engl J Med [Internet]. 1975 [cited 2024 Sep 26];292(6):344-. Available from: https://cir.nii.ac.jp/crid/1572824499452267136

39. Giannini M, Ellezam B, Leclair V, Lefebvre F, Troyanov Y, Hudson M, et al. Scleromyositis: A distinct novel entity within the systemic sclerosis and autoimmune myositis spectrum. Implications for care and pathogenesis. Front Immunol [Internet]. 2023 Jan 26 [cited 2024 Sep 26];13. Available from: https://pubmed.ncbi.nlm.nih.gov/36776390/

40. Carter LM, Gordon C, Yee CS, Bruce I, Isenberg D, Skeoch S, et al. Easy-BILAG: a new tool for simplified recording of SLE disease activity using BILAG-2004 index.

Rheumatology (Oxford) [Internet]. 2022 Oct 6 [cited 2023 Jan 17];61(10):4006. Available from: https://pubmed.ncbi.nlm.nih.gov/35077529/

41. Joy GM, Arbiv OA, Wong CK, Lok SD, Adderley NA, Dobosz KM, et al. Prevalence, imaging patterns and risk factors of interstitial lung disease in connective tissue disease: a systematic review and meta-analysis. Eur Respir Rev [Internet]. 2023 Mar 1 [cited 2024 Sep 26];32(167). Available from: https://pubmed.ncbi.nlm.nih.gov/36889782/

42. Isenberg DA, Rowe D, Shearer M, Novick D, Beverley PCL. Localization of interferons and interleukin 2 in polymyositis and muscular dystrophy. Clin Exp Immunol [Internet]. 1986 [cited 2024 Sep 26];63(2):450. Available from: https://pubmed.ncbi.nlm.nih.gov/2421951/

43. Li M, Zhang Y, Zhang W, Sun J, Liu R, Pan Z, et al. Type 1 interferon signature in peripheral blood mononuclear cells and monocytes of idiopathic inflammatory myopathy patients with different myositis-specific autoantibodies. Front Immunol. 2023 May 9;14:1169057.

44. Wong D, Kea B, Pesich R, Higgs BW, Zhu W, Brown P, et al. Interferon and biologic signatures in dermatomyositis skin: specificity and heterogeneity across diseases. PLoS One [Internet]. 2012 Jan 3 [cited 2024 Sep 26];7(1). Available from: https://pubmed.ncbi.nlm.nih.gov/22235269/

45. Meyer A, Laverny G, Allenbach Y, Grelet E, Ueberschlag V, Echaniz-Laguna A, et al. IFN-β-induced reactive oxygen species and mitochondrial damage contribute to muscle impairment and inflammation maintenance in dermatomyositis. Acta Neuropathol [Internet]. 2017 Oct 1 [cited 2024 Sep 26];134(4):655–66. Available from: https://pubmed.ncbi.nlm.nih.gov/28623559/

46. Ladislau L, Suárez-Calvet X, Toquet S, Landon-Cardinal O, Amelin D, Depp M, et al. JAK inhibitor improves type I interferon induced damage: proof of concept in

dermatomyositis. Brain [Internet]. 2018 Jun 1 [cited 2024 Sep 26];141(6):1609–21. Available from: https://pubmed.ncbi.nlm.nih.gov/29741608/

47. Bolko L, Jiang W, Tawara N, Landon-Cardinal O, Anquetil C, Benveniste O, et al. The role of interferons type I, II and III in myositis: A review. Brain Pathology. 2021 May 1;31(3).

48. MYOJAK - Health Research Authority [Internet]. [cited 2024 Sep 26]. Available from: https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/myojak/

49. Oldroyd AGS, Lilleker JB, Amin T, Aragon O, Bechman K, Cuthbert V, et al. British Society for Rheumatology guideline on management of paediatric, adolescent and adult patients with idiopathic inflammatory myopathy. Rheumatology [Internet]. 2022 May 5 [cited 2024 Sep 26];61(5):1760–8. Available from: https://dx.doi.org/10.1093/rheumatology/keac115

50. Dobloug GC, Svensson J, Lundberg IE, Holmqvist M. Mortality in idiopathic inflammatory myopathy: results from a Swedish nationwide population-based cohort study. Ann Rheum Dis [Internet]. 2018 Jan 1 [cited 2024 Sep 26];77(1):40–7. Available from: https://pubmed.ncbi.nlm.nih.gov/28814428/

51. Lundberg IE, Fujimoto M, Vencovsky J, Aggarwal R, Holmqvist M, Christopher-Stine L, et al. Idiopathic inflammatory myopathies. Nature Reviews Disease Primers 2021 7:1 [Internet]. 2021 Dec 2 [cited 2024 Sep 26];7(1):1–22. Available from: https://www.nature.com/articles/s41572-021-00321-x

52. Nikolopoulos D, Loukogiannaki C, Sentis G, Garantziotis P, Manolakou T, Kapsala N, et al. OP0139 The molecular landscape of antiphospholipid syndrome in systemic lupus erythematosus: distinct transcriptional profiles in venous versus arterial events, with potential therapeutic implications. 2024 Jun;23.2-24.

53. Barbhaiya M, Zuily S, Naden R, Hendry A, Manneville F, Amigo MC, et al. 2023 ACR/EULAR antiphospholipid syndrome classification criteria. Ann Rheum Dis [Internet]. 2023 Oct 1 [cited 2024 Sep 26];82(10):1258–70. Available from: https://ard.bmj.com/content/82/10/1258

54. Clunie GPR, Wilkinson N, Nikiphorou E, Jadon D, editors. Oxford Handbook of Rheumatology. Oxford University Press; 2018.

55. Palli E, Kravvariti E, Tektonidou MG. Type I Interferon Signature in Primary Antiphospholipid Syndrome: Clinical and Laboratory Associations. Front Immunol [Internet]. 2019 [cited 2024 Sep 26];10(MAR):487. Available from: https://pubmed.ncbi.nlm.nih.gov/30930907/

56. Rosa dos Santos AP, de Oliveira Vaz C, Hounkpe BW, Jacintho BC, Oliveira JD, Tripiquia Vechiatto Mesquita GL, et al. Association between interferon-I producing plasmacytoid dendritic cells and thrombotic antiphospholipid syndrome. https://doi.org/101177/09612033221101731 [Internet]. 2022 May 25 [cited 2024 Sep 26];31(9):1067–77. Available from: https://journals.sagepub.com/doi/abs/10.1177/09612033221101731?casa_token=6JKsLYRTZWUAAAAA%3AH_kMFWvlqXoez4iNa18VBpatplyiRhry0szubuZ7ZIhlYSqFgn88s72bpX15_OAaiFkvrcEdV6Xw&journalCode=lupa

57. Grenn RC, Yalavarthi S, Gandhi AA, Kazzaz NM, Núñez-Álvarez C, Hernández-Ramírez D, et al. Endothelial progenitor dysfunction associates with a type I interferon signature in primary antiphospholipid syndrome. Ann Rheum Dis [Internet]. 2017 Feb 1 [cited 2024 Sep 26];76(2):450–7. Available from: https://ard.bmj.com/content/76/2/450

58. Meier FMP, Frommer KW, Dinser R, Walker UA, Czirjak L, Denton CP, et al. Update on the profile of the EUSTAR cohort: an analysis of the EULAR Scleroderma Trials and Research group database. Ann Rheum Dis. 2012 Aug;71(8):1355–60.

59. Rosendahl AH, Schönborn K, Krieg T. Pathophysiology of systemic sclerosis (scleroderma). Kaohsiung J Med Sci [Internet]. 2022 Mar 1 [cited 2024 Sep 27];38(3):187–95. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/kjm2.12505

60. George PM, Oliver E, Dorfmuller P, Dubois OD, Reed DM, Kirkby NS, et al. Evidence for the involvement of type I interferon in pulmonary arterial hypertension. Circ Res. 2014 Feb 14;114(4):677–88.

61. Liu X, Mayes MD, Tan FK, Wu M, Reveille JD, Harper BE, et al. Correlation of interferon-inducible chemokine plasma levels with disease severity in systemic sclerosis. Arthritis Rheum. 2013 Jan;65(1):226–35.

62. Eloranta ML, Franck-Larsson K, Lövgren T, Kalamajski S, Rönnblom A, Rubin K, et al. Type I interferon system activation and association with disease manifestations in systemic sclerosis. Ann Rheum Dis. 2010 Jul;69(7):1396–402.

63. Kim D, Peck A, Santer D, Patole P, Schwartz SM, Molitor JA, et al. Induction of interferon-alpha by scleroderma sera containing autoantibodies to topoisomerase I: association of higher interferon-alpha activity with lung fibrosis. Arthritis Rheum. 2008 Jul;58(7):2163–73.

64. Kakkar V, Assassi S, Allanore Y, Kuwana M, Denton CP, Khanna D, et al. Type 1 interferon activation in systemic sclerosis: a biomarker, a target or the culprit. Curr Opin Rheumatol [Internet]. 2022 Nov 1 [cited 2024 Sep 27];34(6):357. Available from: https://pubmed.ncbi.nlm.nih.gov/36125916/

65. Farina G, Lafyatis D, Lemaire R, Lafyatis R. A four-gene biomarker predicts skin disease in patients with diffuse cutaneous systemic sclerosis. Arthritis Rheum. 2010 Feb;62(2):580–8.

66. Christmann RB, Sampaio-Barros P, Stifano G, Borges CL, de Carvalho CR, Kairalla R, et al. Association of Interferon- and transforming growth factor β-regulated genes and macrophage activation with systemic sclerosis-related progressive lung fibrosis. Arthritis Rheumatol. 2014 Mar;66(3):714–25.

67. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 Classification Criteria for Systemic Sclerosis: An American College of Rheumatology/European League Against Rheumatism Collaborative Initiative. Arthritis Rheum. 2013 Nov 3;65(11):2737–47.

68. Poudel DR, Jayakumar D, Danve A, Sehra ST, Derk CT. Determinants of mortality in systemic sclerosis: a focused review. Rheumatol Int [Internet]. 2018 Oct 1 [cited 2024 Sep 27];38(10):1847–58. Available from: https://link.springer.com/article/10.1007/s00296-017-3826-y

69. Rubio-Rivas M, Royo C, Simeón CP, Corbella X, Fonollosa V. Mortality and survival in systemic sclerosis: systematic review and meta-analysis. Semin Arthritis Rheum. 2014 Oct;44(2):208–19.

70. Sharp GC, Irvin WS, Tan EM, Gould RG, Holman HR. Mixed connective tissue disease--an apparently distinct rheumatic disease syndrome associated with a specific antibody to an extractable nuclear antigen (ENA). Am J Med [Internet]. 1972 [cited 2024 Sep 26];52(2):148–59. Available from: https://pubmed.ncbi.nlm.nih.gov/4621694/

71. Alves MR, Isenberg DA. “Mixed connective tissue disease”: a condition in search of an identity. Clin Exp Med [Internet]. 2020 May 1 [cited 2024 Sep 26];20(2):159–66. Available from: https://pubmed.ncbi.nlm.nih.gov/32130548/

72. Aringer M, Steiner G, Smolen JS. Does mixed connective tissue disease exist? Yes. Rheum Dis Clin North Am [Internet]. 2005 Aug [cited 2024 Sep 26];31(3):411–20. Available from: https://pubmed.ncbi.nlm.nih.gov/16084315/

73. Burdt MA, Hoffman RW, Deutscher SL, Wang GS, Johnson JC, Sharp GC. Long-term outcome in mixed connective tissue disease: longitudinal clinical and serologic findings. Arthritis Rheum. 1999 May;42(5):899–909.

74. Nimelstein SH, Brody S, McShane D, Holman HR. Mixed connective tissue disease: a subsequent evaluation of the original 25 patients. Medicine. 1980 Jul;59(4):239–48.

75. Flåm ST, Gunnarsson R, Garen T, Norwegian MCTD Study Group, Lie BA, Molberg Ø. The HLA profiles of mixed connective tissue disease differ distinctly from the profiles of clinically related connective tissue diseases. Rheumatology (Oxford). 2015 Mar;54(3):528–35.

76. John KJ, Sadiq M, George T, Gunasekaran K, Francis N, Rajadurai E, et al. Clinical and Immunological Profile of Mixed Connective Tissue Disease and a Comparison of Four Diagnostic Criteria. Int J Rheumatol [Internet]. 2020 [cited 2024 Oct 7];2020. Available from: https://pubmed.ncbi.nlm.nih.gov/32411251/

77. John KJ, Sadiq M, George T, Gunasekaran K, Francis N, Rajadurai E, et al. Clinical and Immunological Profile of Mixed Connective Tissue Disease and a Comparison of Four Diagnostic Criteria. Int J Rheumatol [Internet]. 2020 [cited 2024 Sep 26];2020. Available from: https://pubmed.ncbi.nlm.nih.gov/32411251/

78. Williams EL, Gadola S, Edwards CJ. Anti-TNF-induced lupus. Rheumatology [Internet]. 2009 Jul 1 [cited 2024 Sep 26];48(7):716–20. Available from: https://dx.doi.org/10.1093/rheumatology/kep080

79. Cappelli S, Bellando Randone S, Martinović D, Tamas MM, Pasalić K, Allanore Y, et al. “To be or not to be,” ten years after: evidence for mixed connective tissue disease as a distinct entity. Semin Arthritis Rheum. 2012 Feb;41(4):589–98.

80. Burdt MA, Hoffman RW, Deutscher SL, Wang GS, Johnson JC, Sharp GC. Long-term outcome in mixed connective tissue disease: longitudinal clinical and serologic findings. Arthritis Rheum. 1999 May;42(5):899–909.

