Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Medicine 1958

Genetic Contributions to Manifestations and Prognosis of Systemic Lupus Erythematosus

SARAH REID

ACTA UNIVERSITATIS
UPSALIENSIS
2023

ISSN 1651-6206
ISBN 978-91-513-1848-6
urn:nbn:se:uu:diva-506789

UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in Rudbecksalen, Rudbecklaboratoriet, Dag Hammarskjölds väg 20, Uppsala, Thursday, 14 September 2023 at 09:00 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English. Faculty examiner: Professor Timothy Niewold (Hospital for Special Surgery, New York).

Abstract

Reid, S. 2023. Genetic Contributions to Manifestations and Prognosis of Systemic Lupus Erythematosus. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 1958. 79 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-1848-6.

Systemic lupus erythematosus (SLE) is a severe autoimmune disease with a multifactorial aetiology and an extremely heterogenous clinical picture. This thesis explores the impact of SLE genetic variants, through their cumulative effect or gene-environment interactions, on SLE manifestations and complications.

Paper I identifies clinical risk factors for permanent organ damage and mortality in 543 Swedish patients with SLE. Several previously undescribed risk factors for organ damage were identified, including pericarditis, lymphopenia and haemolytic anaemia. After a mean disease duration of 17 years, more than half of the patients were found to have developed organ damage, demonstrating that long-term SLE consequences remain a critical challenge.

Paper II highlights the impact of a high cumulative genetic risk on disease severity, in 1001 Swedish patients with SLE. A genetic risk score (GRS) based on 57 genetic variants with established SLE association was constructed, and patients with high and low GRSs were compared with regard to clinical manifestations and complications. Patients with a high GRS were found to develop SLE earlier in life, to more often be affected by organ damage, cardiovascular events, nephritis and end-stage renal disease, and to have a higher mortality risk.

Paper III demonstrates gene-smoking interactions in two separate samples of Scandinavian patients with SLE, totalling 1 612 patients. In both cohorts, the combination of two SLE risk alleles located within the Signal Transducer and Activator of Transcription 4 (STAT4) and Interleukin-12A (IL12A) genes, which share a connection through the IL12-STAT4 signalling pathway, was associated with a substantial increase in the risk of myocardial infarction (MI). We further demonstrated that smoking enhances the effect of the STAT4 risk allele on MI and is associated with higher levels of phosphorylated STAT4 in CD8+ T cells. Together, these results indicate that IL12-STAT4 pathway activation is associated with MI in SLE.

Paper IV describes the cumulative impact of 57 SLE risk variants on disease manifestations in a large biobank cohort comprised primarily of individuals without SLE. A high genetic SLE predisposition was associated with several manifestations resembling SLE, including skin and pleural manifestations, arthritis and hematologic disorders. Furthermore, based on the findings in the biobank cohort, GRSs for several SLE-like manifestations could be constructed which were associated with the corresponding manifestation in patients with SLE.

In conclusion, this thesis provides new evidence for the importance of both genetic and clinical factors, and their interaction, for SLE prognosis.

Keywords: Systemic lupus erythematosus, genetics, cardiovascular disease, polygenic risk score, smoking

Sarah Reid, Department of Medical Sciences, Rheumatology, Akademiska sjukhuset, Uppsala University, SE-75185 Uppsala, Sweden.

ISSN 1651-6206
ISBN 978-91-513-1848-6
URN urn:nbn:se:uu:diva-506789 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-506789)

To Anthony, Ethan and Alba

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I. Frodlund, A., **Reid, S.**, Wetterö, J., Dahlstrom, Ö., Sjöwall, C., Leonard D. (2019) The majority of Swedish systemic lupus erythematosus patients are still affected by irreversible organ impairment: factors related to damage accrual in two regional cohorts. *Lupus*, 28(10):1261–1272

II. **Reid, S**., Alexsson, A., Frodlund, M., Morris, D., Sandling, J.K., Bolin, K., Svenungsson, E., Jönsen, A., Bengtsson, C., Gunnarsson, I., Illescas Rodriguez, V., Bengtsson, A., Arve, S., Rantapää-Dahlqvist, S., Eloranta, M.L., Syvänen, A.C., Sjöwall, C., Vyse, T.J., Rönnblom, L., Leonard, D. (2020) High genetic risk score is associated with early disease onset, damage accrual and decreased survival in systemic lupus erythematosus. *Ann Rheum Dis*, 79(3):363– 369.

III. **Reid, S**., Hagberg, N., Sandling, J.K., Alexsson, A., Pucholt, P., Sjöwall, C., Lerang, K., Jönsen, A., Gunnarsson, I., Syvänen, A.C., Troldborg, A.M., Voss, A., Bengtsson, A.A., Molberg, Ø., Jacobsen, S., Svenungsson, E., Rönnblom, L., Leonard, D. (2021) Interaction between the STAT4 rs11889341(T) risk allele and smoking confers increased risk of myocardial infarction and nephritis in patients with systemic lupus erythematosus. *Ann Rheum Dis*, 80(9):1183–1189

IV. **Reid, S**., Sandling, J.K., Pucholt, P., Sayadi, A., Sjöwall, C., Frodlund, M., Lerang, K., Jönsen, A., Syvänen, A.C., Bengtsson, A.A., Molberg, Ø., Svenungsson, E., Gunnarsson, I., Rudin, A., Rantapää-Dahlqvist, S., Rönnblom, L., Leonard, D. (2023) From genetic predisposition to clinical outcome in systemic lupus erythematosus: construction and validation of subphenotype-specific genetic risk scores. *Manuscript.*

Supporting publications:

Segerberg F., Lundtoft C., **Reid S.**, Hjorton K., Leonard D., Nordmark G., Carlsten M., Hagberg N. (2019) Autoantibodies to Killer Cell Immunoglobulin-Like Receptors in Patients With Systemic Lupus Erythematosus Induce Natural Killer Cell Hyporesponsiveness. *Front Immunol*. 10:2164.

Hagberg N., Joelsson M., Leonard D., **Reid S.**, Eloranta M.L., Mo J., Nilsson M.K., Syvänen A.C., Bryceson Y.T., Rönnblom, L. (2018) The STAT4 SLE risk allele rs7574865[T] is associated with increased IL-12-induced IF-γ production in T cells from patients with SLE. *Ann Rheum Dis*. 77(7):1070-1077.

Yavuz S., Cansu D.U., Nikolopoulos D., Crisafulli F., Antunes A.M., Adamichou C., **Reid S.**, Stagnaro C., Andreoli L., Tincani A., Moraes-Fontes M.F., Mosca M., Leonard D., Jönsen A., Bengtsson A., Svenungsson E., Gunnarsson I., Dahlqvist S.R., Sjöwall C., Bertsias G., Fanouriakis A., Rönnblom L. (2020) Lymphopenia as a risk factor for neurologic involvement and organ damage accrual in patients with systemic lupus erythematosus: A multi-center observational study. *Semin Arthritis Rheum*. 50(6):1387-1393

Reprints were made with permission from the respective publishers.

Contents

Abbreviations

A	Adenine
aCL	Anti-cardiolipin
ACR	American College of Rheumatology
AE	Arterial events
AIHA	Autoimmune haemolytic anaemia
ANA	Anti-nuclear antibodies
APC	Antigen-presenting cell
aPL	Anti-phospholipid antibodies
APS	Anti-phospholipid syndrome
ATG5	Autophagy Related 5
ATXN2	Ataxin 2
AUC	Area under the ROC Curve
aβ2GP1	Anti-β2-Glycoprotein 1
BANK1	B Cell Scaffold Protein with Ankyrin Repeats 1
BLK	BLK proto-oncogene
C	Cytosine
C1q	Complement component 1q
CA	Custom array
CAD	Coronary artery disease
CD4	Cluster of differentiation 4
CD8	Cluster of differentiation 8
CH50	50% haemolytic complement
CKD	Chronic kidney disease
CLEC16A	C-Type Lectin Domain Containing 16A
cSLE	Childhood-onset SLE
CVD	Cardiovascular disease
CVE	Cardiovascular events
DNA	Deoxyribonucleic Acid
dsDNA	Double-stranded DNA
DZ	Dizygotic
ECG	Electrocardiogram
ELISA	Enzyme-linked immunosorbent assay
ESRD	End-Stage Renal Disease
G	Guanine
GFR	Glomerular filtration rate

GRS	Genetic risk score
GSA	Global Screening Array
GWAS	Genome-wide association study
HLA	Human Leukocyte Antigen
IC	Immunarray chip
ICD	International classification of diseases
ICVD	Ischemic cerebrovascular disease
IFN	Interferon
Ig	Immunoglobulin
IKZF1	Ikaros Family Zinc Finger 1
IL-10	Interleukin 10
IL-12	Interleukin 12
IL12A	Interleukin 12A
IL12B	Interleukin 12B
IL12RB2	Interleukin 12 Receptor Subunit Beta 2
ILD	Interstitial lung disease
IRF5	Interferon regulatory factor 5
IRF8	Interferon regulatory factor 8
ITGAM	Integrin Subunit Alpha M
IV	Instrumental variable
IVW	Inverse variance weighted
JAK	Janus kinase
JAK2	Janus kinase 2
LA	Lupus anticoagulant
LD	Linkage disequilibrium
LE	Lupus erythematosus
LN	Lupus nephritis
MAF	Minor allele frequency
MI	Myocardial infarction
MS	Multiple sclerosis
MZ	Monozygotic
NET	Neutrophil extracellular trap
OR	Odds ratio
PRS	Polygenic risk score
pSTAT4	Phosphorylated STAT4
RA	Rheumatoid arthritis
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
RPR	Rapid plasma regain
RR	Relative risk
SDI	SLICC damage index
SH2B3	SH2B Adaptor Protein 3
SLE	Systemic Lupus Erythematosus
SLICC	Systemic Lupus International Collaborating Clinics

Sm	Smith
SNP	Single Nucleotide Polymorphism
SOCS1	Suppressor of Cytokine Signalling 1
SS	Sjogren's syndrome
SSA	Anti–Sjogren's syndrome-related antigen A
SSB	Anti–Sjogren's syndrome-related antigen B
STAT4	Signal Transducer and Activator of Transcription
T	Thymine
T1D	Type 1 diabetes mellitus
T2D	Type 2 diabetes mellitus
TNFSF4	TNF Superfamily Member 4
TYK2	Tyrosine Kinase 2
UCLH	University College London Hospital
UK	United Kingdom
UVR	Ultraviolet radiation
VTE	Venous thromboembolism
WHO	World Health Organization

Systemic Lupus Erythematosus

Systemic Lupus Erythematosus (SLE) is a severe, chronic, autoimmune disease characterized by defective elimination of apoptotic cellular material, an activated type I interferon (IFN) system and exaggerated B and T cell responses.[1-5] Humoral autoimmunity, with antibodies mainly directed against double-stranded deoxyribonucleic acid (dsDNA) and small nuclear ribonucleic acid (RNA)-binding proteins, occurs in nearly all individuals with the disease and precedes clinical symptoms by 3-9 years.[6-8] The aetiology of SLE is multifactorial, with environmental and hormonal factors believed to trigger the disease in genetically susceptible individuals (figure 1).[6, 9] The relative risk (RR) of developing SLE is higher in women than in men, with female-to-male ratios ranging between 8:1 and 15:1.[10, 11] This female-to-male predominance varies with age and is the most pronounced between puberty and menopause, indicating a possible importance of sex hormones in the disease development.

Epidemiology

From a global perspective, the occurrence of SLE varies substantially between different geographical locations, with the incidence ranging between 0.3–23.7 cases per 100 000 and the prevalence between 6.5-178 cases per 100 000 adults.[12] These variations may be due both to differences in health care standards between different parts of the world, with likely under-diagnosis of the disease in developing countries, and to differences in susceptibility to the disease between different ethnicities.[13] Whilst the prevalence of SLE is low in Africa,[14] studies performed on the north American and British populations have consistently reported a higher prevalence amongst African American, African-Caribbean, Hispanic, Asian and middle-Eastern individuals than in Caucasians.[10, 15-17] In Scandinavia, the incidence has been estimated to 2.6-5.4/100 000, and in a study from Sweden, the prevalence varied between 46-85/100 000 depending on the definition used for diagnosis.[18-24]

Clinical presentation

The clinical presentation of SLE displays immense variation between individuals, with symptoms ranging from mild skin and joint manifestations to severe cytopenia or life-threatening involvement of internal organs.[9] The disease course usually follows a relapsing-remitting pattern, with periods of higher disease activity followed by periods where clinical symptoms improve and laboratory measures, such as complement levels and anti-dsDNA, normalize. In certain individuals, however, prolonged remission or a persistently active disease may be observed.[25]

Clinical diagnosis

The diagnosis of SLE is clinical, supported by laboratory findings, and requires ruling out other causes of each manifestation. Sometimes the disease is easily identifiable, with the patient presenting with several distinct symptoms such as painful, swollen joints, a malar rash, or signs of serositis and neurologic manifestations such as a seizure or psychosis. Laboratory findings may include leukopenia, thrombocytopenia, anaemia or proteinuria. In some cases, however, the initial presentation may be more diffuse and involve constitutional symptoms such as fatigue and malaise, low-grade fever, myalgia or arthralgia, weight loss, headache and cognitive complaints.[9, 26, 27] As the vast majority of individuals with SLE carry anti-nuclear antibodies (ANA), this analysis has a high negative predictive value for the disease. The positive predictive value is lower, however, as ANA may also occur secondarily to other conditions including infection, malignancy and other autoimmune diseases and is present in moderate titres in around 5% of the population.[27, 28] For this reason, the analysis of ANA should only be carried out on a clinical suspicion and not be used for general screening purposes. In the case of individuals presenting with diffuse symptoms, an analysis of ANA may thus not be carried out immediately and consequently, the diagnosis of SLE may be delayed by months or even years.[9]

The American College of Rheumatology classification criteria

Whilst the Fries and Holman's diagnostic principle, requiring manifestations from 2 separate, defined organ systems in the presence of ANA, are often used in a clinical setting,[26] several sets of classification criteria have been developed to assist in clinical diagnosis and to aid in classifying patients for research purposes, where the identification of a relatively homogenous cohort with an established diagnosis is desired.[29] The 1982 American College of Rheumatology (ACR-82) classification criteria for SLE are widely used for these purposes.[30] The ACR-82 criteria (table 1) consist of 11 different categories including manifestations from the skin, joints, serosae, kidneys and

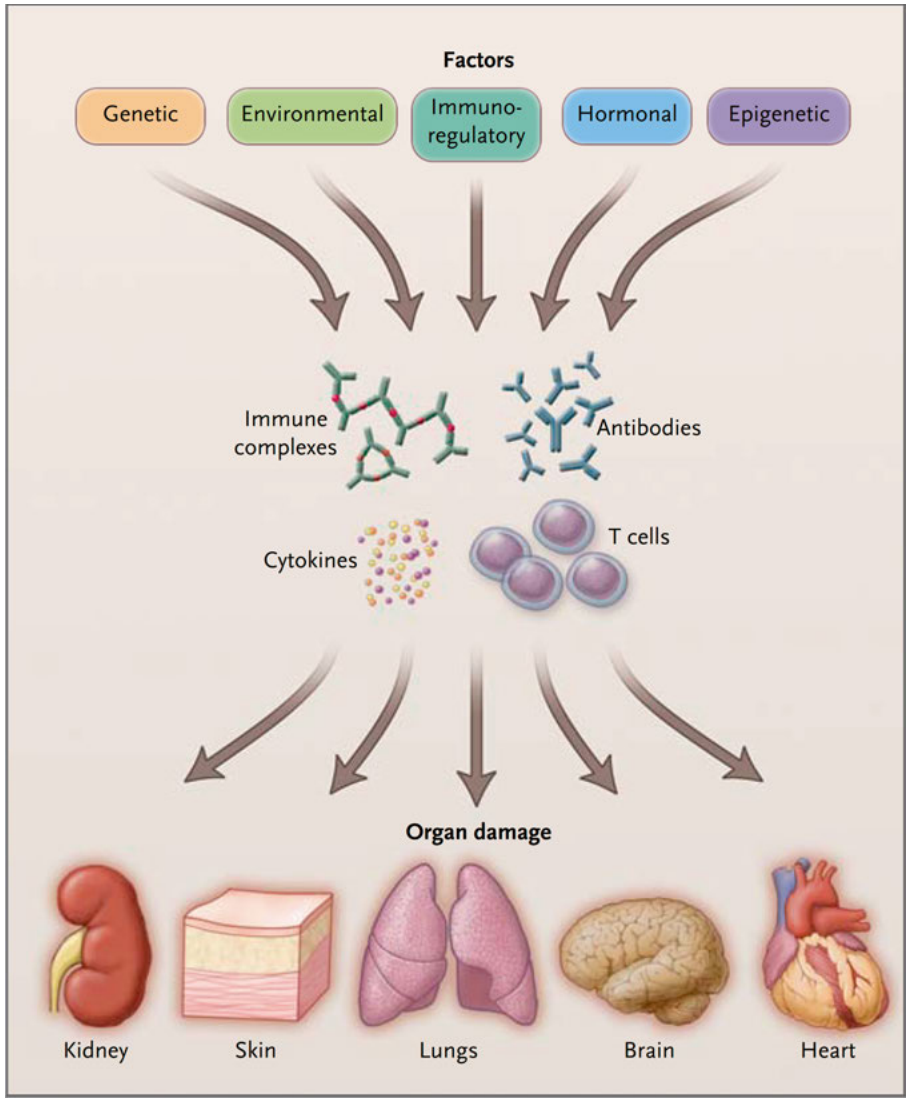

Figure 1. The aetiology and consequences of SLE. Reproduced with permission from George C. Tsokos. Systemic Lupus Erythematosus. N Engl J Med 2011; 365:2110-2121; Copyright Massachusetts Medical Society.

central nervous system, as well as haematological and immunological manifestations.[30] In 1997, the criteria were revised due to certain sub-criteria, such as the analysis of Lupus Erythematosus (LE) cells and the Wasserman test for Syphilis, falling out of use. Instead, a positive Lupus anticoagulant (LA) test and the presence of immunoglobulin (Ig) G or M anti-cardiolipin (aCL) antibodies were added to the immunological criterion.[31] In both sets

of classification criteria, 4 of the 11 main criteria must be fulfilled, simultaneously or successively, and each criteria must not be better explained by other causes.[30]

Table 1. The 1982 American college of Rheumatology (ACR) classification criteria.

Malar rash	Fixed erythema over the malar eminences, often sparing the nasolabial folds
Discoid rash	Erythematous raised patches with adherent keratotic scaling and follicular plugging; atrophic scarring may occur in older lesions
Photosensitivity	Skin rash due to unusual reaction to sunlight, by patient history or physician observation
Oral ulcers	Oral or nasopharyngeal ulceration, usually painless, observed by physician
Arthritis	Nonerosive arthritis involving >2 peripheral joints, characterized by tenderness, swelling, or effusion
Serositis	a) Pleuritis-convincing history of pleuritic pain or rubbing heard by a physician or evidence of pleural effusion OR b) Pericarditis documented by ECG or rub or evidence of pericardial effusion
Renal disorder	a) Persistent proteinuria >0.5 grams/day or >3+ if quantitation not performed OR b) Cellular casts (red cell, haemoglobin, granular, tubular, or mixed)
Neurologic disorder	a) Seizures OR b) Psychosis; both in the absence of offending drugs or known metabolic derangements
Hematologic disorder	a) AIHA OR b) Leukopenia (<4,000/mm^3 total on ≥2 occasions) OR c) Lymphopenia (<1,500/mm^3 on ≥2 occasions) OR d) Thrombocytopenia (<100,000/mm^3 in the absence of offending drugs)
Immunologic disorder	a) Positive LE cell preparation OR b) Anti-dsDNA in abnormal titre OR c) Presence of anti-Sm OR d) False positive serologic test for syphilis for ≥6 months and confirmed by Treponema pallidum immobilization or fluorescent treponemal antibody absorption test
ANA	Abnormal ANA titre ever, by immunofluorescence or equivalent assays, in absence of drug-induced Lupus

ECG: electrocardiogram. AIHA: Autoimmune haemolytic anaemia. LE: Lupus Erythematosus. dsDNA: double-stranded DNA. Sm: Smith antigen. ANA: Antinuclear antibody. For the purpose of identifying patients in clinical studies, a person shall be said to have SLE if any 4 or more of the 11 criteria are present, serially or simultaneously, during any interval of observation. Table adapted from Tan *et al*. *Arthritis Rheum* (1982).[30]

The Systemic Lupus International Collaborating Clinics classification criteria

Since the publication of the ACR-82 criteria, critique has been raised for its large emphasis of the skin domain, with 4 of the 11 criteria involving skin

manifestations. In cases where all of these criteria are fulfilled, the ACR-82 criteria have been suggested to over-estimate the diagnosis, since neither involvement of a second organ system, nor the presence of ANA, would be required for an individual to be classified as having SLE.[27, 32] For this reason, the 2012 Systemic Lupus International Collaborating Clinics (SLICC-2012) classification criteria were introduced, where the skin criteria were refined whilst the total weight of the skin domain was decreased (table 2). In addition, criteria definitions were changed and new neuropsychiatric manifestations and hypocomplementemia were added.[32, 33] To be classified as having SLE according to the SLICC-2012 criteria, an individual must fulfil at least 4 criteria, including at least 1 clinical and 1 immunologic criterion (table 2). Due to its high specificity for SLE, histology-proven nephritis compatible with SLE in the presence of anti-dsDNA or ANA, is also considered sufficient for classification (table 2).[32] Compared with the ACR criteria, the SLICC-2012 criteria has increased sensitivity, however the specificity is lower than for the 1997 ACR criteria.[34, 35]

In 2019, a new set of classification criteria was introduced by the European League Against Rheumatism (EULAR) and the ACR, with ANA considered an entry criterion and fever added as an additional criterion. Furthermore, individual weights were assigned to each criterion, reducing the contribution of manifestations like oral ulcers and non-scarring alopecia, to the overall score. Conversely, the weight assigned to other manifestations, such as biopsy-verified lupus nephritis (LN) or joint involvement, have been increased, meaning that fewer of these criteria are required for diagnosis.[36] The EULAR/ACR classification criteria have been demonstrated to have a similar specificity as the ACR-97 classification criteria, whereas the sensitivity is higher than for the ACR-97 and similar to the SLICC-2012 classification criteria.[36]

Serological differences and subphenotypes of SLE

Since only a small number of each of the sets of classification criteria for SLE require fulfilment in order for an individual to be classified as having SLE, the classification of the disease allows for a large number of combinations of manifestations and the possibility of any 2 individuals with SLE fulfilling entirely different criteria for the disease. Whilst several attempts have been made to categorize patients based on clinical or serological findings, a neat categorization of patients into distinct subphenotype groups has proven difficult.[37, 38] There are, however, several serological findings which are associated with clinical manifestations on a group level:

Around half of patients with SLE carry anti–Sjogren's-syndrome-related antigen A/B (SSA/SSB) auto-antibodies. These patients more often have a diagnosis of secondary Sjogren’s syndrome (SS) in addition to SLE. Patients with

SSA antibodies are also typically older at disease onset and more often display skin manifestations compared with SSA/SSB negative individuals.[37, 39]

Table 2. SLICC-2012 criteria

Clinical criteria	
Acute cutaneous lupus	Lupus malar rash, bullous lupus, toxic epidermal necrolysis, maculopapular rash, photosensitive rash or subacute cutaneous lupus
Chronic cutaneous lupus	Localized or generalized discoid rash, hypertrophic lupus, lupus panniculitis, mucosal lupus, lupus erythematosus tumidus, chilblains lupus or discoid lupus/lichen planus overlap
Oral ulcers	Palate, buccal, tongue, or nasal ulcers in absence of other causes.
Nonscarring alopecia	In absence of other causes
Synovitis	Involving ≥2 joints, characterized by swelling/effusion OR tenderness and ≥30 minutes of morning stiffness
Serositis	Pleurisy >1 day, pleural effusions/rub; typical pericardial pain >1 day, pericardial effusion/rub, or pericarditis by ECG
Renal	Urine protein/creatinine (or 24 hr urine protein) representing 500 mg of protein/24 hr, or RBC casts
Neurologic	Seizures, psychosis, mononeuritis multiplex, myelitis, peripheral/cranial neuropathy, acute confusional state.
AIHA	In absence of other causes
Leukopenia or lymphopenia	<4,000/mm³ (leukopenia) or <1,000/mm³ (lymphopenia) at least once.
Thrombocytopenia	<100,000/mm³ at least once.
Immunological criteria	
ANA	Above laboratory reference range
Anti-dsDNA	Above (ELISA: 2x above) laboratory reference range
Anti-Sm	
Antiphospholipid antibody	Lupus anticoagulant, false positive RPR, medium or high titre aCL/ aβ_2GP1 (IgA/IgG/IgM)
Low complement	Low C3, low C4, low CH50
Positive direct Coombs test	In the absence of AIHA
Lupus nephritis	Biopsy-proven nephritis compatible with SLE *and* ANA or anti-dsDNA

ECG: electrocardiogram. RBC: red blood cell. AIHA: autoimmune haemolytic anaemia. dsDNA: double-stranded DNA. ELISA: Enzyme-linked immunosorbent assay. Sm: Smith. RPR: Rapid plasma regain. aCL: anti-cardiolipin. aβ_2GP1: anti-β_2-glycoprotein 1. CH50: 50% haemolytic complement. At least 1 clinical, at least 1 immunologic and 4 total, OR both lupus nephritis criteria, are required for an individual to be classified as having SLE. Table adapted from Petri *et al*. *Arthritis Rheum* (2012).[33]

Around a third of patients with SLE carry anti-phospholipid antibodies (aPL), including aCL or anti-β_2-glycoprotein 1 (aβ_2GP1) antibodies as well as a positive LA test. These patients more often also fulfil the classification criteria for antiphospholipid syndrome (APS),[40, 41] which is defined by arterial events (AE), venous thromboembolism (VTE) or obstetric complications such as miscarriages or premature birth, in the presence of aPL. Patients positive for aPL also more often fulfil the hematologic criterion due to the development of autoimmune haematolytic anaemia (AIHA) or thrombocytopenia. In addition, they are more likely to be affected by neurologic manifestations, renal microangiopathy, alveolar haemorrhage, myelitis and adrenal insufficiency.[41] Consequently, the overall risk of organ damage and premature death, is higher in patients with secondary APS diagnosis.[42, 43]

Around 2/3 of patients with SLE are positive for anti-dsDNA. These antibodies are strongly associated with the development of nephritis, and their titres correlate positively with disease activity.[44]

Mortality, prognosis and complications

In the last 60 years, the mortality rate of SLE has decreased, likely due to both new treatment strategies and to factors such as an improved recognition of milder cases.[45] Whilst renal involvement historically has been a major reason for premature death in SLE, changes in strategies for monitoring and treatment of patients with LN have led to a remarkable improvement in prognosis. The mortality rate in SLE, however, still remains 2- to 3-fold higher than in the general population, and this figure has been relatively stable since the 1990s, with cardiovascular disease (CVD) and infections now representing the major causes of death.[45-49]

The mortality pattern in SLE is bimodal, with early mortality found to be associated with disease activity and infections, and late mortality with CVD.[50] The mortality risk in SLE has been demonstrated to be higher in Asia, Latin America and North America compared with in Europe.[51] This may partly reflect differences in socioeconomic factors, with non-Caucasian groups having a poorer access to health care compared with Caucasians. Non-Caucasian ethnicities are also known to accrue damage earlier in the disease course and to have a higher risk of nephritis. In addition, the onset of the disease has been demonstrated to occur earlier in life in non-Caucasian groups - a factor which may reflect a higher genetic load.[10, 13, 52]

Permanent organ damage

As a direct consequence of the inflammation associated with active disease, or indirect disease-related factors such as prolonged corticosteroid use, the majority of patients eventually develop permanent organ damage. To monitor the prognosis of patients and assess the impact of SLE, the validated SLICC/ACR Damage Index (SDI) is often used (table 3).[53] This instrument records permanent organ damages, as defined by at least 6 months of disease duration, from 12 different organ domains. The SDI does not only demonstrate the disease progress so far, but also indicates a prognosis for the individual patient as a higher score on the SDI is predictive of further damage.[54] Importantly, unlike the classification criteria for SLE, organ damage thought to occur for other reasons than SLE itself should nevertheless be recorded on the SDI.[53] As the risk of many included items, such as osteoporosis, diabetes mellitus or malignancy, is high in the elderly population in general, the SDI does not only reflect co-morbidity related to SLE, but also to age.