81. Hassan AB, Hozayen RF, Mustafa ZS, Lundberg IE, Jahrami HA. The prevalence of pulmonary arterial hypertension in patients with mixed connective tissue disease: a systematic review and meta-analysis. Clin Exp Rheumatol [Internet]. 2023 Nov 1 [cited 2024 Sep 27];41(11):2301–11. Available from: https://pubmed.ncbi.nlm.nih.gov/37497727/

82. Hajas A, Szodoray P, Nakken B, Gaal J, Zöld E, Laczik R, et al. Clinical course, prognosis, and causes of death in mixed connective tissue disease. J Rheumatol. 2013 Jul;40(7):1134–42.

83. Rubio J, Kyttaris VC. Undifferentiated Connective Tissue Disease: Comprehensive Review. Curr Rheumatol Rep [Internet]. 2023 May 1 [cited 2024 Sep 27];25(5):98–106. Available from: https://link.springer.com/article/10.1007/s11926-023-01099-5

84. Mosca M, Tani C, Vagnani S, Carli L, Bombardieri S. The diagnosis and classification of undifferentiated connective tissue diseases. J Autoimmun. 2014;48–49:50–2.

85. Drehmel KR, Erickson AR, England BR, Michaud KD, Sayles HR, Hearth-Holmes MP. Applying SLICC and ACR/EULAR systemic lupus erythematosus classification criteria in a cohort of patients with undifferentiated connective tissue disease. Lupus [Internet]. 2021 Feb 1 [cited 2024 Sep 27];30(2):280–4. Available from: https://journals.sagepub.com/doi/10.1177/0961203320976939

86. Erden A, Apaydln H, Fanouriakis A, Güven SC, Armagan B, Akyüz Dağll P, et al. Performance of the systemic lupus erythematosus risk probability index in a cohort of undifferentiated connective tissue disease. Rheumatology [Internet]. 2022 Aug 29 [cited 2024 Oct 23];61(9):3606–13. Available from: https://dx.doi.org/10.1093/rheumatology/keac005

87. Adamichou C, Genitsaridi I, Nikolopoulos D, Nikoloudaki M, Repa A, Bortoluzzi A, et al. Lupus or not? SLE Risk Probability Index (SLERPI): a simple, clinician-friendly machine learning-based model to assist the diagnosis of systemic lupus erythematosus. Ann Rheum Dis [Internet]. 2021 Jun 1 [cited 2024 Oct 23];80(6):758–66. Available from: https://ard.bmj.com/content/80/6/758

88. Radin M, Rubini E, Cecchi I, Foddai SG, Barinotti A, Rossi D, et al. Disease evolution in a long-term follow-up of 104 undifferentiated connective tissue disease patients. Clin Exp Rheumatol. 2022 Mar 1;40(3):575–80.

89. Guerrero LF, Rueda JC, Arciniegas R, Rueda JM. Undifferentiated connective tissue disease in a rheumatology center in Cali, Colombia: clinical features of 94 patients followed for a year. Rheumatol Int. 2013 Apr 25;33(4):1085–8.

90. Mosca M, Tani C, Carli L, Della Rossa A, Talarico R, Baldini C, et al. Analysis of the evolution of UCTD to defined CTD after a long term follow-up. Clin Exp Rheumatol. 2013;31(3):471.

91. Tani C, Trentin F, Parma A, Zucchi D, Cardelli C, Stagnaro C, et al. Disease evolution and organ damage accrual in patients with stable UCTD: a long-term monocentric inception cohort. RMD Open [Internet]. 2024 Apr 25 [cited 2024 Sep 27];10(2). Available from: https://pubmed.ncbi.nlm.nih.gov/38670557/

92. Wandstrat AE, Carr-Johnson F, Branch V, Gray H, Fairhurst AM, Reimold A, et al. Autoantibody profiling to identify individuals at risk for systemic lupus erythematosus. J Autoimmun [Internet]. 2006 Nov [cited 2024 Sep 5];27(3):153–60. Available from: https://pubmed.ncbi.nlm.nih.gov/17052888/

93. Fernandez SA ssuncion V, Lobo AZ oghbi C, Oliveira ZN ajjar P de, Fukumori LM aria I, P rigo AM arques, Rivitti EA. Prevalence of antinuclear autoantibodies in the serum of normal blood dornors. Rev Hosp Clin Fac Med Sao Paulo [Internet]. 2003 Nov 1

[cited 2024 Sep 5];58(6):315–9. Available from: https://pubmed.ncbi.nlm.nih.gov/14762490/

94. Chae DH, Martz CD, Fuller-Rowell TE, Spears EC, Smith TTG, Hunter EA, et al. Racial Discrimination, Disease Activity, and Organ Damage: The Black Women's Experiences Living With Lupus (BeWELL) Study. Am J Epidemiol [Internet]. 2019 Aug 1 [cited 2024 Sep 5];188(8):1434. Available from: https://pubmed.ncbi.nlm.nih.gov/31062841/

95. Yusof MYM, Psarras A, El-Sherbiny YM, Hensor EMA, Dutton K, Ul-Hassan S, et al. Prediction of autoimmune connective tissue disease in an at-risk cohort: prognostic value of a novel two-score system for interferon status. Ann Rheum Dis [Internet]. 2018 Oct 1 [cited 2024 Sep 17];77(10):1432–9. Available from: https://ard.bmj.com/content/77/10/1432

96. Kim ST, Muñoz-Grajales C, Dunn SE, Schneider R, Johnson SR, Touma Z, et al. Interferon and interferon-induced cytokines as markers of impending clinical progression in ANA+ individuals without a systemic autoimmune rheumatic disease diagnosis. Arthritis Res Ther [Internet]. 2023 Dec 1 [cited 2024 Sep 26];25(1):21. Available from: https://pubmed.ncbi.nlm.nih.gov/36765391/

97. Fanouriakis, Antonis et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Ann Rheum Dis. 2019 Jun 1;78(6):736–45.

98. Md Yusof, Md Yuzaiful et al. Predicting Severe Infection and Effects of Hypogammaglobulinemia During Therapy With Rituximab in Rheumatic and Musculoskeletal Diseases. Arthritis Rheumatol. 2019 Nov 1;71(11):1812–23.

99. Gaubitz M. Epidemiology of connective tissue disorders. Rheumatology. 2006 Oct 1;45(suppl_3):iii3–4.

100. Isabelle Marie. Morbidity and mortality in adult polymyositis and dermatomyositis. Curr Rheumatol Rep. 2012 Jun;14(3):275–85.

101. Rubio-Rivas, Manuel et al . Mortality and survival in systemic sclerosis: systematic review and meta-analysis. Semin Arthritis Rheum. 2014 Oct 1;44(2):208–19.

102. D'Silva KM, Li L, Lu N, Ogdie A, Avina-Zubieta JA, Choi HK. Persistent premature mortality gap in dermatomyositis and polymyositis: a United Kingdom general population-based cohort study. Rheumatology. 2021 Jun 18;60(6):2653–60.

103. Haque S, Gordon C, Isenberg D, Rahman A, Lanyon P, Bell A, et al. Risk factors for clinical coronary heart disease in systemic lupus erythematosus: The Lupus and Atherosclerosis Evaluation of Risk (LASER) study. Journal of Rheumatology. 2010;37(2):322–9.

104. Manzi, S et al. Age-specific incidence rates of myocardial infarction and angina in women with systemic lupus erythematosus: comparison with the Framingham Study. Am J Epidemiol. 1997;145(5):408–15.

105. Khamashta, M A et al. The cost of care of systemic lupus erythematosus (SLE) in the UK: annual direct costs for adult SLE patients with active autoantibody-positive disease. Lupus. 2014 Mar;23(3):273–83.

106. Sutcliffe, N et al. Total costs and predictors of costs in patients with systemic lupus erythematosus. Rheumatology (Oxford). 2001;40(1):37–47.

107. Meacock, Rachel et al . The humanistic and economic burden of systemic lupus erythematosus : a systematic review. Pharmacoeconomics. 2013 Jan;31(1):49–61.

108. Katz, Patricia et al. Patient-reported flare frequency is associated with diminished quality of life and family role functioning in systemic lupus erythematosus. Qual Life Res. 2020 Dec 1;29(12):3251–61.

109. Almehed K, Carlsten H, Forsblad-D'Elia H. Health-related quality of life in systemic lupus erythematosus and its association with disease and work disability. Scand J Rheumatol. 2010;39(1):58–62.

110. Reddy V, Jayne D, Close D, Isenberg D. B-cell depletion in SLE: Clinical and trial experience with rituximab and ocrelizumab and implications for study design. Vol. 15, Arthritis Research and Therapy. BioMed Central; 2013. p. S2.

111. Cornet A, Andersen J, Myllys K, Edwards A, Arnaud L. Living with systemic lupus erythematosus in 2020: a European patient survey. Lupus Sci Med. 2021 Apr 1;8(1):e000469.

112. Zayat AS, Mahmoud K, Md Yusof MY, Mukherjee S, D'Agostino MA, Hensor EMA, et al. Defining inflammatory musculoskeletal manifestations in systemic lupus erythematosus. Rheumatology (Oxford). 2019 Feb 1;58(2):304–12.

113. Mahmoud, Khaled et al . Ultrasound to identify SLE patients with musculoskeletal symptoms who respond best to therapy: The USEFUL longitudinal multicentre study. Rheumatology (Oxford). 2021 Apr 1;

114. Mahmoud K, Zayat A, Yusof MYM, Ciurtin C, Yee CS, Isenberg D, et al. FRI0599 USEFUL II derivation of the lupus arthritis and musculoskeletal disease activity score (lamda) using data from a multicentre longitudinal study. Ann Rheum Dis. 2020 Jun 1;79(Suppl 1):905–6.

115. Ceccarelli F, Perricone C, Cipriano E, Massaro L, Natalucci F, Capalbo G, et al. Joint involvement in systemic lupus erythematosus: From pathogenesis to clinical assessment. Semin Arthritis Rheum. 2017 Aug 1;47(1):53–64.

116. Ball EMA, Bell AL. Lupus arthritis--do we have a clinically useful classification? Rheumatology (Oxford). 2012;51(5):771–9.

117. Five-year follow-up of 665 Hungarian patients with undifferentiated connective tissue disease (UCTD) [Internet]. [cited 2022 Feb 23]. Available from: https://www.clinexprheumatol.org/abstract.asp?a=2057

118. Fauchais AL, Ouattara B, Gondran G, Lalloué F, Petit D, Ly K, et al. Articular manifestations in primary Sjögren's syndrome: clinical significance and prognosis of 188 patients. Rheumatology. 2010 Jun 1;49(6):1164–72.

119. Narváez J, Sánchez-Fernández SÁ, Seoane-Mato D, Díaz-González F, Bustabad S. Prevalence of Sjögren's syndrome in the general adult population in Spain: estimating the proportion of undiagnosed cases. Scientific Reports 2020 10:1. 2020 Jun 30;10(1):1–10.

120. Mirouse A, Seror R, Vicaut E, Mariette X, Dougados M, Fauchais AL, et al. Arthritis in primary Sjögren's syndrome: Characteristics, outcome and treatment from French multicenter retrospective study. Autoimmun Rev. 2019 Jan 1;18(1):9–14.

121. Klein M, Mann H, Vencovský J. Arthritis in Idiopathic Inflammatory Myopathies. Curr Rheumatol Rep. 2019 Dec 1;21(12):1–9.

122. Parker MJS, Oldroyd A, Roberts ME, Ollier WE, New RP, Cooper RG, et al. Increasing incidence of adult idiopathic inflammatory myopathies in the City of Salford, UK: a 10-year epidemiological study. Rheumatol Adv Pract. 2018 Jul 1;2(2):1–7.

123. Suri D, Bhattad S. Undifferentiated Connective Tissue Disease, Mixed Connective Tissue Disease, and Overlap Syndromes in Rheumatology. Mo Med. 2016 Jan 1;113(2):136.

124. Gunnarsson R, Molberg Ø, Gilboe IM, Gran JT, Lexberg ÅS, Time K, et al. The prevalence and incidence of mixed connective tissue disease: a national multicentre survey of Norwegian patients. Ann Rheum Dis. 2011 Jun;70(6):1047–51.

125. Mucke J, Alarcon-Riquelme M, Andersen J, Aringer M, Bombardieri S, Brinks R, et al. What are the topics you care about making trials in lupus more effective? Results of an Open Space meeting of international lupus experts. Lupus Sci Med [Internet]. 2021 May 20 [cited 2024 Sep 17];8(1):506. Available from: https://pubmed.ncbi.nlm.nih.gov/34016719/

126. Brown S, Navarro Coy N, Pitzalis C, Emery P, Pavitt S, Gray J, et al. The TRACTISS Protocol: A randomised double blind placebo controlled clinical TRial of Anti-B-Cell Therapy in patients with primary Sjögren's Syndrome. BMC Musculoskelet Disord [Internet]. 2014 Jan 17 [cited 2021 Jul 5];15(1). Available from: https://pubmed.ncbi.nlm.nih.gov/24438039/

127. Devauchelle-Pensec V, Mariette X, Jousse-Joulin S, Berthelot JM, Perdriger A, Puéchal X, et al. Treatment of primary Sjögren syndrome with rituximab: a randomized trial. Ann Intern Med [Internet]. 2014 Feb 18 [cited 2024 Sep 17];160(4):233–42. Available from: https://pubmed.ncbi.nlm.nih.gov/24727841/

128. Tarn JR, Howard-Tripp N, Lendrem DW, Mariette X, Saraux A, Devauchelle-Pensec V, et al. Symptom-based stratification of patients with primary Sjögren's syndrome: multi-dimensional characterisation of international observational cohorts and reanalyses of randomised clinical trials. Lancet Rheumatol. 2019;1(2):e85–94.