Table 3. The systemic lupus collaborating clinics/American college of rheumatology (SLICC/ACR) Damage Index

Ocular	Any cataract; retinal change OR optic atrophy
Neuropsychiatric	Cognitive impairment OR major psychosis; Seizures requiring therapy for > 6 months; Cerebral vascular accident OR resection not for malignancy; cranial or peripheral neuropathy; transverse myelitis
Renal	GFR<50%; Proteinuria >3.5 g/24 h; ESRD
Pulmonary	Pulmonary hypertension; pulmonary fibrosis; Shrinking lung; Pleural fibrosis; pulmonary infarction OR resection not for malignancy
Cardiovascular	Angina OR coronary bypass; Myocardial infarction; Cardiomyopathy; Valvular disease; Pericarditis OR pericardiectomy
Peripheral vascular	Claudication for >6 months; minor tissue loss; significant tissue loss ever; Venous thrombosis with swelling, ulceration OR venous stasis
Gastrointestinal	Infarction or resection of bowel below duodenum, spleen, liver or gall bladder; mesenteric insufficiency; chronic peritonitis; stricture OR upper gastrointestinal tract surgery ever; Pancreatic insufficiency requiring enzyme replacement OR with pseudocyst
Musculoskeletal	Muscle atrophy OR weakness; Deforming or erosive arthritis; Osteoporosis with fracture or vertebral collapse; Avascular necrosis; Osteomyelitis;
Skin	Scarring chronic alopecia; extensive scarring of panniculum other than scalp and pulp space; Skin ulceration (excluding thrombosis).
Gonadal	Premature gonadal failure
Endocrine	Diabetes requiring therapy, regardless of treatment
Malignancy	Malignancy (exclude dysplasia)

GFR: glomerular filtration rate. ESRD: End-stage renal disease. Table adapted from Gladman *et al. Arthritis Rheum*. (1996)[53]

Cardiovascular disease

CVD is a set of heterogenous heart and circulatory system disorders which are responsible for around half of all deaths in Europe.[55] CVD is more common in individuals with SLE compared with the general population, with strong evidence of a 2- to 3-fold increase in the RR of myocardial infarction (MI), congestive heart failure, cerebrovascular disease, and overall CVD mortality.[56, 57] Whilst the absolute risk of CVD increases with age both in individuals with SLE and in the general population, the RR in SLE compared with the general population is the highest in younger women and in the first few years following SLE onset.[58, 59]

The traditional Framingham risk factors, which explain a substantial amount of CVD in the general population and include important risk factors like smoking and hypertension, are important also in patients with SLE, but cannot fully account for the excess risk.[60, 61] SLE-specific risk factors like high disease activity, renal disease and prolonged corticosteroid use, as well as genetic risk factors, have been suggested to contribute to the excess risk.[56, 62] Several mechanisms important in SLE have further been suggested to contribute to endothelial dysfunction, which is thought to contribute to the development of CVD. This includes enhanced type I IFN activity − which has been shown to impair endothelial function even in patients with low disease activity − and the presence of autoantibodies and immune complexes, neutrophil responses and abnormal lipoprotein function.[62, 63]

Renal involvement

LN affects between 30-80% of patients with SLE, depending on the basic characteristics of the sample studied, with the prevalence in Sweden estimated to be 38%.[64] In general, a higher prevalence can be observed amongst men, individuals of non-European ancestry and individuals with childhood-onset SLE (cSLE).[52, 65, 66] Whilst improvements in treatment of LN have historically resulted in increasing survival rates in developed countries, this reduction in mortality reached a plateau in the 1990s. Presently, the mortality rate of SLE with LN remains 2-fold higher compared with that of patients without LN. The reasons for this increased risk are several, with LN being associated with a higher risk of accumulating organ damage and developing CVD. Furthermore, in 5-30% of LN patients the nephritis progresses to end-stage renal disease (ESRD).[65, 67, 68] Due to these risks associated with the renal inflammation, early diagnosis and adequate treatment of LN in patients with SLE is crucial for preventing permanent renal damage.

In the vast majority of cases, LN is either present at the time of SLE diagnosis, or has its onset within the first 5 years of disease, with the highest incidence

in the first year.[67, 69] Since the early clinical manifestations of the condition can be mild or absent, regular screening for proteinuria, haematuria and hypertension is important for early detection, especially during the first years after SLE diagnosis.[67, 69] The gold standard for diagnosis of LN is by renal biopsy. As the histological picture of LN varies, and the prognosis is worse in some cases than others, the treatment modality is often adjusted according to the degree of nephritis. Therefore, a renal biopsy and subsequent classification is of importance not just for diagnosis, but for guidance on treatment, and should be performed upon clinical suspicion of LN. The 1995 World Health Organization (WHO) revised classification criteria (table 4) or the 2003 International Society of Nephrology/Renal Pathology society (ISN/RPS) criteria are often used for classification of LN.[70, 71] Proliferative nephritis (WHO or ISN/RPS class III-IV) is associated with the worst prognosis and a higher risk of developing ESRD, and requires a more intense immunosuppressive therapy in combination with antihypertensive treatment.[71]

Table 4. World Health Organization (WHO) histological classification of lupus nephritis (1995)

Class I	Normal glomeruli
	Normal by all techniques
	Normal by light microscopy, but deposits by electron or immunofluorescence microscopy
Class II	Pure mesangial alterations
	Mesangial widening and/or mild hypercellularity
	Moderate hypercellularity
Class III	Focal segmental glomerulonephritis (associated with mild or moderate mesangial alterations)
	With “active” necrotizing lesions
	With “active” and sclerosing lesions
	With sclerosing lesions
Class IV	Diffuse glomerulonephritis (severe mesangial, endocapillary, or mesangiocapillary proliferation and/or extensive subendothelial deposits)
	Without segmental lesions
	With “active” necrotizing lesions
	With “active” and sclerosing lesions
	With sclerosing lesions
Class V	Diffuse membranous glomerulonephritis
	Pure membranous glomerulonephritis
	Associated with lesions of class II
	Associated with lesions of class III
	Associated with lesions of class IV
Class VI	Advanced sclerosing glomerulonephritis

Table adapted from Chow *et al*. *Malays J Pathol*. (2015)[71]

Genetics

Genetics is the scientific study of genetic variation, and of heredity – how the variation is transferred between individuals, resulting in both similarities between family members and diversity between, and within, species. Whilst humans have shown both curiosity and a certain understanding for the concept of heredity for thousands of years, as demonstrated by ancient practices of animal breeding and cultivation of crops, the scientific comprehension of its mechanisms is only just over a century old. The laws of monogenetic inheritance, originally described by Gregor Mendel but subsequently fallen into oblivion for decades, were re-discovered around the turn of the last century, amid microscopy studies identifying a mysterious nuclear structure which appeared to be transmitted from mother to daughter cell during cell division. Eventually, these research paths would converge, with the realization that the nuclear substance carried the very blueprint of both life and disease: the gene.

The human genome and genetic variation

Our cells contain complete instructions for generating a human being, with the genes, consisting of DNA, encoding each of our proteins. The DNA is packed into 23 chromosome pairs in most individuals, each consisting of 1 chromosome derived from each parent, and genes thus occur in 2 copies. The DNA molecule is constructed from deoxyribose, a phosphate group and nitrogenous bases which come in 4 different versions, represented by the first letter of their names: C (cytosine), G (guanine), A (adenine) and T (thymine). The human genome consists of 3 billion nucleotide pairs, with 2 bases consistently bound together; C − G, and A − T. The nuclear DNA contains around 20,000 protein-coding genes, in which the sequence of these bases determines the sequence of amino acids used in the protein construction. In addition, 22,000 genes in the human DNA encode RNA molecules involved in various cellular activities, including protein translation and regulation of gene expression.[72]

All genetic variation originates from changes in the DNA sequence. As a result of these changes, the DNA sequence can differ between individuals and, when present in the germline cells, the change can be transmitted from parent to offspring. The differing sequences, constituting only 0.1% of the human

genome and yet responsible for all genetic variation in our species, are known as alleles. When 2 alleles at the same locus are both common in the population, they are referred to as polymorphisms, with single nucleotide polymorphisms (SNPs) being the most common variation. As each chromosome occurs in 2 copies, for each bi-allelic SNP it is possible to be homozygotic, i.e. having 2 identical copies, for either the minor or major allele, or to be heterozygotic and carry 1 copy of each allele.[72] Changes in the DNA sequence are often functionally silent, with no consequence for the affected individual. In some cases, however, a single point mutation may result in anything from a small change of the function, to a complete loss-of-function of the expressed protein. Moreover, in non-protein-coding segments of the DNA, variations can affect the regulation of a gene's expression, leading to alterations in the amount of the protein produced.[72]

Monogenetic and multifactorial traits

Monogenetic disorders, also referred to as Mendelian disorders, are estimated to affect 1% of the population and are associated with prominent and consistent familial inheritance patterns. Whilst more than 6000 known monogenetic phenotypes are currently known, the prevalence of each individual monogenetic disease is typically very low. For reasons that are not yet fully understood, the same monogenetic mutation can cause different degrees of disease in different individuals, even in individuals within the same family. These phenomena are known as incomplete penetrance or variable expressivity.[73]

Whilst the study of monogenetic disorders has been crucial for our understanding of genetics and heredity, the majority of common diseases have a multifactorial aetiology, involving both multiple genes and environmental factors. Although these diseases are typically associated with a certain increase in the risk of recurrence in siblings or close relatives of an affected individual, they lack the distinct family pattern of inheritance observed in monogenetic diseases.[73, 74] Furthermore, some features can be observed in these diseases, providing evidence of polygenic inheritance and differentiating them from monogenetic diseases with reduced penetrance. For example, the risk of family recurrence of a multifactorial disorder is typically higher when multiple family members are affected, if the disease severity in the affected proband is greater than average, or if the affected proband is of the less commonly affected sex.[74]

The heritability of multifactorial traits

Whilst the environmental and genetic components responsible for causing multifactorial diseases can be hard to separate from each other – with a widely-accepted view that disease is a result of certain environmental factors triggering the disease in genetically susceptible individuals – family studies can be used to calculate the genetic and environmental components of a particular disease. In twin studies, the disease concordance between monozygotic (MZ) and dizygotic (DZ) twins, can be compared, provided that the amount of shared environmental influence is the same in both types of twins. This assumption, however, has been shown not to always be accurate, and to fully account for these differences, adoption studies looking at twins raised apart from each other, can be of large value.[75] An obvious limitation of these studies is the small number of MZ twins available, and unusual situation of twins being separated early in life. For that reason, family studies investigating disease concordance between full siblings or parent-child units are more commonly used, especially in diseases with a low population prevalence.[74, 75]

The genetics of multifactorial traits

Whilst twin- and family studies can provide information about heritability, other methods are required to identify the specific disease-causing gene. Before 2005, there were two main methods used to identify the variants associated with genetic diseases. Firstly, in linkage studies, genomic regions containing genes which predispose to disease were identified by studying members of families with a history of a particular disease. The method relies on linkage between genomic regions, and the method allows the identification of regions responsible for causing the disease, even in the absence of hypothesis about which genes may be involved. These studies are often used in the first stage in the genetic investigation of a disease, and are primarily useful for identifying monogenetic traits.[76] Secondly, in candidate gene studies, a small number of genetic variants were compared between individuals with and without a particular trait or disease. The variants explored were limited and hand-picked by the investigators due to an a priori hypothesis, based on knowledge about the involvement of the gene in a pathway, or based on the gene being located in a particular chromosomal region, thought to be involved in the trait of interest. This approach conferred some important limitations, including that the choice of candidate genes may be based on an erroneous hypothesis, that the causative genes may be located upstream or downstream of the candidate gene in the signalling pathway, and that genes associated with relevant but unknown mechanisms in the disease will not be examined.[77, 78]

In the last 20 years, the cost of genotyping has decreased dramatically, enabling comparisons of hundreds of thousands of genetic loci across the genome, in thousands of individuals, in so-called genome-wide association studies (GWASs). Typically, the genetic variants used are SNPs. In spite of this large number of SNPs included, only a fraction of all our SNPs are normally compared in a GWAS. This is possible due to linkage disequilibrium (LD), a phenomenon causing certain variants to be inherited together more often than would be expected by chance. Due to LD across regions in our DNA, the exact variant causing a disease does not have to be measured in a GWAS, as SNPs inherited together with that variant will cause a signal.[79]

Since the publication of the first GWAS in 2005, these studies have led to the identification of more than 250,000 genetic variants associated with traits and diseases at genome-wide significance ($p<5\times10^{-8}$).[79, 80] The majority of these genetic variants are common variants, defined as having a minor allele frequency (MAF) > 5%. These results have had an extensive impact on our understanding of the aetiology of several diseases, through the identification of disease-causing mechanisms.[81-85] Furthermore, they have aided the discovery of novel drug targets, with clinical trials supported by GWAS findings showing a doubled success rate.[86, 87]

Assessment of polygenic risk

Before the GWAS era, a common belief was that polygenic diseases were influenced by a relatively small number of genetic variants, meaning that many disease-causing genetic variants would have a significant impact on the disease risk, making them both relatively easy to identify in a GWAS and useful in predicting the occurrence of a disease in an individual. However, with thousands of GWASs now performed, evidence has emerged that very few variants confer a risk of disease larger than an odds ratio (OR) of 1.5, with most associations displaying even smaller ORs.[81-83, 85] With this realization follows, that to explain the whole genetic heritability of a typical, multifactorial disease, thousands of genes may have to be identified. For this reason, polygenic models that consider the cumulative risk of anything from a few, to thousands, of genetic variants, are gaining increasing interest.

The construction of a genetic risk score (GRS) is a method of estimating an individual's cumulative genetic risk of a particular phenotype. When the genetic factors included in a GRS are based on genome-wide markers, the score is often referred to as a polygenic risk score (PRS).[88] These scores can be calculated in several ways. A simple method is to count the number of risk alleles for a selected number of risk loci in an individual, and to let the total sum of these risk alleles be defined as the GRS. More commonly, however, each locus is also weighted by multiplying the number of risk alleles by the

previously reported effect size (or sometimes the natural logarithm of the effect size) for that allele. All weighted alleles are subsequently summarized.[89] The equation of this calculation for an individual ***j*** can be written as:

$$PRS_j = \sum_i^N \beta_i * dosage_{ij}$$

The performance of a GRS can be evaluated in a number of ways. For example, the pseudo-R^2, providing an estimate of the phenotypic variance explained by the GRS, may be used. Another commonly used method is to calculate the area under the receiver operating characteristics (ROC) curve (AUC), which compares rates of true positives (sensitivity) and false positives (1 – specificity) and provides an aggregate measure of performance across all possible classification thresholds. The AUC can have a value ranging from 0 to 1, where a value of 0.5 indicates a performance equal to coin-flipping and a value of 1 indicates a perfect separation of positive and negative cases.[84, 85, 89]

GRSs have been found to serve some usefulness in several common diseases. In type 2 diabetes mellitus (T2D), studies have demonstrated up to 9-fold odds of developing the condition for individuals with a GRS in the top percentiles compared to the individuals with the lowest scores, and the American Diabetes Association risk tool has been shown to generate a significantly higher predictive accuracy when including a GRS in the model.[90-92] Similarly, GRSs for coronary artery disease (CAD) have demonstrated the ability of identifying groups with a strongly increased risk of both CAD and subclinical atherosclerosis, with effect sizes comparable or larger than those of traditional risk factors. In addition, numerous studies have found that the predictive ability of the American college of cardiology atherosclerotic CVD risk calculator, improves significantly with the addition of CAD GRSs.[90, 91, 93, 94] For obesity, GRSs have demonstrated a difference in average weight of 13 kilograms, as well as a 25-fold change in the risk of severe obesity, when comparing individuals the top and bottom deciles.[91] Similar results have been found for a range of other diseases, including VTE, atrial fibrillation, breast cancer and type 1 diabetes (T1D), with AUCs for several of tested GRSs approaching 0.75 – a figure that has been suggested as a minimum for informative screening of individuals with an increased risk of a disease.[85, 95, 96]

When constructing GRSs, there are some important technical factors to consider. For example, any sample overlap between the cohort used for discovery of genetic variants and the cohort where the GRS is tested, can inflate the association between the GRS and the measured phenotype. In addition, population differences between the discovery and target cohorts may affect the

measured performance of the GRS, with previous studies demonstrating a better performance when the scores are tested in the same ancestry for which they were derived.[97, 98]

Etiopathogenesis of SLE

The strong genetic component of SLE is evident from studies of twins and families. Familial aggregation studies have estimated the sibling risk ratio (lambda-s), defined as the increased RR for a sibling of an individual with SLE compared to a member of the general population, to between 8 and 29. MZ twins are approximately 10-fold more likely to be concordant for the disease compared to DZ twins or siblings (20–40% compared with 2–5%).[99, 100] This knowledge has been used in attempts to quantify the genetic component of SLE. These estimations, whilst difficult to carry out with precise results, suggest that the genetic component constitutes 44% to 66% of the total SLE risk. The environmental component of the SLE risk can be further divided into shared- and non-shared environments, each accounting for approximately half of the non-genetic risk of SLE.[101]

Genetic factors

In rare cases, SLE has been observed to run in families with an inheritance pattern corresponding to a monogenetic disease. Those affected often present with cSLE and typically display a more aggressive disease phenotype. Interestingly, the female predominance normally observed in the disease, is generally absent in these cases.[102] With the exception of these rare monogenetic instances, SLE is highly polygenic, with each genetic locus conferring only a small increase in the risk of developing the disease.

To date, over 20 GWASs, large-scale replication studies or meta-analyses have been conducted on SLE, mainly involving individuals of European or Asian ancestry.[13, 103-123] Through these studies, more than 180 independent genetic regions associated with the disease have been identified (some of which are illustrated in figure 2) and combined, these regions have been estimated to explain 30% of the disease risk.[13, 101, 124] Consequently, a discrepancy remains between the estimated heritability of SLE and the identified genetic aetiology. The exact reasons for these comprehensive studies' inability to identify the remaining genetic risk factors remain unclear, however, this "missing heritability" problem has been observed in virtually all common diseases. Possible explanations include over-estimations of the heredity of SLE,

imperfect coverage of GWAS arrays, interactions between certain genetic factors, or large contributions from risk variants which, due to rarity or small effect sizes, require higher statistical power than studies have achieved thus far.[79]

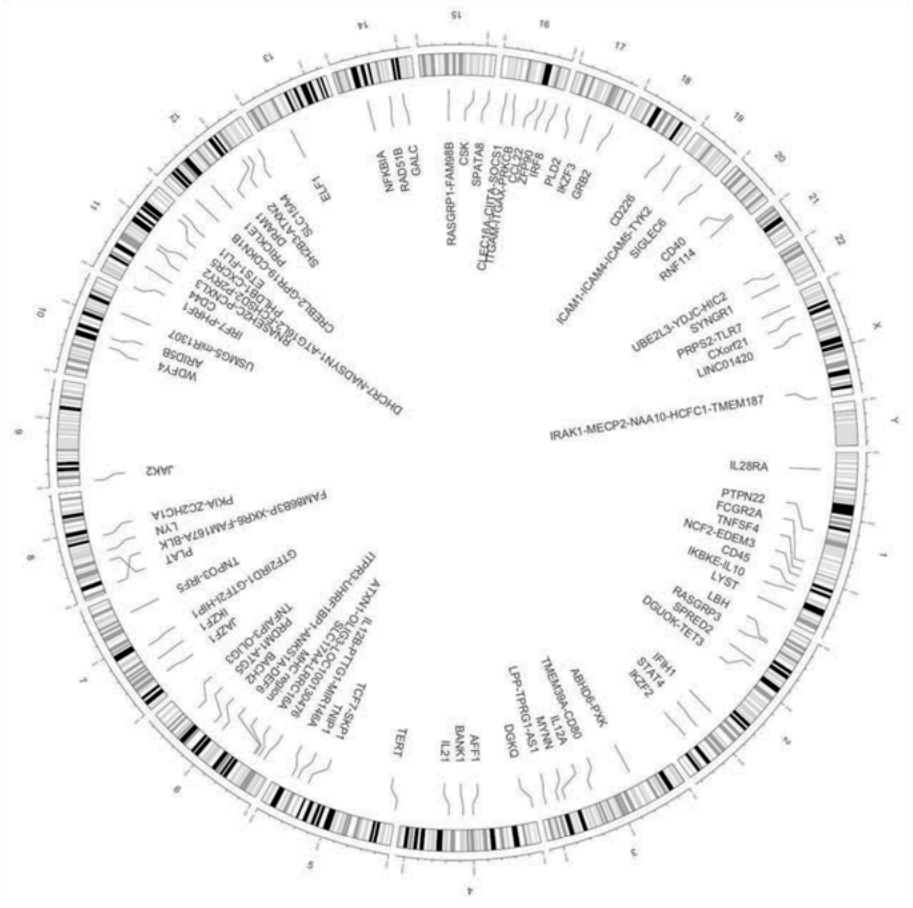

Figure 2. Genetic loci associated with SLE development. The figure shows genes located within 84 genetic loci with established SLE association. The 23 chromosomes are represented by the outside numbers. Reproduced with permission from Chen et al. Genetic advance

The strongest genetic associations with SLE, giving rise to a 2-3-fold increase in disease risk, have been found in the human leucocyte antigen (HLA) region. This region, also known as the major histocompatibility complex region, is located on chromosome 6, is highly polymorphic and, due to high LD across the region, entails difficulties in pinpointing the genetic variants causally associate with SLE development.[110] Whereas the DRB1*03:01 and

DRB1*15:01 HLA alleles have displayed association with SLE in multi-ancestry analyses, other HLA alleles may be specific to certain ancestries due to variations in LD patterns between different ethnic groups.[110]

Most genetic variants associated with SLE are located outside of the HLA region, and their effect sizes are usually smaller, with ORs for SLE development often ranging between 1.1-1.5.[101]

The role of SLE risk genes in SLE pathophysiology

Of the approximately 180 independent SLE risk variants discovered in GWASs so far, many have provided opportunities to confirm, or more deeply understand, the pathophysiology of SLE. Of the SNPs with a confirmed or probable function identified, many have been found to have a role in B- and T-cell signalling, apoptosis, immune complex clearance and type I IFN signalling.[125] However, linking the genetic variants identified in GWASs to functional consequences can be challenging for several reasons. Many of the identified SLE risk loci are located outside of the coding regions of the DNA, and thus do not directly affect protein transcription, but likely possess a regulatory role. The difficulty in distinguishing the causative variants from those in LD with these variants, combined with the rapid pace of new genetic variant identification, leaves the role of some identified variants on SLE pathogenesis unknown.[101]

Cell death and immune complex clearance

A major contributor to the pathology of SLE, is the imbalance between the production and clearance of cellular debris. Defects in cell death mechanisms, including apoptosis and a neutrophil-specific form of death involving neutrophil extracellular traps (NETs), termed NETosis, expose the immune system to autoantigens and damage-associated molecular patterns. Under normal circumstances, apoptotic debris is rapidly cleared following apoptosis, facilitated by opsonization by complement components.[126] In SLE, however, an increased production of apoptotic material and decreased levels of classical complement components, result in sustained levels of debris for long enough to trigger autoimmune and inflammatory responses.[127, 128]

The monogenetic forms of SLE clearly demonstrate the importance of the complement system for disease development, as they include several forms of deficiencies in complement components 1q (C1q), 2 and 4. Among these, C1q deficiency entails the strongest absolute risk for SLE, exceeding 90%.[102, 129] Additionally, genes involved in immune complex clearance and apoptosis also play an important role in polygenic SLE (figure 3).[130] In addition to impairments in apoptosis and NETosis, there is evidence of other forms of cell deaths, including necroptosis and pyroptosis, to be implicated in the SLE

pathogenesis. The role of autophagy, a process for digesting and recycling unnecessary or dysfunctional cellular organelles, is less explored. However, GWASs have identified several autophagy-related genes, including Autophagy Related 5 (*ATG5*) and C-Type Lectin Domain Containing 16A (*CLEC16A)*, as associated with SLE.[127]

Auto-antibodies and cellular immunity

Due to disrupted central B cell tolerance mechanisms, SLE patients have an increased number of hyperreactive and autoreactive B cells.[131] Although the exact mechanisms of the loss of tolerance have not yet been fully elucidated, gene variants in B cell signalling pathways, including BLK Proto-Oncogene (*BLK*) and B Cell Scaffold Protein With Ankyrin Repeats 1 (*BANK1*), have been shown to contribute to this hyperactivity (figure 3).[107, 109] The auto-reactive B cells are responsible for a sustained production of a profuse number of ANAs, which primarily target nucleic acid or related proteins. However, B cells have multiple additional roles besides autoantibody production, including antigen presentation and cytokine secretion. Furthermore, a particular subset of cells, the B regulatory cells, serve to maintain immune homeostasis, and an impairment in this cell type has been pointed out as important in the development of SLE. B regulatory cells largely achieve their effect through the production of inhibitory cytokines like Interleukin 10 (IL-10), and SNPs associated with IL-10 have found to contribute to SLE development.[131]

In order for B cells to function, costimulatory signals from cluster of differentiation 4 positive (CD4+) helper T cells are required. CD4+ T cells are thus crucial for the development of the autoantibody response, and activated CD4+ cells have further been identified in kidney and skin tissue and are thought to directly contribute to damage in these organs. Simultaneously, both the balance between, and function of, certain subsets of CD4+ cells is altered in SLE, with T helper 17 cells contributing both to SLE pathogenesis and the development of LN whilst T regulatory cells are found in lower numbers and fail to limit inflammation.[132, 133] Another CD4+ cell subset, T follicular helper cells, are necessary for an optimal antibody response, and uncontrolled expansion of these cells have been found in SLE. Several of these alterations have been found to be particularly pronounced in active SLE. In addition to the CD4+ cell subsets, cluster of differentiation 8 positive (CD8+) cytotoxic T cells demonstrate reduced activity in SLE, leading to both an increase in infection susceptibility and to sustained autoimmunity. Self-reactive CD8+ cells have also been shown to lose their CD8 positivity, and thereby their CD8-dependent immunosuppressive potential, and convert to double-negative T cells, which are capable of infiltrating tissues.[132, 133]

Many genes identified as associated with SLE development have been found to be associated with important T cell pathways. This includes the Signal Transducer and Activator of Transcription (*STAT4*) gene, which is crucial for the differentiation of naïve CD4+ cells into Th1 cells, TNF Superfamily Member 4 (*TNFSF4*), which is involved in interactions between T cells and antigen-presenting cells (APCs), and Interleukin-10 (*IL10*) (figure 3).

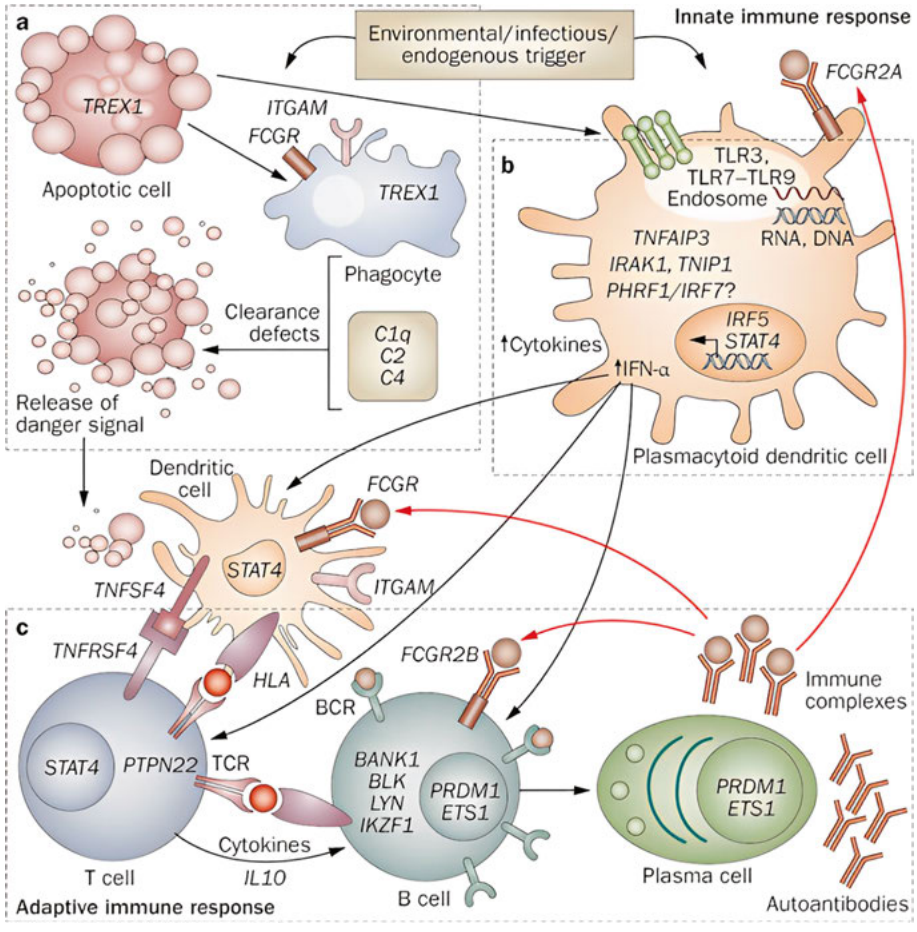

Figure 3. The role of genetics in SLE pathogenesis. Reproduced with permission from Deng et al. Genetic susceptibility to systemic lupus erythematosus in the genomic era. Nat Rev Rheumatol 6:683–692 (2010). Copyright Springer Nature.

The type I interferon system

Due to increased apoptosis, decreased clearance of apoptotic material and production of auto-antibodies, nucleic acid-containing immune complexes are formed in SLE. These immune complexes, together with the neutrophil extracellular traps (NETs) formed during NETosis, trigger type I IFN production in plasmacytoid dendritic cells. This production, in turn, affects numerous

components and events in the immune system, including maturation of dendritic cells into APCs as well as activation, differentiation and increased survival of auto-reactive B and T cells.[125, 134] Most patients with SLE exhibit an upregulation of a large number of type I IFN genes, known as the type I IFN-signature, which, along with evidence of treatment with type I IFN leading to an SLE-like phenotype in certain individuals and of success in the newly approved type I IFN-receptor-blocking antibody anifrolumab, provides strong evidence of a causal role of type I IFN in SLE pathogenesis. Through the events triggered by type I IFN, immune tolerance is disrupted, promoting perpetual autoimmunity.[125, 134-137] Not unexpectedly, GWASs have shown a large over-representation of genes involved in type I IFN signalling, production and response, including Interferon regulatory factor 5 (*IRF5*) and *STAT4* (figure 3).[138, 139]

Associations between SLE genetic variants and other diseases

A number of SLE-associated genetic variants are known from large GWASs to be associated also with other diseases. An interesting example is the SH2B Adaptor Protein 3-Ataxin 2 (*SH2B3-ATXN2*) risk variant rs10774625, which has shown association with a wide range of phenotypes in addition to SLE, including several CVD variables, glomerular filtration rates (GFR), asthma and cancer.[140-144] Currently, the strongest association between this SNP and phenotypes reported in the GWAS catalogue, is for smoking initiation, which is in turn a risk factor for some of these diseases.[145] The pleiotropic effects of certain other SLE risk SNPs appear limited to other autoimmune diseases. A famous example of this is that variants in the *Signal Transducer and Activator of Transcription 4* (*STAT4)* gene, have displayed strong association with a wide range of auto-immune diseases in addition to SLE, including rheumatoid arthritis (RA), primary biliary cholangitis, juvenile idiopathic arthritis, SS, multiple sclerosis (MS) and ulcerative colitis.[146-151] Moreover, in certain GWASs of traits resembling manifestations observed in SLE, a notable number of SLE-associated genes have come up amongst the strongest observations. Examples include a large GWAS and meta-analysis of oral ulcers, that demonstrated strong associations with several well-established SLE risk loci including *STAT4*, Interleukin 12A (*Il12A*), *IRF5*, Interferon regulatory factor 8 (*IRF8*), Ikaros Family Zinc Finger 1 (*IKZF1*) and *IL10*.[152]

Associations between genetic variants and SLE manifestations

In addition to the observations made in the general population, several studies have attempted to analyse whether the individual SLE genetic variants are associated with specific manifestations of SLE, with many associations reported. For example, variants in *STAT4* have been reported to be associated with nephritis, severe renal insufficiency, earlier SLE onset, higher prevalence

of anti-dsDNA positivity, as well as the presence of aPL and ischemic stroke.[153-157] Other examples include associations of Integrin Subunit Alpha M (*ITGAM*) variants with discoid lupus, renal disease and immunological manifestations, an association between a variant in *IRF8* and CVD, as well as multiple variants in SLE- or non-SLE risk genes and LN.[158-161]

Many of the studies performed to date, however, carry some important limitations. For example, several associations reported from GWASs comparing patients with SLE with and without the manifestation of interest, only reach suggestive significance, whilst others are the results from comparisons of SLE patients with a particular phenotype and healthy controls. In the latter case, it may be unclear whether the associations are driven by the phenotype, or by the disease itself. In addition, many studies that employ a candidate-gene approach to examine associations between risk alleles and manifestations are restricted by the limitations of such studies.[158] For these reasons, the knowledge of how a genetic SLE predisposition affects the development of specific subphenotypes, remains relatively limited.