129. Merrill JT, Buyon JP, Furie RA, Latinis KM, Gordon C, Hsieh HJ, et al. Assessment of flares in lupus patients enrolled in a phase II/III study of rituximab (EXPLORER). Lupus [Internet]. 2011 Jun [cited 2024 Sep 17];20(7):709–16. Available from: https://pubmed.ncbi.nlm.nih.gov/21478286/

130. Rovin BH, Furie R, Latinis K, Looney RJ, Fervenza FC, Sanchez-Guerrero J, et al. Efficacy and safety of rituximab in patients with active proliferative lupus nephritis: the Lupus Nephritis Assessment with Rituximab study. Arthritis Rheum [Internet]. 2012

[cited 2024 Sep 17];64(4):1215–26. Available from: https://pubmed.ncbi.nlm.nih.gov/22231479/

131. Reddy V, Jayne D, Close D, Isenberg D. B-cell depletion in SLE: clinical and trial experience with rituximab and ocrelizumab and implications for study design. Arthritis Res Ther [Internet]. 2013 Feb 11 [cited 2024 Sep 17];15 Suppl 1(Suppl 1). Available from: https://pubmed.ncbi.nlm.nih.gov/23566295/

132. Mahieu MA, Strand V, Simon LS, Lipsky PE, Ramsey-Goldman R. A critical review of clinical trials in systemic lupus erythematosus. Lupus [Internet]. 2016 Sep 1 [cited 2024 Sep 17];25(10):1122–40. Available from: https://pubmed.ncbi.nlm.nih.gov/27497257/

133. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: The randomized, double-blind, phase ii/iii systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum [Internet]. 2010 Jan [cited 2020 Sep 1];62(1):222–33. Available from: http://doi.wiley.com/10.1002/art.27233

134. Thanou A, Chakravarty E, James JA, Merrill JT. Which outcome measures in SLE clinical trials best reflect medical judgment? Lupus Sci Med [Internet]. 2014 Jun 1 [cited 2020 Oct 27];1(1). Available from: https://pubmed.ncbi.nlm.nih.gov/25396057/

135. Connelly K, Vettivel J, Golder V, Kandane-Rathnayake R, Morand EF. Measurement of specific organ domains in lupus randomized controlled trials: a scoping review. Rheumatology (Oxford) [Internet]. 2022 Apr 1 [cited 2024 Sep 17];61(4):1341–53. Available from: https://pubmed.ncbi.nlm.nih.gov/34664636/

136. Seror R, Bowman SJ, Brito-Zeron P, Theander E, Bootsma H, Tzioufas A, et al. EULAR Sjögren's syndrome disease activity index (ESSDAI): a user guide. RMD Open [Internet]. 2015 Feb 1 [cited 2024 Sep 17];1(1):e000022. Available from: https://rmdopen.bmj.com/content/1/1/e000022

137. Fox RI, Fox CM. Sjögren Syndrome: Why Do Clinical Trials Fail? Rheumatic Disease Clinics of North America. 2016 Aug 1;42(3):519–30.

138. Previous Diagnosis of Sjögren's Syndrome As Rheumatoid Arthritis or Systemic Lupus Erythematosus - ACR Meeting Abstracts [Internet]. [cited 2024 Sep 17]. Available from: https://acrabstracts.org/abstract/previous-diagnosis-of-sjogrens-syndrome-as-rheumatoid-arthritis-or-systemic-lupus-erythematosus/

139. Mucke J, Alarcon-Riquelme M, Andersen J, Aringer M, Bombardieri S, Brinks R, et al. What are the topics you care about making trials in lupus more effective? Results of an Open Space meeting of international lupus experts. Lupus Sci Med. 2021 May 20;8(1):506.

140. Brito-Zerón P, Ramos-Casals M, Bove A, Sentis J, Font J. Predicting adverse outcomes in primary Sjögren's syndrome: identification of prognostic factors. Rheumatology. 2007 Aug 1;46(8):1359–62.

141. Goules A V., Tzioufas AG. Primary Sjögren's syndrome: clinical phenotypes, outcome and the development of biomarkers. Immunologic Research 2016 65:1. 2016 Jul 21;65(1):331–44.

142. B S, W T, T R, JM L, P H, D P, et al. Prevalence, severity, and predictors of fatigue in subjects with primary Sjögren's syndrome. Arthritis Rheum. 2008 Dec 15;59(12):1780–7.

143. Silk ME, Wallace DJ, Furie RA, Tanaka Y, Kalunian KC, Mosca M, et al. 200 Efficacy and safety of belimumab in patients of black race with systemic lupus erythematosus: results from the EMBRACE study. Lupus Sci Med. 2019 Apr 1;6(Suppl 1):A149.2-A150.

144. Maslen T, Bruce IN, D'cruz D, Ianosev M, Bass DL, Wilkinson C, et al. Efficacy of belimumab in two serologically distinct high disease activity subgroups of patients with

systemic lupus erythematosus: post-hoc analysis of data from the phase III programme. Lupus Sci Med. 2021 Feb 10;8(1):459.

145. Furie R, Khamashta M, Merrill JT, Werth VP, Kalunian K, Brohawn P, et al. Anifrolumab, an Anti–Interferon-α Receptor Monoclonal Antibody, in Moderate-to-Severe Systemic Lupus Erythematosus. Arthritis and Rheumatology. 2017 Feb 1;69(2):376–86.

146. Merrill JT, Furie R, Werth VP, Khamashta M, Drappa J, Wang L, et al. Anifrolumab effects on rash and arthritis: Impact of the type i interferon gene signature in the phase IIb MUSE study in patients with systemic lupus erythematosus. Lupus Sci Med. 2018 Nov 1;5(1).

147. Morand EF, Furie R, Tanaka Y, Bruce IN, Askanase AD, Richez C, et al. Trial of Anifrolumab in Active Systemic Lupus Erythematosus. New England Journal of Medicine. 2020 Jan 16;382(3):211–21.

148. Wilkinson C, Henderson RB, Jones-Leone AR, Flint SM, Lennon M, Levy RA, et al. The role of baseline BLyS levels and type 1 interferon-inducible gene signature status in determining belimumab response in systemic lupus erythematosus: A post hoc meta-analysis. Arthritis Res Ther. 2020 May 4;22(1):1–11.

149. Carter LM, Alase A, Wigston Z, Psarras A, Burska A, MD Yusof MY, et al. OP0134 Novel interferon gene expression scores predict refractory severe cutaneous disease following rituximab therapy in sle. Ann Rheum Dis. 2021 Jun 1;80(Suppl 1):78.1-78.

150. Vital EM, Wittmann M, Edward S, Md Yusof MY, MacIver H, Pease CT, et al. Brief report: Responses to rituximab suggest B cell-independent inflammation in cutaneous systemic lupus erythematosus. Arthritis and Rheumatology. 2015 Jun 1;67(6):1586–91.

151. Tarn JR, Howard-Tripp N, Lendrem DW, Mariette X, Saraux A, Devauchelle-Pensec V, et al. Symptom-based stratification of patients with primary Sjögren's syndrome: multi-dimensional characterisation of international observational cohorts and reanalyses of randomised clinical trials. Lancet Rheumatol. 2019;1(2):e85–94.

152. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. 2017 Jan 1;45(D1):D896–901.

153. Farh KKH, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015 Feb 19;518(7539):337–43.

154. Barturen G, Beretta L, Cervera R, Van Vollenhoven R, Alarcón-Riquelme ME. Moving towards a molecular taxonomy of autoimmune rheumatic diseases. Vol. 14, Nature Reviews Rheumatology. Nature Publishing Group; 2018. p. 75–93.

155. Fernando MMA, Stevens CR, Sabeti PC, Walsh EC, McWhinnie AJM, Shah A, et al. Identification of two independent risk factors for lupus within the MHC in United Kingdom families. PLoS Genet. 2007;3(11):2109–21.

156. Morris DL, Taylor KE, Fernando MMA, Nititham J, Alarcón-Riquelme ME, Barcellos LF, et al. Unraveling multiple MHC gene associations with systemic lupus erythematosus: model choice indicates a role for HLA alleles and non-HLA genes in Europeans. Am J Hum Genet. 2012 Nov 2;91(5):778–93.

157. Kim K, Bang SY, Lee HS, Okada Y, Han B, Saw WY, et al. The HLA-DRβ1 amino acid positions 11-13-26 explain the majority of SLE-MHC associations. Nat Commun. 2014;5:5902.

158. Sun C, Molineros JE, Looger LL, Zhou XJ, Kim K, Okada Y, et al. High-density genotyping of immune-related loci identifies new SLE risk variants in individuals with Asian ancestry. Nat Genet. 2016 Mar 1;48(3):323–30.

159. Mayes MD, Bossini-Castillo L, Gorlova O, Martin JE, Zhou X, Chen W V., et al. Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. Am J Hum Genet. 2014 Jan 2;94(1):47–61.

160. Arnett FC, Gourh P, Shete S, Ahn CW, Honey RE, Agarwal SK, et al. Major histocompatibility complex (MHC) class II alleles, haplotypes and epitopes which confer susceptibility or protection in systemic sclerosis: analyses in 1300 Caucasian, African-American and Hispanic cases and 1000 controls. Ann Rheum Dis. 2010;69(5):822–7.

161. Gladman DD, Kung TN, Siannis F, Pellett F, Farewell VT, Lee P. HLA markers for susceptibility and expression in scleroderma. J Rheumatol. 2005;32(8).

162. Lessard CJ, Li H, Adrianto I, Ice JA, Rasmussen A, Grundahl KM, et al. Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjögren's syndrome. Nat Genet. 2013 Nov;45(11):1284–94.

163. Cobb BL, Lessard CJ, Harley JB, Moser KL. Genes and Sjögren's syndrome. Rheum Dis Clin North Am. 2008 Nov;34(4):847–68.

164. Brkic Z, Van Bon L, Cossu M, Van Helden-Meeuwsen CG, Vonk MC, Knaapen H, et al. The interferon type I signature is present in systemic sclerosis before overt fibrosis and might contribute to its pathogenesis through high BAFF gene expression and high collagen synthesis. Ann Rheum Dis. 2016 Aug 1;75(8):1567–73.

165. Brkic Z, Maria NI, Van Helden-Meeuwsen CG, Van De Merwe JP, Van Daele PL, Dalm VA, et al. Prevalence of interferon type I signature in CD14 monocytes of

patients with Sjogren's syndrome and association with disease activity and BAFF gene expression. Ann Rheum Dis. 2013 May;72(5):728–35.

166. Carnero-Montoro E, Barturen G, Povedano E, Kerick M, Martinez-Bueno M, Ballestar E, et al. Epigenome-Wide Comparative Study Reveals Key Differences Between Mixed Connective Tissue Disease and Related Systemic Autoimmune Diseases. Front Immunol. 2019;10(AUG).

167. Chen W, Cao H, Lin J, Olsen N, Zheng SG. Biomarkers for Primary Sjögren's Syndrome. Genomics Proteomics Bioinformatics. 2015 Aug 1;13(4):219–23.

168. El-Sherbiny YM, Md Yusof MY, Psarras A, Hensor EMA, Kabba KZ, Dutton K, et al. B Cell Tetherin: A Flow Cytometric Cell-Specific Assay for Response to Type I Interferon Predicts Clinical Features and Flares in Systemic Lupus Erythematosus. Arthritis and Rheumatology. 2020 May 1;72(5):769–79.

169. Kim ST, Muñoz-Grajales C, Dunn SE, Schneider R, Johnson SR, Touma Z, et al. Interferon and interferon-induced cytokines as markers of impending clinical progression in ANA+ individuals without a systemic autoimmune rheumatic disease diagnosis. Arthritis Res Ther. 2023 Dec 1;25(1).

170. Psarras A, Emery P, Vital EM. Type I interferon–mediated autoimmune diseases: pathogenesis, diagnosis and targeted therapy. Rheumatology. 2017 Oct 1;56(10):1662–75.

171. Segal B, Thomas W, Rogers T, Leon JM, Hughes P, Patel D, et al. Prevalence, Severity and Predictors of Fatigue in Primary Sjogren's Syndrome. Arthritis Rheum. 2008 Dec 15;59(12):1780.

172. Vodencarevic A, Tascilar K, Hartmann F, Reiser M, Hueber AJ, Haschka J, et al. Advanced machine learning for predicting individual risk of flares in rheumatoid arthritis patients tapering biologic drugs. Arthritis Res Ther. 2021 Dec 1;23(1).

173. Figgett WA, Monaghan K, Ng M, Alhamdoosh M, Maraskovsky E, Wilson NJ, et al. Machine learning applied to whole-blood RNA-sequencing data uncovers distinct subsets of patients with systemic lupus erythematosus. Clin Transl Immunology. 2019 Jan 1;8(12).

174. Hoi A, Nim HT, Koelmeyer R, Sun Y, Kao A, Gunther O, et al. Algorithm for calculating high disease activity in SLE. Rheumatology. 2021 Feb 12;

175. Garantziotis P, Nikolopoulos D, Adamichou C, Fanouriakis A, Repa A, Boumpas D, et al. POS0165 Machine learning-based approaches towards stratification of disease severity and organ damage in patients with systemic lupus erythematosus (SLE). 2024 Jun;400.2-401.

176. Garantziotis P, Nikolakis D, Doumas S, Frangou E, Sentis G, Filia A, et al. Molecular Taxonomy of Systemic Lupus Erythematosus Through Data-Driven Patient Stratification: Molecular Endotypes and Cluster-Tailored Drugs. Front Immunol [Internet]. 2022 May 9 [cited 2024 Oct 23];13:860726. Available from: https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2022.860726/full

177. Robinson GA, Peng J, Dönnes P, Coelewij L, Radziszewska A, Wincup C, et al. A machine learning approach for precision diagnosis of juvenile-onset SLE. medRxiv 19007765; doi: https://doi.org/10.1101/19007765

178. Ceccarelli F, Sciandrone M, Perricone C, Galvan G, Cipriano E, Galligari A, et al. Biomarkers of erosive arthritis in systemic lupus erythematosus: Application of machine learning models. PLoS One. 2018 Dec 1;13(12):e0207926.