Environmental and hormonal risk factors

The moderate concordance rate in MZ twins clearly reveals that environmental factors also play an important role in the development of SLE. Due to the low incidence rate of SLE, however, prospective studies are difficult to conduct as they require following a very large number of individuals in order to achieve enough statistical power to detect associations. Instead, many studies aiming at investigating environmental factors in SLE use a retrospective case-control approach. Such studies often entail a risk of recall bias, or of measuring differences in laboratory parameters or epigenetic changes which may have occurred as a consequence of, rather than prior to, diagnosis. With few exceptions, the evidence of environmental factors suggested to cause SLE remains insufficient. In the case of Silica dust, population- and occupational based prospective studies have provided consistent evidence of an association.[162-164] There is also some support for other respiratory exposures such as asbestos and air pollution playing a part in SLE development.[164] Cigarette smoking has been evaluated as a risk factor for SLE in several studies, and may be linked in particular to anti-dsDNA positive SLE.[165-167] In addition to respiratory exposures, certain drugs can induce an SLE-like state, which typically is transient and ceases when the treatment is discontinued.[168] Ultraviolet radiation (UVR) is a trigger of apoptosis and can cause exacerbation of SLE, with a photosensitive rash being included as a classification criterion in both the ACR-82 and SLICC-2012 criteria. Whether it is also associated with onset of the disease, however, remains unclear.[169]

Other suggested risk factors for SLE include viral infections, in particular Epstein-Barr Virus infections, and vitamin D deficiency,[170, 171] although more prospective studies would be needed to confirm these associations.

Since SLE displays a clear female predominance, which is accentuated after puberty and attenuated after the onset of menopause, female sex hormones appear to play an important role in the disease development. This is further supported by studies investigating the role of oral contraceptives containing oestrogen, which seems to increase the risk of SLE flares.[172] In addition to differences in sex hormones, the female X chromosome contains several genes important for the immune system, and some of these have displayed association with SLE in GWASs.[173] In addition, impaired X-inactivation appears to be involved in the disease aetiology, since men with Klinefelter's syndrome carrying the XXY karyotype are at increased risk of SLE compared with XY males.[174]

Present investigation

Aims of the thesis

The general aim of this thesis is to enhance the understanding of how clinical and genetic risk factors contribute to the extreme heterogeneity of SLE manifestations, with a special focus on more severe disease phenotypes, the development of permanent organ damage, and mortality.

The specific aims of each paper were:

Paper I - to characterize the accumulation of organ damage and causes of death in patients with SLE, and to investigate clinical risk factors for the accumulation of organ damage.

Paper II - to assess the relationship between a high cumulative genetic risk of SLE and clinical manifestations associated with a more severe disease phenotype, including the accumulation of organ damage, cardiovascular events (CVE) and end-stage renal disease.

Paper III - to investigate the association between smoking and clinical manifestations of SLE, and to examine how the effect of smoking on these manifestations is modulated by the presence of genetic variants that contribute to the development of these manifestations.

Paper IV – to investigate the possibility of constructing genetic risk scores for manifestations of SLE, by considering associations between SLE risk SNPs and SLE manifestations in a large cohort of individuals without SLE.

Patients and methods

Patients and controls

In paper I, 543 patients with SLE from the University clinics of Linköping and Uppsala were included. In paper II and IV, we included 1001 patients with SLE from the University clinics in Uppsala, Linköping, Lund, Karolinska Institute (Stockholm) and from the 4 northern-most counties in Sweden. In addition, paper II included 5 524 individuals with SLE with European ancestry as a validation cohort,[175] and paper IV included 486 additional patients with SLE from the University clinics in Uppsala, Lund, Linköping, Gothenburg and Oslo. In paper III, 774 patients with SLE from Sweden and 836 patients from Norway and Denmark were included in the study. A complete list of the number of patients from each clinical centre included in each paper, is available in table 5.

Table 5. Number of patients included from each research centre in each paper

	Paper I	Paper II	Paper III	Paper IV
Uppsala University	296	172 (IC)	171 (IC)	172 (IC) + 10 (GSA)
Linköping University	247	165 (IC)	165 (IC)	165 (IC) + 12 (GSA)
Lund University	-	141 (IC)	141 (IC)	141 (IC) + 118 (GSA)
Karolinska Institute	-	299 (IC)	299 (IC)	299 (IC)
Northern Sweden	-	224 (IC)	-	224 (IC)
Gothenburg University	-	-	-	77 (GSA)
Oslo University	-	-	289 (CA)	269 (GSA)
Copenhagen University	-	-	175 (CA)	-
Odense University	-	-	220 (CA)	-
Aarhus University	-	-	152 (CA)	-
Non-Swedish Europeans [110]	-	5 524 (IC)	-	-

IC: Immunoarray chip. GSA: Global Screening Array. CA: Custom Array. The number of patients included in each of the papers, from each clinical centre. When genotyping of patients was performed, the genotyping chip used is given in brackets.

In paper I, no healthy controls were studied. Paper II and IV included 2802 control individuals who were either blood donors from Uppsala and Lund, or population-based controls from the 4 northernmost counties of Sweden. In addition, the validation cohort in paper II included 9859 healthy controls of European ancestry.[110] Paper III includes 45 blood donors from the Uppsala Bioresource.

Collection of clinical data and classification of SLE

All clinical data for the clinical cohorts included in the papers of this thesis, was retrieved from medical charts, with the exception of causes of death in

paper I, which was retrieved from death certificates. In paper I, patients were included in the studies provided that they had a clinical diagnosis of SLE according to their medical records, and that they fulfilled at least 4 ACR-82 or SLICC-2012 classification criteria for SLE.[30, 33] In the discovery and validation cohorts included in paper II and III, patients who had a clinical diagnosis of SLE according to their medical records and fulfilled at least 4 ACR-82 criteria were included.[30] In paper IV, all patients who fulfilled at least 4 classification criteria for SLE according to the ACR-82, ACR-97 or SLICC-2012 criteria, were included.

Definitions of clinical variables

In all papers included in this thesis, the age at disease onset was defined as patients' age at the time of the first clinical diagnosis of SLE according to the medical file and the disease duration was considered the number of years from SLE clinical diagnosis to the time of follow-up. The age at follow-up was defined as the patients age at the time when the medical chart was reviewed or, for deceased patients, the age at death. In paper I and II, organ damage was defined according to the SDI, and extensive organ damage was defined as having an SDI score ≥ 3 points.[53] In paper I, we classified a patient as having hypertension, hyperthyroidism, hyperlipidaemia or depression, respectively, if the patient was receiving antihypertensive therapy, levothyroxine therapy, statins or anti-depressive therapy according to their medical record. Furthermore, in this paper a diagnosis of lymphoma required verification by biopsy, whereas a diagnosis of Secondary SS required a clinical diagnosis set by a rheumatologist according to classification criteria. A Myositis diagnosis was defined as having clinical symptoms with proximal muscle weakness associated with elevated creatinine kinase and/or electromyogram changes and/or biopsy showing myositis, and a diabetes mellitus diagnosis was defined as fasting blood glucose $\geq$7.0 mmol/L and/or glycated haemoglobin A1c $\geq$48 mmol/mol.

In paper II-III, the identification of the patients' age at nephritis diagnosis was based on the point of when they first fulfilled the ACR-82 nephritis criterion according to their medical charts.[30] The Modification of Diet in Renal Disease formula was used to calculate GFRs, and the chronic kidney disease (CKD) system was used to categorize renal function.[176] Classification of renal biopsies was performed using the 1995 WHO- or the 2003 ISN/RPS classification systems.[70, 71] In accordance with this system, mesangial nephritis was defined as class I-II, proliferative nephritis as class III-IV and membranous nephritis as class V.[70] Biopsies indicating signs of nephritis but not fulfilling the criteria for any of the above classes, were classified as other.[70] Patients were classified as having ESRD if they had undergone a renal transplantation or were receiving dialysis. The ACR-82 classification criteria were used to classify anti-dsDNA and anti-Smith antibody (anti-Sm)

positivity.[30] A clinical diagnosis of APS was defined as fulfilled if a patient had positive tests for IgG aCL, IgG aß2GP1 or LA in conjunction with thrombosis and/or obstetric complications in accordance with Miyakis et al.[40] Triple-positivity for aPL was considered to occur if tests for aCL (IgG *and/or* IgM) *and* aß2GP1 (IgG *and/or* IgM) *and* LA were positive, whereas positivity for any aPL required positivity for any of the 3 test variables. A confirmed MI required electrocardiography evidence and a rise in plasma creatine kinase, muscle and brain fraction or troponin T. Ischemic cerebrovascular disease (ICVD) was confirmed by computer tomography or magnetic resonance imaging. VTE, defined as a deep vein thrombosis and/or pulmonary embolism, was identified through venography or ultrasonography, and/or through radionuclide lung scanning or angiogram, respectively. CVE was defined to encompass any of the above events, whereas AE included MI and ICVD only.

In paper III, smoking status was collected from medical charts and patients were categorized as "ever-smokers" if they had ever told their healthcare provider that they were a current or past smoker, or "never-smokers" if they had denied ever smoking according to their medical chart. For patients whose medical file did not reference smoking, the data was considered missing for the variable.

Ethics approvals

Oral and/or written informed consent have been obtained from all individuals included in the papers of this thesis. For paper I, the study protocol was approved by the Regional Ethical Review Board in Linköping (DNR M75-08/2008) and Uppsala (DNR 2016/155). For paper II, the study protocol was approved by the Regional Ethical Review Board in Uppsala (DNR 2016/155) and the local ethics committees. The Regional Ethical Review Board in Uppsala and the local ethics committees also approved the study protocol for paper III (DNR 2020-05065) and paper IV (DNR 2009/013 and 2020-05065). Paper IV is partly based on analysis using datasets from the FinnGen consortium. The University of Helsinki is responsible for the FinnGen research project, and the Coordinating Ethics Committee of the Helsinki and Uusimaa Hospital District have provided ethical approval of the project.[177]

Genotyping and quality control procedures and selection of SNPs

In paper II−IV, patients (and healthy controls, in paper II and IV) were genotyped using a 200 K Immunoarray chip (IC) array (Illumina). In paper IV, an additional 486 patients were genotyped using the Global Screening Array (GSA, Illumina), table 5. All samples were checked for autosomal heterozygosity rate, mislabelled gender, European ancestry and relatedness, and all

ambiguous samples were excluded. SNPs with more than 5% missing data or a MAF below 1% were excluded, and 61 (paper II), 60 (paper III) and 55 (paper IV) SNPs with previous GWAS association ($p<10e-8$) with SLE development were selected for further analysis.[101] Patients missing data for any of the selected SNPs were excluded from further analyses.

The patients and controls of the validation cohort in paper II were genotyped on the custom-designed Immunochip Illumina Infinium Assay according to Illumina's protocols, using the Illumina iScan scanner.[178] In total, 6,748 cases and 11,516 controls with European ancestry were genotyped and passed quality control for the Immunochip, and 120,873 SNPs passed quality control analyses in ≥1 ancestry. Of these, data for the 57 genetic variants studied in paper II, located outside the HLA region, were selected. Individuals with <100% genotype success rate of the 57 SNPs were excluded from the validation cohort, and due to an overlap in this cohort with the Swedish SLE cohort, subjects identical to any individual included in the discovery cohort (pi-hat > 0.9) were identified and excluded, leaving 5 524 patients and 9 859 controls available for analysis.

Genotyping of the replication cohort in paper III was performed using the iPLEX chemistry on a MassARRAY system (Agena Bioscience, San Diego, CA, USA). Quality control procedures included a minimum per sample call rate of 90% and a per variant call rate of 90%.

Genetic risk score construction

In paper II, a logistic regression model was used to investigate associations between each of the 57 genetic variants located outside of the HLA region included in the study, and SLE. In addition, a separate HLA-GRS, based on 4 SNPs located in the HLA region, was constructed. For each SNP, the natural logarithm of the OR for SLE development, based on the patients and controls in the discovery cohort, was multiplied by the number of risk alleles (0, 1 or 2) in each individual. The sum of all weighted SNP values in each individual, was defined as the GRS. In paper IV, the same 57 non-HLA genetic variants as in paper II, were used to create 9 different weighted GRSs, representing 9 different clinical manifestations. In this paper, each SNP received a weight corresponding to the effect size of the SNP-manifestation association based on each of the datasets. When the SNP-manifestation associations displayed a direction of the effect opposite to that between the SNP and SLE according to previous GWAS results, the weight was defined as zero. The 9 GRSs were subsequently standardized by subtracting the mean value performing division with the mean GRS value, to facilitate the interpretation of ORs.

Mendelian Randomization analysis

Mendelian Randomization (MR) is a genetics-based instrumental variable (IV) approach that relies on the random assignment of genetic variants at conception, to investigate the causal effect of an exposure on an outcome.[179, 180] In order for the MR results to be valid, the included IVs must fulfil 3 core assumptions; 1) the IVs must be strongly associated with the exposure; 2) the IVs must not be related to any confounder which affects both the risk of the exposure and the outcome; 3) the IVs must not influence the outcome in any other way than through the exposure. Whilst fulfilment of the first assumption can be directly ensured by selecting well-investigated genetic variants displaying clear association with the exposure, evaluation of the second and third assumptions requires indirect investigation through additional MR sensitivity analysis and tests of pleiotropic effects.[179, 180]

Exposure dataset

Since in paper II, the included SLE risk SNPs were extensively validated in a validation cohort, we determined these SNPs to fulfil the relevance assumption, and therefore used these SNPs a representation of the genetic SLE predisposition in paper IV. For each of these SNPs, the original GWAS was used to retrieve summary level statistics, including risk alleles, effect sizes and standard errors.[13, 103, 110, 113, 117, 175, 181-190]

Outcome data sources

Paper IV, utilizes the FinnGen consortium database (Data Freeze 5, 2021) for identification of datasets which as closely as possible resemble the ACR-82 classification criteria.[30, 177, 191] To prevent inclusion of datasets not relevant for SLE, the search was narrowed down to datasets including a number of SLE patients representing at least twice of what would be expected based on the prevalence of SLE in the entire FinnGen database. Datasets were further limited to include the International Classification of Diseases (ICD)–10 codes corresponding to *diseases of the skin and subcutaneous tissue* (L00–L99), *diseases of oral cavity, salivary glands and jaws* (K00–K14), *diseases of the musculoskeletal system and connective tissue* (M00–M99), *other forms of heart disease* (I30–I52), *other diseases of pleura* (J90–J94), *glomerular diseases* (N00–N08), *diseases of the nervous system* (G00–G99), *diseases of the blood and blood–forming organs and certain disorders involving the immune mechanism* (D50–D89), and *abnormal findings on examination of blood, without diagnosi*s (R70–R79).[192] Moreover, a SLE GWAS performed on 538 patients and 176,590 controls was included.[191]

Interleukin-12-induced phosphorylation of STAT4

In paper III, data from a previous study examining Interleukin 12 (IL-12)-induced phosphorylation of STAT4 (pSTAT4) in 72 healthy blood donors from Uppsala Bioresource, was included.[193] Of the 72 blood donors previously analysed, we included a subset of donors (n=45) who had data on smoking. Of these individuals, 20 were past or current smokers and 25 were non-smokers.