179. Banchereau R, Hong S, Cantarel B, Rossello-Urgell J, Wright T, Correspondence VP. Personalized Immunomonitoring Uncovers Molecular Networks that Stratify Lupus Patients. 2016;

180. Toro-Domínguez D, Martorell-Marugán J, Goldman D, Petri M, Carmona-Sáez P, Alarcón-Riquelme ME. Stratification of Systemic Lupus Erythematosus Patients Into Three Groups of Disease Activity Progression According to Longitudinal Gene Expression. Arthritis and Rheumatology. 2018 Dec 1;70(12):2025–35.

181. Guthridge JM, Lu R, Tran LTH, Arriens C, Aberle T, Kamp S, et al. Adults with systemic lupus exhibit distinct molecular phenotypes in a cross-sectional study. EClinicalMedicine. 2020 Mar 1;20.

182. Toro-Domínguez, Daniel et al. Stratification of Systemic Lupus Erythematosus Patients Into Three Groups of Disease Activity Progression According to Longitudinal Gene Expression. Arthritis Rheumatol. 2018 Dec 1;70(12):2025–35.

183. Antunes M, Scirè CA, Talarico R, Alexander T, Avcin T, Belocchi C, et al. Undifferentiated connective tissue disease: State of the art on clinical practice guidelines. Vol. 4, RMD Open. BMJ Publishing Group; 2019. p. 786.

184. Sciascia S, Roccatello D, Radin M, Parodis I, Yazdany J, Pons-Estel G, et al. Differentiating between UCTD and early-stage SLE: from definitions to clinical approach. Nat Rev Rheumatol. 2022 Jan 1;18(1):9–21.

185. Mosca M, Tani C, Neri C, Della Rossa A, Baldini C, Talarico R, et al. Analysis of the evolution to defined connective tissue diseases of patients with “early unidifferentiated connective tissue diseases (UCTD).” Reumatismo. 2011 Sep 12;60(1).

186. Kim KJ, Tagkopoulos I. Application of machine learning in rheumatic disease research. Korean J Intern Med. 2019;34(4):708–22.

187. Hügle M, Omoumi P, van Laar JM, Boedecker J, Hügle T. Applied machine learning and artificial intelligence in rheumatology. Rheumatol Adv Pract. 2020;4(1).

188. Foulquier N, Redou P, Pers JO, Saraux A. New criteria and new methodological tools for devising criteria sets of inflammatory rheumatic diseases. Vol. 38, Clinical and

Experimental Rheumatology. Clinical and Experimental Rheumatology S.A.S.; 2020. p. 776–82.

189. Robinson GA, Peng J, Dönnes P, Coelewij L, Naja M, Radziszewska A, et al. Disease-associated and patient-specific immune cell signatures in juvenile-onset systemic lupus erythematosus: patient stratification using a machine-learning approach. Lancet Rheumatol. 2020 Aug 1;2(8):e485–96.

190. L MG, J P, NL T, GA R, M N, H P, et al. Two shared immune cell signatures stratify patients with Sjögren's syndrome and systemic lupus erythematosus with potential therapeutic implications. Arthritis Rheumatol. 2021 Mar;

191. Foulquier N, Le Dantec C, Bettacchioli E, Jamin C, Beretta L, Vigone B, et al. Machine learning identifies a common signature for anti-SSA/Ro60 antibody expression across autoimmune diseases. Arthritis & Rheumatology. 2022 May 30;

192. Le Bon A, Thompson C, Kamphuis E, Durand V, Rossmann C, Kalinke U, et al. Cutting edge: enhancement of antibody responses through direct stimulation of B and T cells by type I IFN. J Immunol [Internet]. 2006 Feb 15 [cited 2024 Sep 16];176(4):2074–8. Available from: https://pubmed.ncbi.nlm.nih.gov/16455962/

193. Longhi MP, Trumpfheller C, Idoyaga J, Caskey M, Matos I, Kluger C, et al. Dendritic cells require a systemic type I interferon response to mature and induce CD4+ Th1 immunity with poly IC as adjuvant. J Exp Med [Internet]. 2009 Jul 6 [cited 2024 Sep 16];206(7):1589–602. Available from: https://pubmed.ncbi.nlm.nih.gov/19564349/

194. Jego G, Palucka AK, Blanck JP, Chalouni C, Pascual V, Banchereau J. Plasmacytoid dendritic cells induce plasma cell differentiation through type I interferon and interleukin 6. Immunity [Internet]. 2003 Aug 1 [cited 2024 Sep 16];19(2):225–34. Available from: https://pubmed.ncbi.nlm.nih.gov/12932356/

195. Psarras A. Type I Interferon Regulation in Health and Autoimmune Disease [Thesis]. [Leeds]: University of Leeds; 2018.

196. Morand EF, Furie R, Tanaka Y, Bruce IN, Askanase AD, Richez C, et al. Trial of Anifrolumab in Active Systemic Lupus Erythematosus. N Engl J Med [Internet]. 2020 Jan 16 [cited 2024 Sep 16];382(3):211–21. Available from: https://pubmed.ncbi.nlm.nih.gov/31851795/

197. Khanna D, Denton CP, Assassi S, Kuwana M, Allanore Y, Domsic RT, et al. A randomised, parallel-group, double-blind, placebo-controlled phase 3 study to Determine the effectiveness of the type I interferon receptor antibody, Anifrolumab, In SYstemic sclerosis: DAISY study design and rationale. Clin Exp Rheumatol [Internet]. 2024 Aug 1 [cited 2024 Sep 16];42(8):1635–44. Available from: https://pubmed.ncbi.nlm.nih.gov/39152751/

198. Study Details | Anifrolumab Treatment for 24 Weeks in Patients With Primary Sjögren's Syndrome | ClinicalTrials.gov [Internet]. [cited 2024 Sep 16]. Available from: https://clinicaltrials.gov/study/NCT05383677

199. Carter LM, Wigston Z, Laws P, Vital EM. Rapid efficacy of anifrolumab across multiple subtypes of recalcitrant cutaneous lupus erythematosus parallels changes in discrete subsets of blood transcriptomic and cellular biomarkers. Br J Dermatol [Internet]. 2023 Aug 1 [cited 2024 Sep 16];189(2):210–8. Available from: https://pubmed.ncbi.nlm.nih.gov/36944572/

200. Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. Arthritis Rheumatol [Internet]. 2014 [cited 2023 Mar 6];66(6):1583–95. Available from: https://pubmed.ncbi.nlm.nih.gov/24644022/

201. Banchereau R, Hong S, Cantarel B, Rossello-Urgell J, Wright T, Correspondence VP. Personalized Immunomonitoring Uncovers Molecular Networks that Stratify Lupus Patients. 2016 [cited 2021 Jul 12]; Available from: http://dx.doi.org/10.1016/j.cell.2016.03.008

202. Jourde-Chiche N mie, Whalen E, Gondouin B, Speake C, Gersuk V, Dussol B, et al. Modular transcriptional repertoire analyses identify a blood neutrophil signature as a candidate biomarker for lupus nephritis. Rheumatology (Oxford) [Internet]. 2017 Mar 1 [cited 2024 Sep 27];56(3):477–87. Available from: https://pubmed.ncbi.nlm.nih.gov/28031441/

203. Carter LM, Alase A, Wigston Z, Psarras A, Burska A, Sutton E, et al. Gene expression and autoantibody analysis reveals distinct ancestry-specific profiles associated with response to rituximab in refractory systemic lupus erythematosus. Arthritis Rheumatol [Internet]. 2022 Nov 21 [cited 2023 Mar 6]; Available from: https://pubmed.ncbi.nlm.nih.gov/36409591/

204. El-Sherbiny YM, Psarras A, Yusof MYM, Hensor EMA, Tooze R, Doody G, et al. A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. Sci Rep [Internet]. 2018 Dec 1 [cited 2023 Mar 6];8(1). Available from: https://pubmed.ncbi.nlm.nih.gov/29643425/

205. Catalina MD, Bachali P, Yeo AE, Geraci NS, Petri MA, Grammer AC, et al. Patient ancestry significantly contributes to molecular heterogeneity of systemic lupus erythematosus. JCI Insight [Internet]. 2020 Aug 6 [cited 2024 Sep 30];5(15). Available from: https://pubmed.ncbi.nlm.nih.gov/32759501/

206. Silk ME, Wallace DJ, Furie RA, Tanaka Y, Kalunian KC, Mosca M, et al. 200 Efficacy and safety of belimumab in patients of black race with systemic lupus erythematosus: results from the EMBRACE study. Lupus Sci Med [Internet]. 2019 Apr 1 [cited 2024

Sep 30];6(Suppl 1):A149.2-A150. Available from: https://lupus.bmj.com/content/6/Suppl_1/A149.2

207. MacQueen, James and others. Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967;1(14):281–97.

208. Patel E, Kushwaha DS. Clustering Cloud Workloads: K-Means vs Gaussian Mixture Model. Procedia Comput Sci. 2020;171(2019):158–67.

209. Kubara K. Towards data science. 2020 [cited 2023 Sep 9]. Gaussian Mixture Models vs K-Means. Which One to Choose? Available from: https://towardsdatascience.com/gaussian-mixture-models-vs-k-means-which-one-to-choose-62f2736025f0

210. Gandhi V. Kaggle. 2019 [cited 2023 Sep 8]. Gaussian Mixture Models Clustering - Explained. Available from: https://www.kaggle.com/code/vipulgandhi/gaussian-mixture-models-clustering-explained

211. Angelini F, Widera P, Mobasheri A, Blair J, Struglics A, Uebelhoer M, et al. Osteoarthritis endotype discovery via clustering of biochemical marker data. Ann Rheum Dis. 2022;81(5):666–75.

212. Pettersson S, Demmelmaier I, Nordgren B, Dufour AB, Opava CH. Identification and Prediction of Fatigue Trajectories in People With Rheumatoid Arthritis. ACR Open Rheumatol. 2022 Feb 1;4(2):111–8.

213. Su R, Li B, Wu R, Xie Y, Gao A, Gao C, et al. Stratified distribution of Th17 and Treg cells in patients with multi-stage rheumatoid arthritis. Arthritis Res Ther. 2023 Dec 1;25(1).

214. Werdyani S, Liu M, Zhang H, Sun G, Furey A, Randell EW, et al. Endotypes of primary osteoarthritis identified by plasma metabolomics analysis. Rheumatology (Oxford). 2021 Jun 1;60(6):2735–44.

215. Choi MY, Chen I, Clarke AE, Fritzler MJ, Buhler KA, Urowitz M, et al. Machine learning identifies clusters of longitudinal autoantibody profiles predictive of systemic lupus erythematosus disease outcomes. Ann Rheum Dis. 2023;82(7):927–36.

216. Gisslander K, White A, Aslett L, Hrušková Z, Lamprecht P, Musiał J, et al. Data-driven subclassification of ANCA-associated vasculitis: model-based clustering of a federated international cohort. Lancet Rheumatol [Internet]. 2024 Nov 1 [cited 2025 Jan 21];6(11). Available from: https://pubmed.ncbi.nlm.nih.gov/39182506/

217. Rezaei E, Hogan D, Trost B, Kusalik AJ, Boire G, Cabral DA, et al. Associations of clinical and inflammatory biomarker clusters with juvenile idiopathic arthritis categories. Rheumatology (Oxford) [Internet]. 2020 May 1 [cited 2025 Jan 21];59(5):1066–75. Available from: https://pubmed.ncbi.nlm.nih.gov/32321162/

218. Yiu C, Griffith JF, Xiao F, Shi L, Zhou B, Wu S, et al. Automated quantification of wrist bone marrow oedema, pre- and post-treatment, in early rheumatoid arthritis. Rheumatol Adv Pract [Internet]. 2024 [cited 2025 Jan 21];8(3). Available from: https://pubmed.ncbi.nlm.nih.gov/38915843/

219. Xie J, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. 33rd International Conference on Machine Learning, ICML 2016. 2016;1:740–9.

220. Von Luxburg U. A tutorial on spectral clustering. Stat Comput. 2007;17(4):395–416.

221. Keerthana V. What, why and how of Spectral Clustering! Data Science Blogathon. 2021;

222. Yang Y, Xu D, Nie F, Yan S, Zhuang Y. Image clustering using local discriminant models and global integration. IEEE Transactions on Image Processing. 2010;19(10):2761–73.

223. Tian F, Gao B, Cui Q, Chen E, Liu TY. Learning deep representations for graph clustering. Proceedings of the National Conference on Artificial Intelligence. 2014;2:1293–9.

224. Chollet F, Allaire JJ. Artificial intelligence, machine learning and deep learning. In: Deep Learning with R. 1st ed. New York: Manning Publications; 2018. p. 1–109.

225. Marino J. Predictive coding, variational autoencoders, and biological connections. Neural Comput. 2021;34(1):1–44.

226. Werbos P. Beyond regression: new tools for prediction and analysis in the behavioral sciences. PhD thesis, Committee on Applied Mathematics, Harvard University, Cambridge, MA. 1974;

227. Rumelhart DE, Hintont GE. Learning Representations by Back-Propagating Errors. Cognitive Modeling. 2019;(2):3–6.

228. Chollet F, Allaire JJ. The engine of neural networks: gradient-based optimisation. In: 1, editor. Deep Learning with R2. 1st ed. New York: Manning Publications; 2018. p. 41–7.

229. Patrikar S. Towards data science. 2019 [cited 2023 Sep 8]. p. 1 Batch, Mini Batch & Stochastic Gradient Descent. Available from: https://towardsdatascience.com/batch-mini-batch-stochastic-gradient-descent-7a62ecba642a

230. Kingma DP, Welling M. Auto-encoding variational bayes. 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings. 2014;

231. Rezende DJ, Mohamed S, Wierstra D. Stochastic backpropagation and approximate inference in deep generative models. 31st International Conference on Machine Learning, ICML 2014. 2014;4:3057–70.