Statistical analysis

In paper I, chi-square tests or Fishers exact test were used to compare frequency distributions, and comparisons on interval/ratio scales were performed using independent t-tests or Mann-Whitney U tests. In addition, Poisson regression was used for uni- and multivariate analysis to investigate associations between clinical variables and organ damage, and these analyses were adjusted for age at diagnosis and disease duration. In addition, the Kruskal-Wallis test was used for comparison of groups with different SDI scores and disease duration.

In paper II-IV, logistic regression models were used to assess differences between groups for all binary outcomes. In paper II, ordinal regression models were employed to assess the association between the GRS and SDI scores or CKD stages, and the generalized Wilcoxon test was used for survival analysis. All logistic and ordinal regression models in paper II were adjusted for the age at follow-up, and statistically significant associations were further adjusted for in a secondary analysis using the age at SLE onset.

In paper III, the associations between each of the included 60 SNPs and MI were analysed in separate logistic regression models including age at follow-up and disease duration as covariates. In subsequent analysis, all SNPs with a direction of the effect corresponding to that of SLE itself, were included in a forwards conditional logistic model with age and disease duration. Paper III also included multiplicative interaction analyses, which were performed by adding an interaction term in the logistic regression; and additive interaction models, which were performed by calculating the attributable proportion due to interaction.[194, 195] In addition, in Paper III, differences in pSTAT4 were assessed using Student's T test as well as a linear regression model to allow for the adjustment of age and the number of *STAT4* risk alleles.

In paper IV, MR analysis was conducted using the TwoSampleMR package.[196, 197] The inverse variance weighted (IVW) model was used as an initial test to determine the association between the SLE genetic risk and each of the 30 investigated datasets. Additional MR methods, specifically the MR

Egger method, simple and weighted median methods and a sign concordance test, were employed for sensitivity analysis.[179, 180] A test of the MR-Egger intercept and assessment of the Q statistics for heterogeneity with respect to IVW, was further carried out.[179, 180] In analysis of the 9 constructed GRSs in this paper, age at follow-up and gender were used as covariates. In all papers included in this thesis, p-values <0.05 were considered statistically significant.

Results and discussion

Paper I

The majority of Swedish systemic lupus erythematosus patients are still affected by irreversible organ impairment: factors related to damage accrual in two regional cohorts

Despite increased understanding of SLE, development of new treatment strategies and decrease in mortality in recent decades, the accrual of irreversible organ damage among patients affected by the disease remains a critical concern. In this paper, the main objectives were to characterize organ damage, as defined by the SDI, and to identify factors associated with damage accrual and death.[53] The study includes 543 patients with SLE from 2 Swedish university clinics in Uppsala and Linköping who were predominantly of Caucasian ethnicity (>90%). More than 95% of the patients had an established disease, defined as more than 1 year of disease duration.

At the time of follow-up, the mean disease duration in the combined study population was 17 years. At this point, 318 of the 543 patients (59%) had ≥ 1 incident of organ damage recorded. Upon retrospective investigation of the accumulation of organ damage from the time of diagnosis to the time of follow-up, the highest rate of damage accrual was observed in the first few years following SLE diagnosis, with around 12% of patients experiencing organ damage in the first year, 30% within the first 5 years and 45% within 10 years (figure 4). These figures are largely in accordance with a previous study from the University College London Hospital (UCLH), in which 50% of patients with SLE were found to have acquired some organ damage after 10 years of disease.[198] Other studies have provided both higher and lower figures, with differences possibly attributable to variations in factors such as age and ethnicity composition of the examined cohorts. In a Portuguese study examining the damage accrual in 976 patients with SLE who were around 5 years younger at follow-up compared with the present study, only 37% presented with damage after 14 years.[199] Conversely, a study of the multi-ancestry SLICC inception cohort exhibited higher rates of damage, with 24% of patients having acquired some damage already after a disease duration of 1.2 years.[200]

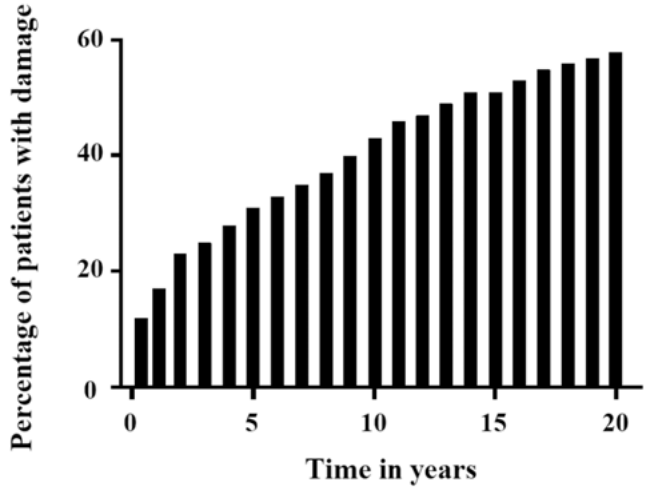

Figure 4. The accumulation of organ damage in the study population from the time of SLE diagnosis until follow-up. Figure adapted from paper I, previously published in Lupus (2019).

Of the individuals with ≥1 point on the SDI in the present study, 137 (43%) had "extensive damage", defined as an SDI score ≥3. The most prevalent damage domains were found to be neuropsychiatric damage, including cerebrovascular accidents and seizures, which affected a quarter of the patients. This was followed by ocular, cardiovascular, and musculoskeletal damage. These findings are essentially in line with other studies, with the aforementioned UCLH and Portuguese studies identifying neuropsychiatric damage as the primary or secondary most common domain.[198, 199]

Upon comparing patients with any damage (SDI>0) to patients without damage, those with damage were found to be older at diagnosis and to more often be of Caucasian ethnicity, possibly reflecting the association between these variables with a higher age at follow-up and longer disease duration, respectively, which are well-known risk factors for organ damage.[10, 52, 200] In addition, patients with any damage were found to fulfil a higher number of SLICC-2012 criteria, with fulfilment of its neurologic disorder criterion being specifically associated with damage accrual. Furthermore, a secondary diagnosis of APS or positive aPL test at any point in the disease history, was associated with an increased risk of having accrued organ damage. Lastly, associations were found between the occurrence of damage and SS, hypertension, hyperlipidaemia, depression and several therapies for SLE including cyclophosphamide, ciclosporin and mycophenolate mofetil. When analysing associations with extensive damage, in addition to the variables above, associations were found with serositis, haematological and renal disorder, and interstitial lung disease (ILD). These results were not unexpected, since APS, renal

disorder and ILD are well-known to be associated with a poorer prognosis in SLE.[199, 201, 202] In addition, treatment with more aggressive immunosuppressants may be indicative of a more severe disease, leading to increased accrual of organ damage.

To identify independent risk factors for organ damage, a multiple Poisson regression analysis was carried out. Upon adjusting this model for disease duration and age at SLE diagnosis, several independent associations with damage accrual were identified, including fulfilling the ACR-82 or SLICC-2012 classification criteria for pericarditis, AIHA, lymphopenia, and neurologic disorder. Treatment with prednisolone, cyclosporine and cyclophosphamide, or for hypertension or hyperlipidaemia, were also associated with an increased risk of organ damage. In addition, a clinical APS diagnosis, myositis and a secondary diagnosis of SS were associated with organ damage. In total, the identified factors were found to explain around half of the variation in SDI scores. Several of these manifestations, including pericarditis, AIHA, lymphopenia and myositis, have to the best of our knowledge not previously been described in this context. Whilst renal disease is a well-known risk factor for the accrual of damage, this association did not show up as significant in the multivariate model. However, treatment with angiotensin converting enzyme inhibitors is usually prescribed to patients with nephritis to reduce proteinuria. Anti-hypertensive treatment, which was significant in the model, may therefore be considered a proxy for nephritis in this analysis.

54 patients (10%) in the cohort were deceased at the time of follow-up, with malignancy being the leading cause of death, followed by infections and CVD. In concordance with previous studies demonstrating a strong relationship between higher scores on the SDI and an increased risk of death,[9, 201] deceased cases were found to have higher SDI scores compared to patients alive at follow-up. The mean age at death among the 54 cases was 70 years, with 10 of the cases found to have died before the age of 60. Importantly, due to the retrospective approach of this study and the fact that recruiting patients to the study sometimes occurred many years after the onset of SLE, it is likely that these figures under-estimate the true mortality risk.

There are several strengths of this study, including a large and well-characterized study population and a high estimated coverage of patients in the 2 included regions of Sweden. There are, however, also some limitations. Whilst we had excellent, retrospective data for the SDI, detailing both the year and the type of damage for each event, we did not have access to information on when the examined risk factors were first fulfilled. Whilst we found many associations between these risk factors and the presence of organ damage, it therefore remains unclear whether the risk factors occurred before or after each event of organ damage. In addition, despite efforts to adjust the multi-

variate model for age- and disease duration related effects, due to a strong correlation between these variables and the age at follow-up, we could not adjust the model for all 3 age variables. Furthermore, some associations in this study may be influenced by similarities in the classification criteria for SLE and the SDI. For example, pericarditis, which was significantly associated with an increased risk of organ damage in the multivariate model, is included as an item in the SDI.[53] However, the impact of this overlap is likely minimal, as only 4 individuals had an event of pericarditis recorded in the SDI. Lastly, some significant differences were present between the 2 populations included in this study, with higher rates of malar rash, photosensitivity, oral ulcers, and positivity for anti-dsDNA and anti-Sm in the Uppsala patients, whereas the Linköping patients included older patients with a shorter disease duration and higher prevalence of aPL and APS. Due to these differences, analysis of each individual cohort separately would have been valuable, although this was not performed due to a risk of insufficient statistical power.

In conclusion, the study demonstrates that the majority of patients with SLE experience severe consequences of their disease. Additionally, we have identified several independent risk factors for the accrual of organ damage, of which some have not previously been described.

Paper II

High genetic risk score is associated with early disease onset, damage accrual and decreased survival in systemic lupus erythematosus

Of the more than 180 independent genetic variants associated with SLE, only a smaller subset has been linked with specific clinical manifestations of the disease.[124] In the few previous publications exploring the association between the cumulative genetic risk of SLE and the occurrence of specific subphenotypes, analyses have been limited to the ACR-82 classification criteria, disease activity, sex and age at onset.[203-206] To enhance the understanding of how a high genetic risk of SLE affects the long-term prognosis of the disease, we investigated the association between a GRS based on 57 established SLE risk SNPs, and factors associated with more severe SLE manifestations. This included the accumulation of organ damage, as defined by the SDI, overall mortality, the development of ESRD and CVE, and the presence of aPL, anti-dsDNA and anti-Sm.[53] The study included 1,001 patients with SLE, for which case-case analyses were conducted to compare clinical variables collected from medical charts. To validate the effect on the risk of developing SLE conferred by the GRS, we also included 2,802 Swedish healthy controls, as well as a replication cohort consisting of 5,524 European patients with SLE and 9,859 healthy controls.

In our case-control analysis, we found a strong association between a high GRS and the probability of having SLE, with a high-to-low quartile comparison showing over 12-fold and 7-fold higher risks of SLE in the discovery and validation cohorts, respectively. A strong correlation between the mean GRS and the prevalence of SLE was further identified upon division of the individuals into smaller groups, with the prevalence of SLE in the validation cohort being ~5% and ~90% in the groups with the lowest and highest GRSs, respectively (figure 5).

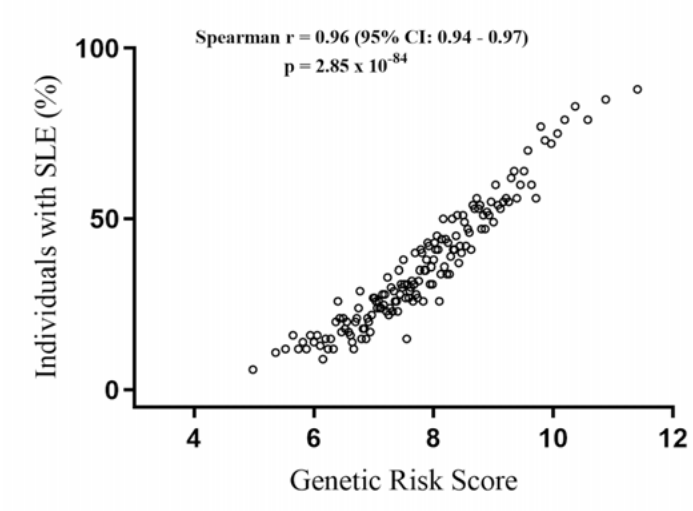

Figure 5. The correlation between the genetic risk score and prevalence of SLE, with each point representing a group of 100 patients and controls from the validation cohort. Figure adapted from paper II, previously published in *Ann Rheum Dis.* (2020)

This paper also confirms a previously reported association between the cumulative genetic risk of SLE and an earlier onset of the disease.[110, 206] When comparing the quartiles of patients with the highest and lowest GRS, the mean age at onset was found to be 33 years in the high quartile, compared with 39 years for individuals with a GRS in the low quartile. Interestingly, however, in our cohort the difference in onset between patients with a high and low GRS began already during childhood and reached its peak at 20 years of age. In contrast, in the largest study of GRSs in European individuals with SLE predating our study, the deviation between the groups did not occur until after age 20 and reached its peak around age 40.[110] Although these differences would need to be scrutinized in future studies, it may be an indication that different SNPs are associated with early- and late onset SLE, respectively, and

that the exact consequence of a high genetic risk for SLE on the onset age, is thus also dependent on individual SNPs.

In analysis of associations between the GRS and clinical outcomes, we found that patients with a high GRS were more likely to have accrued organ damage and to be deceased at the time of follow-up. In analysis of the ACR-82 criteria, renal variables and CVE, we could further demonstrate that a high GRS was associated with doubled odds of fulfilling the ACR-82 nephritis criterion and more than 5-fold elevated odds of having ESRD. Furthermore, we investigated renal biopsy results from 222 patients with positive renal biopsies, and found a significant association between a high GRS and the risk of proliferative nephritis (WHO class III-IV). In addition, we found that a high GRS was associated with a decreased time to the first CVE, and in particular AE, as well as with an increased risk of having tested positive for aPL.[207] As renal involvement, CVD and the presence of aPL are important causes of permanent organ damage and death in SLE, these associations may be the explanation of the association between a high GRS and organ damage or death.[48, 57, 62]

An observation made in family studies of many multifactorial diseases is that when the least affected sex is affected by a disease, the recurrence risk is higher, presumably because the genetic liability is higher in these individuals.[74] In accordance with this hypothesis, a previous study examining GRSs in SLE showed a very strong association was found between male gender and a GRS for SLE.[205] Contrarily, however, another study showed the opposite association, with a high GRS strongly associated with female gender.[204] Importantly, both of these studies used sex-dependent weights acquired from the same cohort in which the scores were tested, which may have introduced bias in the analyses. In our study, no association between a high GRS and male gender could be identified. However, as the GRS used in this study did not include variants on the X-chromosome and several SLE risk variants are located on this chromosome, this study does not rule out that such an association is present.[101]

Whilst most of the associations between the GRS and clinical manifestations identified in this paper were positively associated with a high GRS, we found that for photosensitivity, a high GRS conferred a protective effect. Whilst the reasons for this finding are not entirely clear, a possible hypothesis may be that patients who are sensitive to this environmental risk factor, require a lower genetic risk in order to develop SLE.

In conclusion, the findings of this study contribute to the comprehension of the genetic aetiology of permanent organ damage, CVE, renal dysfunction and overall mortality in patients with SLE. Furthermore, the results provide an indication that GRSs may be potential tools for prediction of disease severity,

which may be useful both in a clinical setting and for stratification of patients in clinical trials.

Paper III

Interaction between the STAT4 rs11889341(T) risk allele and smoking confers increased risk of myocardial infarction and nephritis in patients with systemic lupus erythematosus

Whilst the prognosis of LN has improved in recent decades, death from CVD remains a major challenge in SLE.[56, 57] It is known that both traditional risk factors like smoking, and SLE-related risk factors such as a high disease activity or nephritis, are important for the development of CVE. A considerable part of the excess risk seen in patients with SLE, however, remains unexplained.[60, 61] In 2 well-characterized cohorts of patients with SLE, with a combined size of over 1200 patients, we investigated how gene-smoking interactions contribute to the development of specific manifestations of SLE.

Whilst smoking in itself was found to be a risk factor for MI, a key finding of this study was that its effect was enhanced in patients carrying the *STAT4* SLE risk allele rs11889341. In the discovery cohort, smokers carrying the risk allele were observed to have a 2.7-fold elevated risk compared with non-smokers carrying the risk allele. In smokers carrying 2 risk alleles, the risk was 8-fold elevated, and in the replication cohort, patients carrying ≥1 risk allele displayed a 6-fold elevated risk. Through both additive and multiplicative interaction analysis, we demonstrated that the effect of the *STAT4* risk allele on MI was dependent upon smoking status amongst the patients, with the proportion of the risk attributable to the interaction estimated to 54%. In addition, we found a similar interaction between smoking and the *STAT4* risk allele on the risk of nephritis. Patients with nephritis are known to have an increased prevalence of CVD and it has been demonstrated in previous studies that the STAT4 risk allele is associated both with nephritis and severe renal insufficiency.[153, 154, 208] We therefore speculated that these results were connected. However, after performing stratification and analysing only patients without nephritis, we could conclude that the 2 associations occurred independently of each other.

It has previously been demonstrated that the *STAT4* risk allele is associated with increased levels of pSTAT4 in activated CD8+ T cells of SLE patients.[193] Increased STAT4 messenger RNA expression has further been associated with increased cardiovascular damage in patients with SLE.[209] We therefore hypothesized that smoking, similarly to carrying the STAT4 risk allele, results in higher levels of pSTAT4. To test this hypothesis, we analysed pSTAT4 levels in CD8+ cells of 45 healthy controls with available smoking

data (ever vs never). We found the levels to be twice as high in the smokers (mean geometric fluorescence intensity: 1063 vs 565). This is in line with results from a previous study demonstrating higher levels of pSTAT4 in bronchial T cells of smokers, compared with non-smokers.[210] Due to these findings, we speculate that for patients with SLE who both smoke and carry the risk allele, the alteration in levels of activated STAT4 is larger, resulting in cardiovascular and renal manifestations.

Lastly, when assessing associations between 60 SNPs with previously confirmed association with SLE ($p<5.0E-8$), we found the Interleukin 12A *IL12A* SLE-risk SNP rs564799 to be associated with MI in both the discovery and replication cohorts. The *IL12A* risk allele is located within the *IL12A* gene, which encodes a subunit of the IL-12 protein (figure 6).[211] The binding of this protein to the IL-12 receptor, leads to phosphorylation of STAT4, which subsequently translocates to the nucleus, constituting a transcription factor for hundreds of genes. Examples of consequences include IFN-γ production, T helper type 1 and -17 differentiation and monocyte activation.[211] These findings point further to the importance of STAT4 activation for the development of MI in patients with SLE. Indeed, when we examined the effect of carrying both the *STAT4* and *IL12A* risk alleles in the discovery cohort, we found that patients homozygous for both the *STAT4* and the *IL12A* risk alleles displayed a statistically significant and more-than 5-fold increased risk of MI, compared with patients with any other combination of the risk alleles. In the replication cohort, the same analysis demonstrated a statistically significant and more than 7-fold increased risk in patients homozygous for both risk alleles.

The findings of this paper are important as they may provide a clue about a possible mechanism behind the increased prevalence of CVD in SLE. CVD is a leading cause of death in the disease, and the identification of pathways involved in the development which could be targeted by future therapies, is therefore of great importance. Besides the *STAT4* and *IL12A* risk alleles, several SNPs associated with genes involved in the STAT4-IL-12 signalling pathway have displayed genome-wide or suggestive association with SLE in previous GWASs, including Janus Kinase 2 (*JAK2)*, Tyrosine Kinase 2 (*TYK2)*, Interleukin 12B (*IL12B*) and Interleukin 12 Receptor Subunit Beta 2 (*IL12RB2*) (figure 6), however their effect on the development on CVD in SLE is unexplored.[110] In previous work, the potential of an efficient blocking of STAT4 phosphorylation by Janus kinase (JAK) inhibitors has been demonstrated, which appears to ameliorate murine lupus and to improve its associated vascular dysfunction.[193, 212] In light of the findings of this study, the effect of the additional SLE risk alleles on the risk of CVD, and of specific blocking of the STAT4-IL-12 pathway in SLE, warrants further investigation.

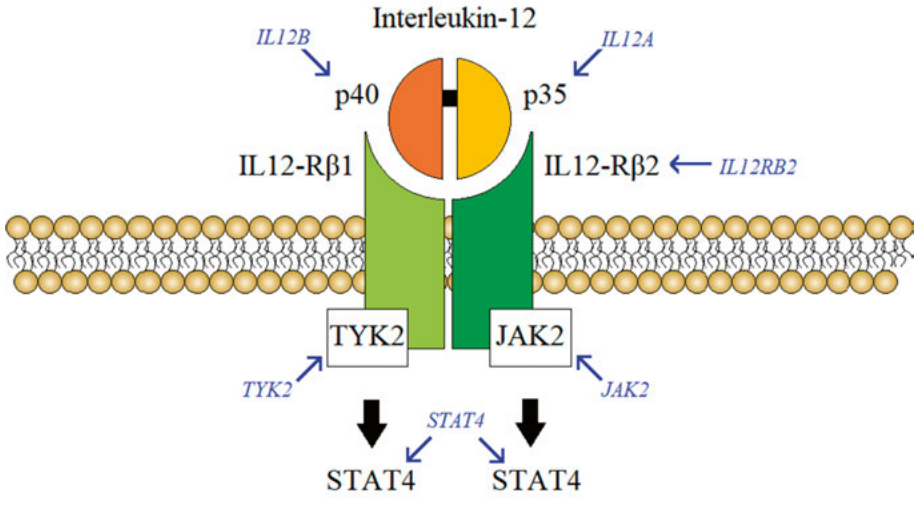

Figure 6. Components of the IL-12-STAT4 signalling pathway, resulting in phosphorylation of STAT4. Genes in genetic loci with genome-wide association ($p<5\times10-8$) for SLE development according to previous studies, are printed in blue.

In conclusion, our results suggest that the IL-12-STAT4 signalling pathway may be of importance in the development of CVE and renal manifestations in patients with SLE. In addition, the findings stress the importance of smoking cessation in patients with SLE.

Paper IV

From genetic predisposition to clinical outcome in systemic lupus erythematosus: construction and validation of subphenotype-specific genetic risk scores

The findings of an association between the cumulative genetic risk for SLE and certain specific manifestations of the disease, as demonstrated in paper II of this thesis, may suggest a potential role of genetic risk prediction in a clinical setting. It is imaginable, however, that the predictive ability of a GRS would be enhanced if the model was based on the risk of each specific manifestation, rather than the risk of SLE itself. In this study, we explored the possibility of increasing the number of manifestations associated with the cumulative genetic risk of SLE, by inferring knowledge of SNP-subphenotype associations from a large non-SLE population. To acquire this knowledge, we explored how a high genetic predisposition towards SLE affects the risk of developing manifestations relevant for the disease, using publicly available GWAS data based on 321,000 individuals from Finland.[177, 191] Using the

FinnGen database, a systematic search for ICD-10 codes representing diagnoses which as closely as possible resemble the ACR-82 classification criteria, was conducted, resulting in the identification of 30 relevant datasets. To explore how a high genetic risk of SLE development, represented by the 57 SLE risk SNPs, affects the risk of being diagnosed with each of these 30 manifestations, MR analysis was conducted.[179, 180] Lastly, we examined whether SNP-manifestation associations identified in these datasets could be validated in a clinical cohort consisting of 1487 patients with SLE from Sweden and Norway.

Of the 30 selected datasets, 5 were related to skin manifestations (figure 7). In the MR analysis, a high genetic risk of SLE was found to increase the probability of being diagnosed with the skin manifestation, in all 5 datasets. Among the manifestations, the analysis for rosacea displayed the highest OR and lowest p-value (figure 7). This result is interesting as it corroborates results from a few previous studies, which have provided support for an association between rosacea and auto-immune disease. For example, in a population-based case-control study involving over 6000 individuals with rosacea and 30 000 controls, rosacea was associated with an over 2-fold elevated risk of developing T1D, celiac disease and RA, and a 65% increased risk of developing MS.[213] In a study from 1970, an SLE-like deposition of immunoglobulins in the skin of patients with rosacea was reported.[214] Moreover, the HLA allele DRB1*03:01 has stood out amongst the strongest genetic factors associated with both rosacea and SLE.[110, 215] The IVs included in the present study were all located outside the HLA region, and the association found between SLE and rosacea may thus be an additional indication of a connection in the aetiology. However, rosacea is also known as a common “SLE mimicker”, with an SLE malar or discoid rash easily mistaken for rosacea.[216] Therefore, it is not entirely clear whether the association represents a shared genetic background between the diseases, or is an indication that individuals with early or incomplete SLE initially tend to be misdiagnosed with rosacea.

For arthritis, we examined 7 datasets, of which 5 were significantly associated with the SLE genetic risk (figure 7). In accordance with the clinical picture seen in SLE, the strongest association was found for polyarthrophathies. Correspondingly, the 2 non-significant datasets were for pyogenic arthritis and reactive arthropathies, which plausibly have different aetiologies than immune-mediated arthritis. Importantly however, in subsequent analyses performed to assess the fulfilment of the MR assumptions,[180] we found indications that some of the individual 57 SNPs affected the risk of polyarthropathies through an alternative pathway to SLE. It is well-known that many IVs included in this study increase the risk of both SLE and other auto immune diseases, including RA, which is also a cause of polyarthropathies.[217]

Hence, the strong association found in this study does not necessarily indicate a causal relationship.

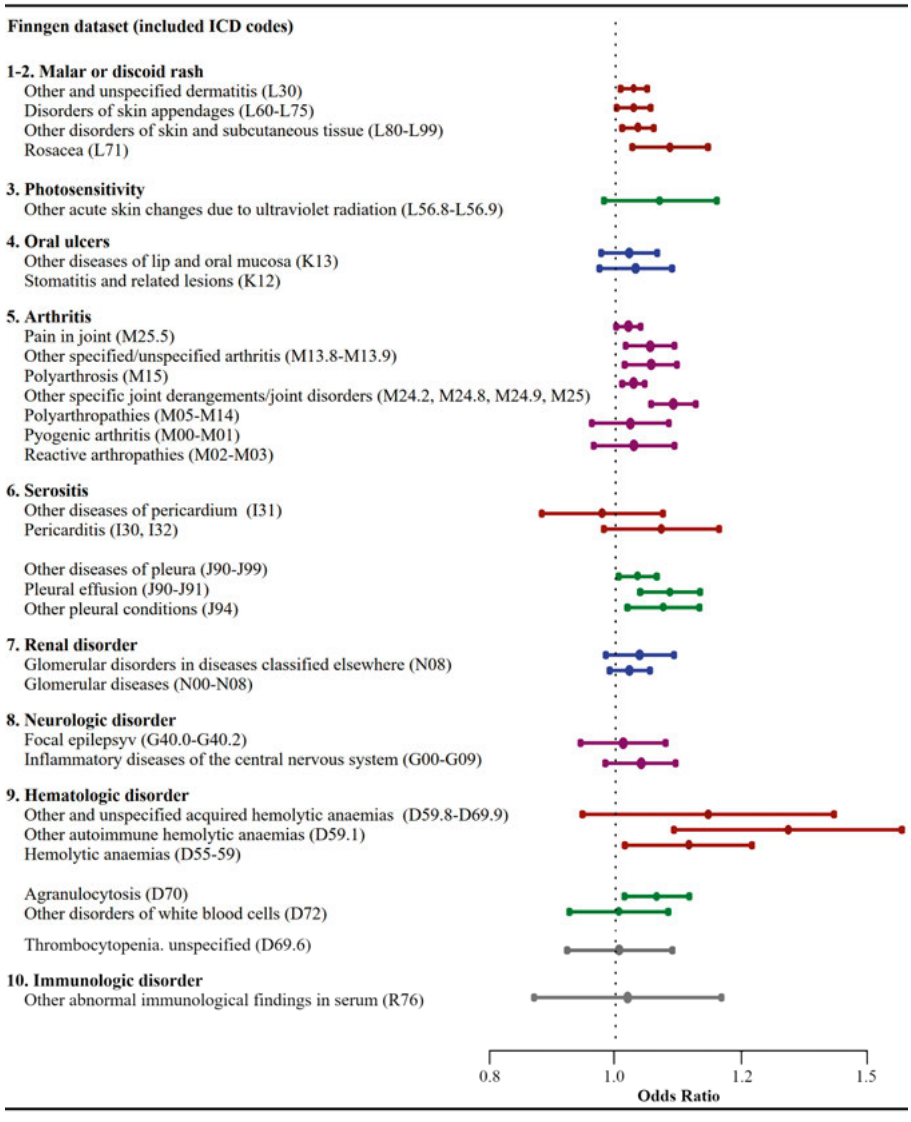

Figure 7. Results from the Mendelian Randomization analysis. Figure adapted from paper IV.

For serositis, we included 2 datasets for pericarditis and 3 for pleuritis (figure 7). Whilst neither of the 2 pericarditis datasets were associated with a high genetic SLE predisposition, we found associations between the SLE genetic risk and all 3 pleuritis datasets. The most significant association was for pleural effusion, which is a common manifestation in SLE.[30]

For hematologic disorder, we found associations between the SLE genetic risk and haemolytic anaemia in 2/3 examined datasets (figure 7). In addition, we found an association with agranulocytosis, whereas the dataset representing other disorders of white blood cells and thrombocytopenia did not display any significant association with the SLE genetic risk. Lastly, for the dataset selected to represent the ACR-82 immunologic disorder criterion, no significant association with the SLE genetic risk was present (figure 7).