232. Chollet F, Allaire JJ. Generative deep learning. In: Deep Learning with R. 1st ed. New York: Manning Publications; 2018. p. 250–92.

233. Singh A, Ogunfunmi T. An Overview of Variational Autoencoders for Source Separation, Finance, and Bio-Signal Applications. Entropy. 2022;24(1).

234. Singh A, Ogunfunmi T. An Overview of Variational Autoencoders for Source Separation, Finance, and Bio-Signal Applications. Entropy. 2022 Jan 1;24(1).

235. Rocca J. Towards data science. 2019 [cited 2023 Sep 8]. Understanding Variational Autoencoders (VAEs). Available from: https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73

236. Colby SM, Nuñez JR, Hodas NO, Corley CD, Renslow RR. Deep Learning to Generate in Silico Chemical Property Libraries and Candidate Molecules for Small Molecule Identification in Complex Samples. Anal Chem. 2020;92(2):1720–9.

237. Grønbech CH, Vording MF, Timshel PN, Sønderby CK, Pers TH, Winther O. ScVAE: Variational auto-encoders for single-cell gene expression data. Bioinformatics. 2020;36(16):4415–22.

238. Pratella D, Saadi SAEM, Bannwarth S, Paquis-fluckinger V, Bottini S. A survey of autoencoder algorithms to pave the diagnosis of rare diseases. Int J Mol Sci. 2021;22(19).

239. Franco EF, Rana P, Cruz A, Calderón V V., Azevedo V, Ramos RTJ, et al. Performance comparison of deep learning autoencoders for cancer subtype detection using multi-omics data. Cancers (Basel). 2021;13(9):1–17.

240. Zhang L, Lv C, Jin Y, Cheng G, Fu Y, Yuan D, et al. Deep learning-based multi-omics data integration reveals two prognostic subtypes in high-risk neuroblastoma. Front Genet. 2018;9(OCT):1–9.

241. Rampášek L, Hidru D, Smirnov P, Haibe-Kains B, Goldenberg A. Dr.VAE: Improving drug response prediction via modeling of drug perturbation effects. Bioinformatics. 2019;35(19):3743–51.

242. Lim KL, Jiang X, Yi C. Deep Clustering with Variational Autoencoder. IEEE Signal Process Lett. 2020;27:231–5.

243. Xu J, Xu J, Meng Y, Lu C, Cai L, Zeng X, et al. Graph embedding and Gaussian mixture variational autoencoder network for end-to-end analysis of single-cell RNA sequencing data. Cell reports methods. 2023 Jan 23;3(1).

244. Chi C, Solomon O, Shiboski C, Taylor KE, Quach H, Quach D, et al. Identification of Sjögren's syndrome patient subgroups by clustering of labial salivary gland DNA methylation profiles. PLoS One. 2023 Mar 1;18(3).

245. Chollet F, Allaire JJ. Deep learning for text and sequences. In: Deep Learning with R. 1st ed. Manning Publications; 2018. p. 164–217.

246. Mikolov T, Chen K, Corrado G, Dean J. Efficient Estimation of Word Representations in Vector Space. 1st International Conference on Learning Representations, ICLR 2013 - Workshop Track Proceedings. 2013 Jan 16;

247. Pennington J, Socher R, Manning CD. GloVe: Global Vectors for Word Representation. In: Empirical Methods in Natural Language Processing (EMNLP). 2014. p. 1532–43.

248. Humbert-Droz M, Izadi Z, Schmajuk G, Gianfrancesco M, Baker MC, Yazdany J, et al. Development of a Natural Language Processing System for Extracting Rheumatoid

Arthritis Outcomes From Clinical Notes Using the National Rheumatology Informatics System for Effectiveness Registry. Arthritis Care Res (Hoboken). 2023;75(3):608–15.

249. Matsuda S, Ohtomo T, Tomizawa S, Miyano Y, Mogi M, Kuriki H, et al. Incorporating unstructured patient narratives and health insurance claims data in pharmacovigilance: natural language processing analysis of patient-generated texts about systemic lupus erythematosus. JMIR Public Health Surveill. 2021;7(6):1–14.

250. Zhao SS, Hong C, Cai T, Xu C, Huang J, Ermann J, et al. Incorporating natural language processing to improve classification of axial spondyloarthritis using electronic health records. Rheumatology (United Kingdom). 2020;59(5):1059–65.

251. Chollet F, Allaire JJ. Conclusions. In: Deep Learning with R. 1st ed. Manning Publications; 2018.

252. Şenbabaoğlu Y, Michailidis G, Li JZ. Critical limitations of consensus clustering in class discovery. Sci Rep [Internet]. 2014 Aug 27 [cited 2024 Sep 17];4. Available from: https://pubmed.ncbi.nlm.nih.gov/25158761/

253. John CR, Watson D, Russ D, Goldmann K, Ehrenstein M, Pitzalis C, et al. M3C: Monte Carlo reference-based consensus clustering. Scientific Reports 2020 10:1 [Internet]. 2020 Feb 4 [cited 2023 Nov 7];10(1):1–14. Available from: https://www.nature.com/articles/s41598-020-58766-1

254. Kapur N. The NHS Long Term Plan. SUSHRUTA Journal of Health Policy & Opinions. 2020;12(1):10–1.

255. Park JJH, Hsu G, Siden EG, Thorlund K, Mills EJ. An overview of precision oncology basket and umbrella trials for clinicians. CA Cancer J Clin. 2020;70(2):125–37.

256. Berry DA. The Brave New World of clinical cancer research: Adaptive biomarker-driven trials integrating clinical practice with clinical research. Mol Oncol. 2015;9(5):951–9.

257. Park JJH, Siden E, Zoratti MJ, Dron L, Harari O, Singer J, et al. Systematic review of basket trials, umbrella trials, and platform trials: A landscape analysis of master protocols. Trials. 2019;20(1):1–10.

258. O'Dwyer PJ, Gray RJ, Flaherty KT, Chen AP, Li S, Wang V, et al. The NCI-MATCH trial: lessons for precision oncology. Nat Med. 2023;29(6):1349–57.

259. Moore KN, Mannel RS. Is the NCI MATCH trial a match for gynecologic oncology? Gynecol Oncol. 2016;140(1):161–6.

260. Chen AP, Kummar S, Moore N, Rubinstein L V., Zhao Y, Williams PM, et al. Molecular Profiling-Based Assignment of Cancer Therapy (NCI-MPACT): A Randomized Multicenter Phase II Trial. JCO Precis Oncol. 2021;(5):133–44.

261. Toro-Domínguez D, Martorell-Marugán J, Martinez-Bueno M, López-Domínguez R, Carnero-Montoro E, Barturen G, et al. Scoring personalized molecular portraits identify Systemic Lupus Erythematosus subtypes and predict individualized drug responses, symptomatology and disease progression. Brief Bioinform [Internet]. 2022 Aug 10 [cited 2024 Sep 17];23(5):1–13. Available from: https://digibug.ugr.es/handle/10481/79808

262. Barturen G, Babaei S, Català-Moll F, Martínez-Bueno M, Makowska Z, Martorell-Marugán J, et al. Integrative Analysis Reveals a Molecular Stratification of Systemic Autoimmune Diseases. Arthritis Rheumatol [Internet]. 2021 Jun 1 [cited 2022 Jun 22];73(6):1073–85. Available from: https://pubmed.ncbi.nlm.nih.gov/33497037/

263. Lorenzon R, Ribet C, Pitoiset F, Aractingi S, Banneville B, Beaugerie L, et al. The universal effects of low-dose interleukin-2 across 13 autoimmune diseases in a basket clinical trial. J Autoimmun. 2024 Apr 1;144:103172.

264. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. New England Journal of Medicine. 2015;372(26):2509–20.

265. Marcus L, Lemery SJ, Keegan P, Pazdur R. FDA approval summary: Pembrolizumab for the treatment of microsatellite instability-high solid tumors. Clinical Cancer Research. 2019;25(13):3753–8.

266. Chen Y, Chi P. Basket trial of TRK inhibitors demonstrates efficacy in TRK fusion-positive cancers. J Hematol Oncol. 2018;11(1):1–5.

267. Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, Demetri GD, et al. Efficacy of Larotrectinib in TRK Fusion–Positive Cancers in Adults and Children . New England Journal of Medicine. 2018;378(8):731–9.

268. Hong DS, DuBois SG, Kummar S, Farago AF, Albert CM, Rohrberg KS, et al. Larotrectinib in patients with TRK fusion-positive solid tumours: a pooled analysis of three phase 1/2 clinical trials. Lancet Oncol. 2020;21(4):531–40.

269. Demetri GD, De Braud F, Drilon A, Siena S, Patel MR, Cho BC, et al. Updated Integrated Analysis of the Efficacy and Safety of Entrectinib in Patients With NTRK Fusion-Positive Solid Tumors. Clinical Cancer Research. 2022;28(7):1302–12.

270. Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay JY, et al. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations . New England Journal of Medicine. 2015;373(8):726–36.

271. Flaherty KR, Wells AU, Cottin V, Devaraj A, Walsh SLF, Inoue Y, et al. Nintedanib in Progressive Fibrosing Interstitial Lung Diseases. N Engl J Med. 2019 Oct 31;381(18):1718–27.

272. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus

erythematosus: The randomized, double-blind, phase ii/iii systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum. 2010 Jan;62(1):222–33.

273. Brown S, Navarro Coy N, Pitzalis C, Emery P, Pavitt S, Gray J, et al. The TRACTISS Protocol: A randomised double blind placebo controlled clinical TRial of Anti-B-Cell Therapy in patients with primary Sjögren's Syndrome. BMC Musculoskelet Disord. 2014 Jan 17;15(1).

274. Oddis C V., Reed AM, Aggarwal R, Rider LG, Ascherman DP, Levesque MC, et al. Rituximab in the treatment of refractory adult and juvenile dermatomyositis and adult polymyositis: A randomized, placebo-phase trial. Arthritis Rheum. 2013 Feb;65(2):314–24.

275. Laigle L, et al. AB1372 Towards reforming the taxonomy of human disease: the precisesads cross sectional study. Ann Rheum Dis [Internet]. 2018 Jun 1 [cited 2024 Sep 17];77(Suppl 2):1771–2. Available from: https://ard.bmj.com/content/77/Suppl_2/1771.3

276. Khatri B, Tessneer KL, Rasmussen A, Aghakhanian F, Reksten TR, Adler A, et al. Genome-wide association study identifies Sjögren's risk loci with functional implications in immune and glandular cells. Nat Commun [Internet]. 2022 Dec 1 [cited 2024 Sep 17];13(1). Available from: https://pubmed.ncbi.nlm.nih.gov/35896530/

277. Foulquier N, Le Dantec C, Bettacchioli E, Jamin C, Alarcón-Riquelme ME, Pers JO. Machine Learning for the Identification of a Common Signature for Anti–SSA/Ro 60 Antibody Expression Across Autoimmune Diseases. Arthritis and Rheumatology [Internet]. 2022 Oct 1 [cited 2024 Sep 17];74(10):1706–19. Available from: https://pubmed.ncbi.nlm.nih.gov/35635731/

278. Parodis I, Lindblom J, Toro-Domínguez D, Beretta L, Borghi MO, Castillo J, et al. Interferon and B-cell Signatures Inform Precision Medicine in Lupus Nephritis. Kidney

Int Rep [Internet]. 2024 Jun 1 [cited 2024 Sep 17];9(6):1817–35. Available from: https://pubmed.ncbi.nlm.nih.gov/38899167/

279. Beretta L, Barturen G, Vigone B, Bellocchi C, Hunzelmann N, De Langhe E, et al. Genome-wide whole blood transcriptome profiling in a large European cohort of systemic sclerosis patients. Ann Rheum Dis. 2020;79(9):1218–26.

280. Furie RA, Petri MA, Wallace DJ, Ginzler EM, Merrill JT, Stohl W, et al. Novel Evidence-Based Systemic Lupus Erythematosus Responder Index. Arthritis Rheum [Internet]. 2009 Sep 9 [cited 2024 Sep 23];61(9):1143. Available from: https://pubmed.ncbi.nlm.nih.gov/19714615/

281. Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol. 2002;29(2).

282. Yee CS, Isenberg DA, Prabu A, Sokoll K, Teh LS, Rahman A, et al. BILAG-2004 index captures systemic lupus erythematosus disease activity better than SLEDAI-2000. Ann Rheum Dis [Internet]. 2008 Jun [cited 2024 Sep 17];67(6):873–6. Available from: https://pubmed.ncbi.nlm.nih.gov/17519277/

283. Sultan SM, Allen E, Oddis C V., Kiely P, Cooper RG, Lundberg IE, et al. Reliability and validity of the myositis disease activity assessment tool. Arthritis Rheum. 2008 Nov;58(11):3593–9.

284. Rider LG, Werth VP, Huber AM, Alexanderson H, Rao AP, Ruperto N, et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis. Arthritis Care Res (Hoboken). 2011 Nov;63(SUPPL. 11).

285. Isenberg DA, Allen E, Farewell V, Ehrenstein MR, Hanna MG, Lundberg IE, et al. International consensus outcome measures for patients with idiopathic inflammatory myopathies. Development and initial validation of myositis activity and damage indices in patients with adult onset disease. Rheumatology (Oxford) [Internet]. 2004

Jan [cited 2024 Sep 23];43(1):49–54. Available from: https://pubmed.ncbi.nlm.nih.gov/12867580/

286. MYOJAK - Health Research Authority [Internet]. [cited 2024 Sep 23]. Available from: https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/myojak/

287. Seror R, Bowman SJ, Brito-Zeron P, Theander E, Bootsma H, Tzioufas A, et al. EULAR Sjögren's syndrome disease activity index (ESSDAI): a user guide. RMD Open [Internet]. 2015 Feb 1 [cited 2024 Sep 23];1(1):e000022. Available from: https://rmdopen.bmj.com/content/1/1/e000022

288. Machado AC, dos Santos LC, Fidelix T, Lekwitch I, Soares SB, Gasparini AF, et al. Effectiveness and safety of abatacept for the treatment of patients with primary Sjögren's syndrome. Clin Rheumatol [Internet]. 2020 Jan 1 [cited 2024 Sep 23];39(1):243–8. Available from: https://pubmed.ncbi.nlm.nih.gov/31420813/

289. Gottenberg JE, Cinquetti G, Larroche C, Combe B, Hachulla E, Meyer O, et al. Efficacy of rituximab in systemic manifestations of primary Sjögren's syndrome: results in 78 patients of the AutoImmune and Rituximab registry. Ann Rheum Dis [Internet]. 2013 Jun 1 [cited 2024 Sep 23];72(6):1026–31. Available from: https://ard.bmj.com/content/72/6/1026

290. A Randomized, Double-blind 2-arm NEPTUNUS Extension Study to Assess the Long-term Safety and Efficacy of Ianalumab in Patients With Sjogrens Syndrome. | Novartis [Internet]. [cited 2024 Sep 23]. Available from: https://www.novartis.com/clinicaltrials/study/nct05985915

291. Chessa E, Piga M, Floris A, Devilliers H, Cauli A, Arnaud L. Use of Physician Global Assessment in systemic lupus erythematosus: a systematic review of its psychometric properties. Rheumatology (United Kingdom). 2020;59(12 Rheumatology (United Kingdom)):3622–32.

292. Rose E, Ferrada MA, Quinn KA, Goodspeed W, Arnaud L, Sharma A, et al. Physician Global Assessment as a Disease Activity Measure for Relapsing Polychondritis. Arthritis Care Res (Hoboken). 2022;74(8):1269–76.

293. Franklyn K, Lau CS, Navarra S V., Louthrenoo W, Lateef A, Hamijoyo L, et al. Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). Ann Rheum Dis. 2016 Sep 1;75(9):1615–21.

294. A L, MH S, J H, J U, T S, WB G, et al. Assessing health-related quality of life in primary Sjögren's syndrome-The PSS-QoL. Semin Arthritis Rheum. 2018 Aug 1;48(1):105–10.

295. 36-Item Short Form Survey (SF-36) | RAND [Internet]. [cited 2024 Sep 5]. Available from: https://www.rand.org/health-care/surveys_tools/mos/36-item-short-form.html

296. Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: Scoping review. SAGE Open Med [Internet]. 2016 [cited 2024 Sep 23];4. Available from: https://pubmed.ncbi.nlm.nih.gov/27757230/

297. sf36: SF-36 questionnaire scoring in lbraglia/QScoring: Miscellaneous scoring functions [Internet]. [cited 2024 Sep 23]. Available from: https://rdrr.io/github/lbraglia/QScoring/man/sf36.html

298. Zhang Q, Wang X, Chen H, Shen B. Sjögren's syndrome is associated with negatively variable impacts on domains of health-related quality of life: evidence from Short Form 36 questionnaire and a meta-analysis. Patient Prefer Adherence [Internet]. 2017 May 10 [cited 2024 Sep 23];11:905–11. Available from: https://pubmed.ncbi.nlm.nih.gov/28546741/

299. Fischer A, Zimovetz E, Ling C, Esser D, Schoof N. Humanistic and cost burden of systemic sclerosis: A review of the literature. Autoimmun Rev [Internet]. 2017 Nov 1

[cited 2024 Sep 23];16(11):1147–54. Available from: https://pubmed.ncbi.nlm.nih.gov/28899803/

300. Frade S, O'Neill S, Greene D, Nutter E, Cameron M. Exercise as adjunctive therapy for systemic lupus erythematosus. Cochrane Database Syst Rev [Internet]. 2023 Apr 19 [cited 2024 Sep 23];4(4). Available from: https://pubmed.ncbi.nlm.nih.gov/37073886/

301. Patrick DL and EP. Health Status and Health Policy: Quality of Life in Health Care Evaluation and Resource Allocation. eweb:124761.

302. EQ-5D-5L | EuroQol [Internet]. [cited 2024 Sep 5]. Available from: https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-5l/

303. Devlin ProfDrN, Parkin D, Janssen B. An Introduction to EQ-5D Instruments and Their Applications. Methods for Analysing and Reporting EQ-5D Data [Internet]. 2020 Jul 21 [cited 2024 Sep 5];1–22. Available from: https://www.ncbi.nlm.nih.gov/books/NBK565680/

304. FACIT-Fatigue [Internet]. [cited 2024 Sep 23]. Available from: https://www.facit.org/measures/facit-fatigue

305. Harel D, Group on behalf of CSR, Thombs BD, Group on behalf of CSR, Hudson M, Group on behalf of CSR, et al. Measuring fatigue in SSc: a comparison of the Short Form-36 Vitality subscale and Functional Assessment of Chronic Illness Therapy–Fatigue scale. Rheumatology [Internet]. 2012 Dec 1 [cited 2024 Sep 5];51(12):2177–85. Available from: https://dx.doi.org/10.1093/rheumatology/kes206

306. Flynn TN, Huynh E, Peters TJ, Al-Janabi H, Clemens S, Moody A, et al. Scoring the Icecap-a capability instrument. Estimation of a UK general population tariff. Health Econ [Internet]. 2015 Mar 1 [cited 2024 Sep 5];24(3):258–69. Available from: https://pubmed.ncbi.nlm.nih.gov/24254584/

307. ICECAP-A | Bristol Medical School: Population Health Sciences | University of Bristol [Internet]. [cited 2024 Sep 5]. Available from: https://www.bristol.ac.uk/population-health-sciences/projects/icecap/icecap-a/

308. Farkas M, Huynh E, Gulácsi L, Zrubka Z, Dobos Á, Kovács L, et al. Development of Population Tariffs for the ICECAP-A Instrument for Hungary and their Comparison With the UK Tariffs. Value in Health. 2021 Dec 1;24(12):1845–52.

309. Rohrbach PJ, Dingemans AE, Groothuis-Oudshoorn CGM, Van Til JA, Essers BA, Van Furth EF, et al. The ICEpop Capability Measure for Adults Instrument for Capabilities: Development of a Tariff for the Dutch General Population. Value in Health. 2022 Jan 1;25(1):125–32.

310. Péntek M, Poór G, Gulácsi L, Zrubka Z, Brodszky V, Rencz F, et al. Musculoskeletal health and capability wellbeing: Associations between the HAQ-DI, ICECAP-A and ICECAP-O measures in a population survey. Musculoskelet Sci Pract [Internet]. 2021 Oct 1 [cited 2024 Sep 23];55. Available from: https://pubmed.ncbi.nlm.nih.gov/34271413/

311. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics [Internet]. 1993 [cited 2024 Sep 5];4(5):353–65. Available from: https://pubmed.ncbi.nlm.nih.gov/10146874/

312. Kronborg C, Gitte AE, Ae H, Axelsen F. Health care costs, work productivity and activity impairment in non-malignant chronic pain patients.

313. Glanz BI, Dégano IR, Rintell DJ, Chitnis T, Weiner HL, Healy BC. Work productivity in relapsing multiple sclerosis: associations with disability, depression, fatigue, anxiety, cognition, and health-related quality of life. Value Health [Internet]. 2012 Dec [cited 2024 Sep 5];15(8):1029–35. Available from: https://pubmed.ncbi.nlm.nih.gov/23244804/

314. Ariza-Ariza R, Hernández-Cruz B, Navarro-Compán V, Leyva Pardo C, Juanola X, Navarro-Sarabia F. A comparison of telephone and paper self-completed questionnaires of main patient-related outcome measures in patients with ankylosing spondylitis and psoriatic arthritis. Rheumatol Int [Internet]. 2013 Nov [cited 2024 Sep 5];33(11):2731–6. Available from: https://pubmed.ncbi.nlm.nih.gov/23793389/

315. Haglund E, Petersson IF, Bremander A, Bergman S. Predictors of presenteeism and activity impairment outside work in patients with spondyloarthritis. J Occup Rehabil [Internet]. 2015 Jun 22 [cited 2024 Sep 5];25(2):288–95. Available from: https://pubmed.ncbi.nlm.nih.gov/25173795/

316. Heller GZ, Manuguerra M, Chow R. How to analyze the Visual Analogue Scale: Myths, truths and clinical relevance. Scand J Pain [Internet]. 2016 Oct 1 [cited 2024 Sep 23];13(1):67–75. Available from: https://www.degruyter.com/document/doi/10.1016/j.sjpain.2016.06.012/html

317. Freyd M. The Graphic Rating Scale. J Educ Psychol. 1923 Feb;14(2):83–102.

318. van Riel PLCM, Renskers L. The Disease Activity Score (DAS) and the Disease Activity Score using 28 joint counts (DAS28) in the management of rheumatoid arthritis. Clin Exp Rheumatol [Internet]. 2016 [cited 2024 Sep 23];34:40–4. Available from: https://www.clinexprheumatol.org/abstract.asp?a=11135

319. Benucci M, Gobbi FL, Fusi P, Damiani A, Russo E, Guiducci S, et al. Different Biomarkers of Response to Treatment with Selective Jak-1 Inhibitors in Rheumatoid Arthritis. Front Biosci (Landmark Ed) [Internet]. 2023 Aug 1 [cited 2024 Sep 23];28(8). Available from: https://pubmed.ncbi.nlm.nih.gov/37664943/

320. Miyauchi K, Fujimoto K, Abe T, Takei M, Ogawa K. Cross-sectional assessment of sleep and fatigue in middle-aged Japanese women with primary Sjogren syndrome or rheumatoid arthritis using self-reports and wrist actigraphy. Medicine [Internet]. 2021

Sep 17 [cited 2024 Sep 23];100(37). Available from: https://pubmed.ncbi.nlm.nih.gov/34664865/

321. Mahmoud K, Zayat AS, Yusof MYM, Dutton K, Teh LS, Yee CS, et al. Ultrasound to identify systemic lupus erythematosus patients with musculoskeletal symptoms who respond best to therapy: the US Evaluation For mUsculoskeletal Lupus longitudinal multicentre study. Rheumatology (Oxford) [Internet]. 2021 Nov 1 [cited 2024 Sep 23];60(11):5194–204. Available from: https://pubmed.ncbi.nlm.nih.gov/33792659/

322. Barbacki A, Petri M, Avina-Zubieta A, Alarcon GS, Bernatsky S. Fatigue Measurements in Systemic Lupus Erythematosus. J Rheumatol [Internet]. 2019 Nov 1 [cited 2024 Sep 23];46(11):1470–7. Available from: https://pubmed.ncbi.nlm.nih.gov/30709953 /

323. Waldheim E, Elkan AC, Bergman S, Frostegård J, Van Vollenhoven R, Henriksson EW. Extent and characteristics of self-reported pain in patients with systemic lupus erythematosus. Lupus [Internet]. 2013 Feb [cited 2024 Sep 23];22(2):136–43. Available from: https://pubmed.ncbi.nlm.nih.gov/23192324/

324. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. Pain [Internet]. 1986 [cited 2024 Sep 23];27(1):117–26. Available from: https://journals.lww.com/pain/fulltext/1986/10000/the_measurement_of_clinical_pain_intensity__a.10.aspx

325. Dexter F, Chestnut DH. Analysis of Statistical Tests to Compare Visual Analog Scale Measurements among Groups. Anesthesiology [Internet]. 1995 Apr 1 [cited 2024 Sep 23];82(4):896-902. Available from: https://dx.doi.org/10.1097/00000542-199504000-00012

326. voom: Transform RNA-Seq Data Ready for Linear Modelling in limma: Linear Models for Microarray Data [Internet]. [cited 2024 Sep 27]. Available from: https://rdrr.io/bioc/limma/man/voom.html

327. El-Sherbiny YM, Md Yusof MY, Psarras A, Hensor EMA, Kabba KZ, Dutton K, et al. B Cell Tetherin: A Flow Cytometric Cell-Specific Assay for Response to Type I Interferon Predicts Clinical Features and Flares in Systemic Lupus Erythematosus. Arthritis and Rheumatology [Internet]. 2020 May 1 [cited 2023 Jul 11];72(5):769–79. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.41187

328. Yusof MYM, Psarras A, El-Sherbiny YM, Hensor EMA, Dutton K, Ul-Hassan S, et al. Prediction of autoimmune connective tissue disease in an at-risk cohort: prognostic value of a novel two-score system for interferon status. Ann Rheum Dis [Internet]. 2018 Jul 20 [cited 2024 Sep 5];77(10):1432–9. Available from: https://pubmed.ncbi.nlm.nih.gov/29929956/

329. Principal Component Analysis (PCA) Explained | Built In [Internet]. [cited 2024 Sep 24]. Available from: https://builtin.com/data-science/step-step-explanation-principal-component-analysis

330. Jollife IT, Cadima J. Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences [Internet]. 2016 Apr 13 [cited 2024 Sep 24];374(2065). Available from: http://dx.doi.org/10.1098/rsta.2015.0202

331. Shlens J. A TUTORIAL ON PRINCIPAL COMPONENT ANALYSIS Derivation, Discussion and Singular Value Decomposition. 2003;

332. Chollet F, Allaire JJ. Generative deep learning. In: Deep Learning with R [Internet]. 1st ed. New York: Manning Publications; 2018. p. 250–92. Available from: https://www.manning.com/books/deep-learning-with-r

333. Sharma S, Sharma S, Athaiya A. ACTIVATION FUNCTIONS IN NEURAL NETWORKS. International Journal of Engineering Applied Sciences and Technology [Internet]. 2020 [cited 2024 Sep 24];4:310–6. Available from: http://www.ijeast.com

334. Glorot X, Bordes A, Bengio Y. Deep Sparse Rectifier Neural Networks [Internet]. JMLR Workshop and Conference Proceedings; 2011 [cited 2024 Sep 24]. p. 315–23. Available from: https://proceedings.mlr.press/v15/glorot11a.html

335. Sun Y, Wang X, Tang X. Deeply learned face representations are sparse, selective, and robust. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition [Internet]. 2014 Dec 3 [cited 2024 Sep 24];07-12-June-2015:2892–900. Available from: https://arxiv.org/abs/1412.1265v1

336. Lu L, Shin Y, Su Y, Karniadakis GE. Dying ReLU and Initialization: Theory and Numerical Examples. Commun Comput Phys [Internet]. 2019 Mar 15 [cited 2024 Sep 24];28(5):1671–706. Available from: http://arxiv.org/abs/1903.06733

337. Chollet F, Allaire JJ. The engine of neural networks: gradient-based optimisation. In: 1, editor. Deep Learning with R2 [Internet]. 1st ed. New York: Manning Publications; 2018. p. 41–7. Available from: https://www.manning.com/books/deep-learning-with-r

338. Kingma DP, Ba JL. Adam: A Method for Stochastic Optimization. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings [Internet]. 2014 Dec 22 [cited 2024 Sep 24]; Available from: https://arxiv.org/abs/1412.6980v9

339. Sun S, Cao Z, Zhu H, Zhao J. A Survey of Optimization Methods from a Machine Learning Perspective. 2019;

340. Ho SB, Haque R, Chai I, Abdullah A. Optimised deep neural network model to predict asthma exacerbation based on personalised weather triggers. F1000Res [Internet].

2021 [cited 2024 Sep 24];10. Available from: https://pubmed.ncbi.nlm.nih.gov/34745565/

341. Abitbol E, Miere A, Excoffier JB, Mehanna CJ, Amoroso F, Kerr S, et al. Deep learning-based classification of retinal vascular diseases using ultra-widefield colour fundus photographs. BMJ Open Ophthalmol [Internet]. 2022 Feb 4 [cited 2024 Sep 24];7(1). Available from: https://pubmed.ncbi.nlm.nih.gov/35141420/

342. What is Binary Cross Entropy? Calculation & Its Significance [Internet]. [cited 2024 Sep 24]. Available from: https://deepchecks.com/glossary/binary-cross-entropy/

343. Mean Square Error - an overview | ScienceDirect Topics [Internet]. [cited 2024 Sep 24]. Available from: https://www.sciencedirect.com/topics/engineering/mean-square-error

344. Dubey A, Abha Choubey A. A Systematic Review on K-Means Clustering Techniques. International Journal of Scientific Research Engineering & Technology (IJSRET) [Internet]. 2017 [cited 2024 Sep 23];6(6). Available from: www.ijsret.org

345. pallu pallavi. A Systematic Review on K-Means Clustering Techniques [Internet]. [cited 2024 Sep 23]. Available from: https://www.academia.edu/35874865/A_Systematic_Review_on_K_Means_Clustering_Techniques

346. Chong Bao. K-means clustering algorithm: a brief review. Academic Journal of Computing & Information Science [Internet]. 2021 Sep 20 [cited 2024 Sep 23];4(5):37–40. Available from: http://www.francis-press.com/index.php/papers/4643

347. Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. Appl Stat. 1979;28(1):100.

348. kmeans function - RDocumentation [Internet]. [cited 2024 Sep 23]. Available from: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/kmeans

349. Patel E, Singh Kushwaha D. ScienceDirect Third International Conference on Computing and Network Communications (CoCoNet'19) Clustering Cloud Workloads: K-Means vs Gaussian Mixture Model. Procedia Comput Sci [Internet]. 2020 [cited 2024 Sep 23];171:158–67. Available from: www.sciencedirect.comwww.sciencedirect.comwww.elsevier.com/locate/procedia

350. Botyarov M, Miller EE. Partitioning around medoids as a systematic approach to generative design solution space reduction. Results in Engineering. 2022 Sep 1;15:100544.

351. Understanding Gaussian Mixture Models: A Comprehensive Guide | by Juan C Olamendy | Medium [Internet]. [cited 2024 Sep 23]. Available from: https://medium.com/@juanc.olamendy/understanding-gaussian-mixture-models-a-comprehensive-guide-df30af59ced7

352. John CR, Watson D, Russ D, Goldmann K, Ehrenstein M, Pitzalis C, et al. M3C: Monte Carlo reference-based consensus clustering. bioRxiv [Internet]. 2019 Dec 8 [cited 2024 Sep 17];377002. Available from: https://www.biorxiv.org/content/10.1101/377002v2

353. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987 Nov 1;20(C):53–65.

354. Understanding K-means Clustering. Unraveling K-means Clustering: From… | by Nermeen Abd El-Hafeez | Medium [Internet]. [cited 2024 Sep 23]. Available from: https://medium.com/@nerminhafeez2002/understanding-k-means-clustering-3716731cc648

355. The Limitation Of Silhouette Score Which Is Often Ignored By Many [Internet]. [cited 2024 Sep 23]. Available from: https://blog.dailydoseofds.com/p/the-limitation-of-silhouette-score

356. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. Mach Learn. 2003 Jul;52(1–2):91–118.

357. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. Mach Learn. 2003 Jul;52(1–2):91–118.

358. The Cancer Genome Atlas Program (TCGA) - NCI [Internet]. [cited 2024 Sep 23]. Available from: https://www.cancer.gov/ccg/research/genome-sequencing/tcga

359. CRAN: Package dagitty [Internet]. [cited 2024 Oct 7]. Available from: https://cran.r-project.org/web/packages/dagitty/index.html

360. Lipsky AM, Greenland S. Causal Directed Acyclic Graphs. JAMA [Internet]. 2022 Mar 15 [cited 2024 Oct 2];327(11):1083–4. Available from: https://jamanetwork.com/journals/jama/fullarticle/2789646

361. Watson ML, Hickman SHM, Dreesbeimdiek KM, Kohler K, Stubbs DJ. Directed acyclic graphs in perioperative observational research-A systematic review and critique against best practice recommendations. PLoS One [Internet]. 2023 Feb 1 [cited 2024 Oct 2];18(2). Available from: https://pubmed.ncbi.nlm.nih.gov/36758007/

362. Digitale JC, Martin JN, Glymour MM. Tutorial on directed acyclic graphs. J Clin Epidemiol [Internet]. 2022 Feb 1 [cited 2024 Oct 2];142:264–7. Available from: http://www.jclinepi.com/article/S0895435621002407/fulltext

363. Tingley D, Yamamoto HT, Hirose K, Keele L, Princeton KI. mediation: R Package for Causal Mediation Analysis. [cited 2024 Oct 7]; Available from: http://CRAN.R-project.org/package=mediation

364. Yu Q, Scribner RA, Leonardi C, Zhang L, Park C, Chen L, et al. Package 'mma''.' Spat Spatiotemporal Epidemiol. 2017 Jun 1;21:13–23.

365. Brown VA. An Introduction to Linear Mixed-Effects Modeling in R. Adv Methods Pract Psychol Sci [Internet]. 2021 Mar 25 [cited 2024 Oct 16];4(1). Available from: https://journals.sagepub.com/doi/10.1177/2515245920960351

366. Verbyla AP. A note on model selection using information criteria for general linear models estimated using REML. Aust N Z J Stat [Internet]. 2019 Mar 1 [cited 2024 Oct 16];61(1):39–50. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/anzs.12254

367. Varady NH, Pareek A, Eckhardt CM, Williams RJ, Madjarova SJ, Ollivier M, et al. Multivariable regression: understanding one of medicine's most fundamental statistical tools. Knee Surgery, Sports Traumatology, Arthroscopy [Internet]. 2023 Jan 1 [cited 2025 Jan 21];31(1):7–11. Available from: https://link.springer.com/article/10.1007/s00167-022-07215-9

368. Ke X, Keenan K, Smith VA. Treatment of missing data in Bayesian network structure learning: an application to linked biomedical and social survey data. BMC Med Res Methodol [Internet]. 2022 Dec 1 [cited 2024 Oct 9];22(1):1–16. Available from: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-022-01781-9

369. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. J Stat Softw. 2011 Dec;45(3):1–67.

370. Soloski MJ, Chrest FJ. Multi-parameter flow cytometry for discovery of disease mechanisms in rheumatic diseases. Arthritis Rheum [Internet]. 2013 May [cited 2024 Oct 11];65(5):1148. Available from: /pmc/articles/PMC3636160/

371. Lu Z, Li W, Tang Y, Da Z, Li X. Lymphocyte subset clustering analysis in treatment-naive patients with systemic lupus erythematosus. Clin Rheumatol [Internet]. 2021 May 1 [cited 2025 Jan 21];40(5):1835–42. Available from: https://link.springer.com/article/10.1007/s10067-020-05480-y

372. CRAN: Package cluster [Internet]. [cited 2024 Oct 14]. Available from: https://cran.r-project.org/web/packages/cluster/index.html

373. Stifano G, Christmann RB. Macrophage Involvement in Systemic Sclerosis: Do We Need More Evidence? Curr Rheumatol Rep [Internet]. 2016 Dec 1 [cited 2024 Oct 16];18(1):1–6. Available from: https://pubmed.ncbi.nlm.nih.gov/26700912/

374. Villanueva-Martin G, Acosta-Herrera M, Carmona EG, Kerick M, Ortego-Centeno N, Callejas-Rubio JL, et al. Non-classical circulating monocytes expressing high levels of microsomal prostaglandin E2 synthase-1 tag an aberrant IFN-response in systemic sclerosis. J Autoimmun [Internet]. 2023 Nov 1 [cited 2024 Oct 16];140. Available from: https://pubmed.ncbi.nlm.nih.gov/37633117/

375. Kubo S, Nakayamada S, Yoshikawa M, Miyazaki Y, Sakata K, Nakano K, et al. Peripheral Immunophenotyping Identifies Three Subgroups Based on T Cell Heterogeneity in Lupus Patients. Arthritis and Rheumatology. 2017 Oct 1;69(10):2029–37.

376. Robinson GA, Peng J, Dönnes P, Coelewij L, Naja M, Radziszewska A, et al. Disease-associated and patient-specific immune cell signatures in juvenile-onset systemic lupus erythematosus: patient stratification using a machine-learning approach. Lancet Rheumatol [Internet]. 2020 Aug 1 [cited 2025 Jan 21];2(8):e485–96. Available from: http://www.thelancet.com/article/S2665991320301685/fulltext

377. Mosca M, Tani C, Vagnani S, Carli L, Bombardieri S. The diagnosis and classification of undifferentiated connective tissue diseases. J Autoimmun. 2014;48–49:50–2.

378. Barturen G, Babaei S, Català-Moll F, Martínez-Bueno M, Makowska Z, Martorell-Marugán J, et al. Integrative Analysis Reveals a Molecular Stratification of Systemic Autoimmune Diseases. Arthritis Rheumatol. 2021 Jun 1;73(6):1073–85.

379. Petri M, Bruce IN, Dörner T, Tanaka Y, Morand EF, Kalunian KC, et al. Baricitinib for systemic lupus erythematosus: a double-blind, randomised, placebo-controlled, phase 3 trial (SLE-BRAVE-II). The Lancet. 2023 Mar 25;401(10381):1011–9.

380. Vital EM, Merrill JT, Morand EF, Furie RA, Bruce IN, Tanaka Y, et al. Anifrolumab efficacy and safety by type I interferon gene signature and clinical subgroups in patients with SLE: post hoc analysis of pooled data from two phase III trials. Ann Rheum Dis. 2022 Jul 1;81(7):951–61.

381. Morand EF, Vital EM, Petri M, van Vollenhoven R, Wallace DJ, Mosca M, et al. Baricitinib for systemic lupus erythematosus: a double-blind, randomised, placebo-controlled, phase 3 trial (SLE-BRAVE-I). The Lancet. 2023 Mar 25;401(10381):1001–10.

382. Furie RA, Morand EF, Bruce IN, Manzi S, Kalunian KC, Vital EM, et al. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. Lancet Rheumatol. 2019 Dec 1;1(4):e208–19.

383. Chiara T, Linda C, Andrea DS, Lucrezia R, Sabrina V, Rosaria T, et al. Impact of Joint Involvement on Patients Reported Outcomes in Systemic Lupus Erythematosus. Curr Rheumatol Rev. 2018 Mar 15;14(2):188–92.

384. Abel GA, Barclay ME, Payne RA. Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates. BMJ Open. 2016 Nov 1;6(11).

385. Charlson ME, Charlson RE, Peterson JC, Marinopoulos SS, Briggs WM, Hollenberg JP. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. J Clin Epidemiol. 2008 Dec;61(12):1234–40.

386. El-Sherbiny YM, Psarras A, Yusof MYM, Hensor EMA, Tooze R, Doody G, et al. A novel two-score system for interferon status segregates autoimmune diseases and correlates with clinical features. Sci Rep. 2018 Dec 1;8(1).

387. Yusof MYM, Psarras A, El-Sherbiny YM, Hensor EMA, Dutton K, Ul-Hassan S, et al. Prediction of autoimmune connective tissue disease in an at-risk cohort: prognostic value of a novel two-score system for interferon status. Ann Rheum Dis. 2018 Jul 20;77(10):1432–9.

388. Carter LM, Alase A, Wigston Z, Psarras A, Burska A, Sutton E, et al. Gene expression and autoantibody analysis reveals distinct ancestry-specific profiles associated with response to rituximab in refractory systemic lupus erythematosus. Arthritis Rheumatol. 2022 Nov 21;

389. SankeyMATIC: Make Beautiful Flow Diagrams [Internet]. [cited 2025 Jan 21]. Available from: https://sankeymatic.com/

390. Furie RA, van Vollenhoven RF, Kalunian K, Navarra S, Romero-Diaz J, Werth VP, et al. Trial of Anti-BDCA2 Antibody Litifilimab for Systemic Lupus Erythematosus. New England Journal of Medicine. 2022 Sep 8;387(10):894–904.

391. Mahmoud K, Zayat A, Yusof MYM, Ciurtin C, Yee CS, Isenberg D, et al. FRI0599 USEFUL II: derivation of the lupus arthritis and musculoskeletal disease activity score (lamda) using data from a multicentre longitudinal study. Ann Rheum Dis. 2020 Jun 1;79(Suppl 1):905–6.

392. Unterhuber M, Kresoja KP, Rommel KP, Besler C, Baragetti A, Klöting N, et al. Proteomics-Enabled Deep Learning Machine Algorithms Can Enhance Prediction of Mortality. J Am Coll Cardiol. 2021 Oct 19;78(16):1621–31.

393. Patrick MT, Stuart PE, Raja K, Gudjonsson JE, Tejasvi T, Yang J, et al. Genetic signature to provide robust risk assessment of psoriatic arthritis development in psoriasis patients. Nat Commun. 2018 Dec 1;9(1).

394. Jorge A, Castro VM, Barnado A, Gainer V, Hong C, Cai T, et al. Identifying lupus patients in electronic health records: Development and validation of machine learning algorithms and application of rule-based algorithms. Semin Arthritis Rheum. 2019 Aug 1;49(1):84–90.

395. Goodfellow I, Bengio Y, Courville A. No Title. In: Deep Learning. MIT Press; 2016. p. 499–523.

396. Li R, Hu B, Liu F, Liu W, Cunningham F, McManus DD, et al. Detection of Bleeding Events in Electronic Health Record Notes Using Convolutional Neural Network Models Enhanced With Recurrent Neural Network Autoencoders: Deep Learning Approach. JMIR Med Inform 2019;7(1):e10788 https://medinform.jmir.org/2019/1/e10788. 2019 Feb 8;7(1):e10788.

397. Miller MI, Orfanoudaki A, Cronin M, Saglam H, So Yeon Kim I, Balogun O, et al. Natural Language Processing of Radiology Reports to Detect Complications of Ischemic Stroke. Neurocrit Care. 2022 Aug 1;37(Suppl 2):291–302.

398. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. [cited 2023 Nov 7]; Available from: www.tensorflow.org.

399. Developers T. TensorFlow [Internet]. Zenodo; 2024. Available from: https://doi.org/10.5281/zenodo.10798587

400. Chollet F, others. Keras. 2015.

401. Developers T. TensorFlow. Zenodo; 2024.

402. Allaire JJ, K Ushey, Tang Y. reticulate: Interface to Python. 2023.

403. Scrucca L, Fraley C, Murphy TB, Raftery AE. Model-Based Clustering, Classification, and Density Estimation Using {mclust} in {R}. Chapman and Hall/CRC; 2023.

404. Lu L, Shin Y, Su Y, Karniadakis GE. Dying ReLU and initialization: Theory and numerical examples. Commun Comput Phys. 2020;28(5):1671–706.

405. Kingma DP, Ba JL. Adam: A method for stochastic optimization. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. 2015;1–15.

406. John CR, Watson D, Russ D, Goldmann K, Ehrenstein M, Pitzalis C, et al. M3C: Monte Carlo reference-based consensus clustering. Scientific Reports 2020 10:1. 2020 Feb 4;10(1):1–14.

407. Therneau TM. A Package for Survival Analysis in R. 2024.

408. Chiche L, Jourde-Chiche N, Whalen E, Presnell S, Gersuk V, Dang K, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. Arthritis Rheumatol. 2014;66(6):1583–95.

409. Shiboski CH, Shiboski SC, Seror R, Criswell LA, Labetoulle M, Lietman TM, et al. 2016 ACR-EULAR Classification Criteria for primary Sjögren's Syndrome: A Consensus and Data-Driven Methodology Involving Three International Patient Cohorts. Arthritis Rheumatol [Internet]. 2017 Jan 1 [cited 2024 Sep 26];69(1):35. Available from: /pmc/articles/PMC5650478/

410. Chaussabel D, Quinn C, Shen J, Patel P, Glaser C, Baldwin N, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. Immunity [Internet]. 2008 Jul 7 [cited 2024 Oct 7];29(1):150. Available from: /pmc/articles/PMC2727981/

411. Merrill JT, Werth VP, Furie R, van Vollenhoven R, Dörner T, Petronijevic M, et al. Phase 2 Trial of Iberdomide in Systemic Lupus Erythematosus. New England Journal of Medicine [Internet]. 2022 Mar 17 [cited 2024 Oct 7];386(11):1034–45. Available from: https://www.nejm.org/doi/full/10.1056/NEJMoa2106535

412. Novartis Pharmaceuticals. PD of VAY736 in Patients With Primary Sjögren's Syndrome (CVAY736X2201). 2024.

413. Anifrolumab Treatment for 24 Weeks in Patients With Primary Sjögren's Syndrome [Internet]. [cited 2024 Oct 7]. Available from: https://ctv.veeva.com/study/anifrolumab-treatment-for-24-weeks-in-patients-with-primary-sjogrens-syndrome

414. Determine effectiveness of Anifrolumab In SYstemic sclerosis (DAISY) [Internet]. [cited 2024 Oct 7]. Available from: https://www.astrazenecaclinicaltrials.com/study/D3460C00002/

415. Toro-Domínguez D, Martorell-Marugán J, Martinez-Bueno M, López-Domínguez R, Carnero-Montoro E, Barturen G, et al. Scoring personalized molecular portraits identify Systemic Lupus Erythematosus subtypes and predict individualized drug responses, symptomatology and disease progression. Brief Bioinform [Internet]. 2022 Aug 10 [cited 2023 Nov 7];23(5):1–13. Available from: https://digibug.ugr.es/handle/10481/79808

416. Home | 3TR [Internet]. [cited 2024 Oct 7]. Available from: https://3tr-imi.eu/

417. Lindblom J, Toro-Domínguez D, Carnero-Montoro E, Beretta L, Borghi MO, Castillo J, et al. Distinct gene dysregulation patterns herald precision medicine potentiality in systemic lupus erythematosus. J Autoimmun. 2023 Apr 1;136:103025.

418. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature 2020 584:7821

[Internet]. 2020 Jul 8 [cited 2024 Jul 17];584(7821):430–6. Available from: https://www.nature.com/articles/s41586-020-2521-4

419. Westreich D, Edwards JK, Van Smeden M. Comment on Williamson et al. (OpenSAFELY): The Table 2 Fallacy in a Study of COVID-19 Mortality Risk Factors. Epidemiology [Internet]. 2021 Jan 1 [cited 2024 Jul 17];32(1):E1–2. Available from: https://pubmed.ncbi.nlm.nih.gov/33065610/

420. Lipsky AM, Greenland S. Causal Directed Acyclic Graphs. JAMA [Internet]. 2022 Mar 15 [cited 2024 Jul 17];327(11):1083–4. Available from: https://jamanetwork.com/journals/jama/fullarticle/2789646

421. CRAN: Package ggplot2 [Internet]. [cited 2024 Jul 30]. Available from: https://cran.r-project.org/web/packages/ggplot2/index.html

422. Introduction to tableone [Internet]. [cited 2024 Oct 7]. Available from: https://cran.r-project.org/web/packages/tableone/vignettes/introduction.html

423. corrplot function - RDocumentation [Internet]. [cited 2024 Oct 7]. Available from: https://www.rdocumentation.org/packages/corrplot/versions/0.94/topics/corrplot

424. heatmap function - RDocumentation [Internet]. [cited 2024 Oct 7]. Available from: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/heatmap

425. stats package - RDocumentation [Internet]. [cited 2025 Jan 21]. Available from: https://www.rdocumentation.org/packages/stats/versions/3.6.2

426. lme4 package - RDocumentation [Internet]. [cited 2024 Jul 30]. Available from: https://www.rdocumentation.org/packages/lme4/versions/1.1-35.5

427. Companion to Applied Regression [R package car version 3.1-3]. 2024 Sep 27 [cited 2024 Oct 7]; Available from: https://CRAN.R-project.org/package=car

428. Chiche L, Seguier J, Gentile S, Burtey S, Dussol B, Halfon P, et al. Unexpected Association between Health-Related Quality of Life and the Blood Interferon Modular Transcriptional Signatures in Patients with Systemic Lupus Erythematosus. 2018 Oct 19 [cited 2024 Oct 14];70(9S):1096. Available from: https://hal.science/hal-01935047

429. Papanikolaou S, Kosmara D, Stathopoulou C, Sidiropoulos P, Konstantopoulos D, Bertsias G. Deciphering the Molecular Mechanism of Flares in Patients with Systemic Lupus Erythematosus through Single-Cell Transcriptome Analysis of the Peripheral Blood. Mediterr J Rheumatol. 2022;33(1):94.

430. Mai L, Asaduzzaman A, Noamani B, Fortin PR, Gladman DD, Touma Z, et al. The baseline interferon signature predicts disease severity over the subsequent 5 years in systemic lupus erythematosus. Arthritis Res Ther [Internet]. 2021 Dec 1 [cited 2024 Oct 9];23(1):1–9. Available from: https://link.springer.com/articles/10.1186/s13075-021-02414-0

431. Hosack T, Thomas T, Ravindran R, Uhlig HH, Travis SPL, Buckley CD. Inflammation across tissues: can shared cell biology help design smarter trials? Nature Reviews Rheumatology 2023 19:10 [Internet]. 2023 Sep 4 [cited 2024 Oct 9];19(10):666–74. Available from: https://www.nature.com/articles/s41584-023-01007-2

432. Heinrich MC, Joensuu H, Demetri GD, Corless CL, Apperley J, Fletcher JA, et al. Phase II, open-label study evaluating the activity of imatinib in treating life-threatening malignancies known to be associated with imatinib-sensitive tyrosine kinases. Clin Cancer Res. 2008 May 1;14(9):2717–25.

433. Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, Demetri GD, et al. Efficacy of Larotrectinib in TRK Fusion–Positive Cancers in Adults and Children. N Engl J Med [Internet]. 2018 Feb 2 [cited 2024 Oct 9];378(8):731. Available from: /pmc/articles/PMC5857389/

434. Barturen G, Beretta L, Cervera R, Van Vollenhoven R, Alarcón-Riquelme ME. Moving towards a molecular taxonomy of autoimmune rheumatic diseases [Internet]. Vol. 14, Nature Reviews Rheumatology. Nature Publishing Group; 2018 [cited 2021 Jun 28]. p. 75–93. Available from: https://pubmed.ncbi.nlm.nih.gov/29362467/

435. Carnero-Montoro E, Barturen G, Povedano E, Kerick M, Martinez-Bueno M, Ballestar E, et al. Epigenome-Wide Comparative Study Reveals Key Differences Between Mixed Connective Tissue Disease and Related Systemic Autoimmune Diseases. Front Immunol [Internet]. 2019 [cited 2023 Aug 14];10(AUG). Available from: https://pubmed.ncbi.nlm.nih.gov/31440254/

436. Toro-Domínguez D, Martorell-Marugán J, Martinez-Bueno M, López-Domínguez R, Carnero-Montoro E, Barturen G, et al. Scoring personalized molecular portraits identify Systemic Lupus Erythematosus subtypes and predict individualized drug responses, symptomatology and disease progression. Brief Bioinform [Internet]. 2022 Sep 20 [cited 2024 Oct 2];23(5):1–13. Available from: https://dx.doi.org/10.1093/bib/bbac332

437. Mahmoud K, Zayat A, Yusof MYM, Ciurtin C, Yee CS, Isenberg D, et al. FRI0599 USEFUL II: Derivation of the lupus arthritis and musculoskeletal disease activity score (lamda) using data from a multicentre longitudinal study. Ann Rheum Dis [Internet]. 2020 Jun 1 [cited 2022 Mar 24];79(Suppl 1):905–6. Available from: https://ard.bmj.com/content/79/Suppl_1/905.2

438. Rovin BH, Furie R, Latinis K, Looney RJ, Fervenza FC, Sanchez-Guerrero J, et al. Efficacy and safety of rituximab in patients with active proliferative lupus nephritis: the Lupus Nephritis Assessment with Rituximab study. Arthritis Rheum. 2012;

439. Merrill JT, Buyon JP, Furie RA, Latinis KM, Gordon C, Hsieh HJ, et al. Assessment of flares in lupus patients enrolled in a phase II/III study of rituximab (EXPLORER). Lupus. 2011 Jun;20(7):709–16.

440. Petri M, Bruce IN, Dörner T, Tanaka Y, Morand EF, Kalunian KC, et al. Baricitinib for systemic lupus erythematosus: a double-blind, randomised, placebo-controlled, phase 3 trial (SLE-BRAVE-II). The Lancet [Internet]. 2023 Mar 25 [cited 2023 Aug 2];401(10381):1011–9. Available from: http://www.thelancet.com/article/S0140673622025466/fulltext

441. Turner CA, Jacobs AD, Marques CK, Oates JC, Kamen DL, Anderson PE, et al. Word2Vec inversion and traditional text classifiers for phenotyping lupus. BMC Med Inform Decis Mak [Internet]. 2017 Aug 22 [cited 2024 Oct 11];17(1):1–11. Available from: https://link.springer.com/articles/10.1186/s12911-017-0518-1

442. Boquet G, Morell A, Serrano J, Vicario JL. A variational autoencoder solution for road traffic forecasting systems: Missing data imputation, dimension reduction, model selection and anomaly detection. Transp Res Part C Emerg Technol. 2020 Jun 1;115:102622.

443. Liu X, Zhang L, Guo Z, Han T, Ju M, Xu B, et al. Medical Image Compression Based on Variational Autoencoder. Math Probl Eng [Internet]. 2022 Jan 1 [cited 2024 Oct 9];2022(1):7088137. Available from: https://onlinelibrary.wiley.com/doi/full/10.1155/2022/7088137