Next, a single dataset per manifestation was selected, narrowing the number of datasets down from 30 to 9, and individual SNP-manifestation associations in the datasets were scrutinized. In most datasets, the number of SNPs statistically significantly associated with the respective manifestation was low, with the exception of arthritis where 15 SNPs with a $p<0.05$ and the same direction of effect as previously reported for SLE, were identified. Of these, 2 SNPs associated with the PTPN22 (rs2476601) and SHB3-ATXN2 (rs10774625) loci, met the GWAS significance threshold of $p<5e-8$.

To assess whether the cumulative effect of the individual SNPs associated with each manifestation, could be used to predict the presence of each corresponding ACR-82 criterion in SLE, a GRS for each of the 9 datasets were constructed and analysed in the clinical cohort using a logistic regression model. For 5 of the 11 ACR-82 criteria, specifically arthritis, nephritis, neurologic disorder, hematologic disorder and immunologic disorder, the GRS was found to be significantly associated with the corresponding ACR-82 criterion. This indicates that, in addition to the association of the cumulative genetic risk of SLE with nephritis and immunologic disorder shown in paper II in this thesis as well as by others,[206, 218] some SNPs within the genetic risk of SLE give rise to these additional subphenotypes.

This study has several strengths, including the use of a large number of well-established SLE risk SNPs and a well-characterized clinical cohort of nearly 1,500 individuals with SLE. However, there are also some limitations to consider. Firstly, this study did not evaluate the possibility of a reversed relationship between the SLE genetic risk and the examined manifestations, due to the low number of GWAS-associated SNPs in most datasets. Secondly, although a number of additional sensitivity analyses have been employed to reduce the possibility of violation of the MR IV assumptions, it is not possible to completely rule out that some SLE risk SNPs are associated with confounders which in turn lead to development of certain manifestations. Our analyses did not find strong evidence of pleiotropy for any of the examined manifestations, however, in analysis of the polyarthropathies dataset we found strong heterogeneity among individual SNPs. Similarly, it is unclear whether certain associations discovered in the MR analysis reflect biological mechanisms or other factors, such as in the case of rosacea described above.

In conclusion, this study highlights the potential of combining analysis of large biobank datasets and a smaller clinical cohort, to overcome the challenges entailed by rare diseases like SLE.

General discussion

The focus of this thesis has been to explore the impact of SLE genetic variants, through their cumulative effect or gene-environment interactions, on SLE manifestations and complications. The recent revolution in the field of genetics, brought about by rapidly decreasing genotyping costs, has enabled the genotyping and analysis of millions of individuals. This has offered substantial benefits to our comprehension of the underlying pathogenic mechanisms of many diseases. In addition, clinical trials supported by GWAS findings are more than twice as likely to yield positive outcomes, emphasizing the importance of identifying additional genetic markers associated both with SLE itself, and with specific subphenotypes.[86, 87] With this in mind, a central discovery presented in this thesis is the potential role of the IL12-STAT4 pathway in the development of CVE and nephritis. This pathway, which may constitute a potential drug target in SLE, is further associated with several additional SLE risk loci, including *JAK2*, *TYK2* and *SOCS1*.[110, 219] Given our findings, further exploration of these variants, their functional consequence and their individual and combined impact on the clinical picture of the disease, may provide additional knowledge about the role of this pathway in SLE.

Considering the challenges of early recognition of SLE, with delays in diagnosis potentially leading to postponements in adequate treatment, the idea of being able to employ a genetic model as a diagnostic aid in the future of precision medicine – especially for patients at risk of developing more severe manifestations – is intriguing. In this context, another key finding presented in this thesis is the strong association between a high GRS for SLE, and the development of the disease itself as well as its renal manifestations. In 2018, when our study on GRSs was initiated, the 57 SNPs included could be considered representative of the vast majority of established non-HLA SLE genetic loci in the European population.[101] However, due to the rapid evolvement in the field of genetics, at present point this GRS comprises less than a third of the ~180 currently identified SLE risk loci. [124] This raises an important question: Will the predictive ability of our model be further improved in the future, and if so, to what extent?

Ultimately, the predictive potential of a GRS for a particular trait is restricted by the trait heritability. This factor determines the highest AUC that could

theoretically be achieved from a ROC curve, under the assumption that all genetic contributors to the disease were known and included in the model.[84] As SLE is a rare disease and has a moderate MZ twin concordance, doubt has been raised about the prospect of using GRSs as a screening tool for the disease, even if extensive knowledge of the genetic aetiology was available.[124] Contrarily to this view, however, mathematical estimations have demonstrated that, assuming an heritability of SLE at the high end of current estimations, an AUC of 1.0 can theoretically be achieved. Moreover, assuming inclusion of variants explaining half of the genetic variability of the disease, an AUC of 0.95 has been estimated to be achievable.[84, 85, 220]

In the ROC curve analysis performed in this thesis, the AUC in the validation cohort was found to be 0.71. In spite of the progress in GWAS studies after the publication of our study, with estimations showing that half of the genetic risk of SLE has been identified, results from subsequently published papers have at best been in line with this figure.[101] In a large study including more than 7,000 patients with SLE with European and Chinese ancestry, the AUC was found to vary between 0.64 and 0.72 when testing GRSs for SLE within European and between the European and Chinese groups.[218] The best-performing PRS in this study was later validated in the UK biobank material, generating a lower AUC of 0.64.[221] Consistent with the latter result, an examination of a GRS based on 80 SLE risk SNPs in a small number of patients from the Nurses' Health Study (NHS), also resulted in an AUC of 0.64.[222] These lower values may reflect a difference in the selection of patients, with the UK Biobank and NHS possibly including more patients with a milder disease compared with studies performed on patients recruited from hospital clinics.

With this discrepancy between the suggested possible accuracy of GRSs and the results of these studies, it remains unclear whether the genetic risk of SLE has a purpose in the future as a diagnostic test. It is, however, possible that the current knowledge may provide some usefulness in certain situations where the pre-test probability of SLE is higher than in the general population. In a study comparing GRSs for different rheumatic diseases in patients with inflammatory arthritis, it was shown that taking GRSs into account could help discriminate true diagnoses across 3 different cohorts, with AUCs for the predictions ranging from 0.69-0.84. Taking genetics into account, the accuracy of an initial rheumatologist evaluation of patients with arthritis was further shown to increase from 39 to 51%.[223]

Furthermore, as demonstrated in this thesis, using GRSs to differentiate between mild and severe cases in patients who have already developed SLE, may be another approach of using genetic risk models in a clinical setting. In our evaluation of predictive ability of the GRS for manifestations of SLE, the

strongest result was found for ESRD, with a sensitivity and specificity at the optimal cut-off level of 63% and 83%, respectively. Calculations of positive and negative predictive values further showed that at a prevalence of ESRD of 11%, it would be possible to divide patients into a group with only a 5% lifetime risk of developing the complication, and a group with a 31% lifetime probability. The accuracy for ESRD is thus approaching values which would be required for clinical utility. However, in light of the variation between AUCs for SLE found in different studies, cautiousness in the interpretation of these findings is important and 2 limitations of the paper, likely affecting the average disease severity of the cohorts in opposite directions, are important to consider: Firstly, the retrospective collection of clinical data, with the inclusion of patients diagnosed with SLE decades ago only possible for patients still alive at the time of DNA sampling, may result in patients with the most aggressive disease courses being included to a smaller extent than expected. Secondly, as all patients included in the study were recruited from university hospital clinics, the patients with the mildest disease can also be expected to be under-represented in the material. Due to these limitations, it is difficult to predict how the GRS would perform on a population level, and further studies exploring this question are indeed warranted. Moreover, some additional limitations of GRSs in general must be overcome before clinical implementation of the method would be possible, of which the most important is the current lack of equitable use of GRSs across ancestries and social groups.[89]

Similar to the prediction of SLE itself, an important question which remains to be answered is to what extent GRSs for specific SLE manifestations can be further improved. On this note, it is important to remember that the GRS examined in this thesis, was based on SNPs associated with SLE itself, rather than its subphenotypes. Provided that more knowledge is acquired about individual SNP-manifestation associations, future GRSs may be constructed with regards to each specific SLE manifestation, and it appears plausible that this would increase their predictive ability. So far, genetic variants identified as associated with SLE subphenotypes are few, and the construction of subphenotype-specific GRSs has proven difficult. In the largest study to date exploring the possibility of creating a LN-specific PRS, the statistical power was found to be too low in spite of 3000 patients being included in the discovery cohort, and neither of the PRSs constructed reached statistical significance in the replication cohort.[218] With these challenges, brought on both by the rarity of SLE and the small effect sizes of each SNP for the manifestations, increased international collaboration between clinical centres, enabling analysis of very large materials of well-characterized patients, may be necessary to enable a more precise analysis of the link between the SLE genetic risk and disease outcomes.

A different way to overcome these challenges, as demonstrated in this thesis, is to investigate associations in materials not limited to patients with SLE. By incorporating large, publicly available datasets in our analyses, we succeeded in re-constructing our original GRS to find associations with the ACR-82 criteria beyond nephritis and anti-dsDNA antibodies, specifically with arthritis, neurologic disorder and hematologic disorder. An important limitation of this study, is that the selected datasets from the FinnGen biobank included patients with a wide range of causes of their examined manifestation. For example, for the dataset selected to represent the ACR-82 nephritis criterion, glomerular disorders secondary to diseases such as amyloidosis, congenital syphilis, multiple myeloma and sepsis were included, some of which may have a very different aetiology compared with LN.[192] As the number of individuals included in publicly available datasets continue to grow, the phenotype data is likely to become more sophisticated, which may enable analysis of manifestations even more closely related to the subphenotypes of SLE.

In addition to methods which incorporate information from the general population, it is likely that GRSs for SLE could improve in several ways. In an interesting study exploring pathway specific PRSs in SLE, it was demonstrated that patients could be clustered into 4 different molecular subsets, and that specifically, patients with high T- and B-cell pathway PRSs displayed increased organ damage.[224] If further elaborated, this method may provide not only clinically relevant information, but also insights on the differences in the disease aetiology between patient groups. Additionally, there are a number of suggested uses for GRSs in SLE which remain largely unexplored, including stratification of patients for medical trials. Another interesting application of GRSs which has been suggested, is to identify patients more likely to carry monogenetic disease variants.[225] As demonstrated in this thesis, patients with an early onset of SLE have a higher GRS compared with patients with and onset later in life. In patients with early-onset SLE lacking this cumulative genetic risk factor, however, it may be speculated that the chance of finding a monogenetic cause of the disease is increased.

Although a high predictive accuracy for SLE itself and its renal manifestations may be attainable using a model based purely on the genetic risk, the heritability of other SLE manifestations, and thereby the maximal potential accuracy of a GRS for the manifestations, may be lower.[84] As exemplified by the interaction between smoking and the *STAT4* risk allele in this thesis, gene-environment interactions may further be responsible for the development of certain manifestations and complications of the disease. For these reasons, predictive models based purely on genetics may not be sufficient to predict all manifestations of SLE, and combined models based on both the genome and exposome may ultimately be required to accurately predict the disease course.

Gene-environment interactions remain poorly studied regarding the development of SLE, and even more so when it comes to specific manifestations, and the field thus warrants further investigation.

Moreover, as the extreme heterogeneity of SLE is seen not only between patients but over time in the same patient,[9] factors not relying solely on lifetime exposure may need to be combined with a PRS in order to achieve an optimal predictive ability. Recently, it was demonstrated that an epigenetic IFN score could predict the presence of the IFN signature in SS and was associated with several clinical phenotypes in patients with the disease.[226, 227] Scores combining the genetic risk with an epigenetic score may thus potentially provide a more optimal combination of both lifetime and current risk of manifestations.

Concluding remarks

SLE is a complex disease in which the clinical picture varies extremely between individuals. Understanding the underlying mechanisms that contribute to both common characteristics and the heterogeneity of the disease, is important for both clinical management and the development of new treatment strategies. This thesis highlights the importance of both genetic and clinical factors, as well as their interaction, for the prognosis of SLE. The results primarily support the idea that the genetic burden in SLE has an impact on the occurrence of more serious manifestations, and that the IL-12-STAT4 signalling pathway may be relevant for the development of cardiovascular complications in SLE. Through this thesis, the role of genetic factors in disease manifestations has been elucidated, thereby enhancing the overall understanding of the disease.

Sammanfattning på Svenska

Systemisk lupus erythematosus (SLE) är en kronisk, autoimmun sjukdom orsakad av både gener och miljöfaktorer. Personer med SLE uppvisar en extrem variation i sin symptombild, som kan vara allt ifrån mild till livshotande. Syftet med denna avhandling är att öka kunskapen om hur genetiska riskfaktorer för SLE, genom samverkan med varandra och med miljöfaktorer, bidrar till uppkomsten av sjukdomens olika symptom.

Trots att överlevnaden för personer med SLE har förbättrats, är risken för organskada och för tidig död fortfarande högre än hos normalbefolkningen. I **delstudie I** studerades utvecklandet av permanent organskada hos svenska individer med SLE. Studien visade att efter i genomsnitt 17 års sjukdomsduration, hade drygt hälften av patienterna utvecklat organskada. Flera kliniska riskfaktorer för organskada identifierades också i studien.

I nuläget saknas pålitliga markörer som kan förutse vilka patienter med SLE som kommer att uppvisa ett allvarligt respektive milt sjukdomsförlopp. I **delstudie II** undersöktes om den sammantagna effekten av 57 kända SLE-riskgener, var kopplad till sjukdomens symptom eller komplikationer. Hos de patienter som visade sig bära på en högre genetisk risk för SLE, fann vi att risken var högre att utveckla permanent organskada, njursvikt och kardiovaskulär sjukdom, jämfört med patienterna med en låg genetisk risk. Vi fann också att risken att avlida var kopplad till en högre genetisk risk. **I delstudie IV** undersöktes effekten av samma 57 riskgener i en stor grupp individer utan SLE. I denna grupp fann vi att en högre genetisk risk för SLE var associerat med utvecklandet av diagnoser som påminner om vissa SLE-manifestationer, såsom led- och lungsäcksinflammation. Sammantaget visar studierna att den sammanlagda genetiska risken för SLE kan utgöra en möjlig framtida markör för sjukdomens symptombild och prognos.

Det har föreslagits att rökning spelar en roll för utvecklandet av SLE, och det är dessutom välkänt att rökning ger upphov till hjärtinfarkt. I **delstudie III** kunde vi visa att sambandet mellan rökning och hjärtinfarkt förstärks hos patienter med SLE som bär på en känd SLE-riskvariant i *Signal Transducer and Activator of Transcription 4* (*STAT4*)-genen, med en mycket kraftig riskökning hos patienter som bar på två kopior av riskvarianten. Vi kunde också visa

att rökning är kopplat till högre nivåer av aktiverat STAT4 i immunceller från friska personer, vilket kan vara förklaringen till hur genen samverkar med rökning. Hjärtinfarkt förekommer i högre utsträckning bland patienter med SLE jämfört med normalbefolkningen, och denna studie belyser både en möjlig orsak till detta, samt vikten av att personer med SLE avstår från att röka.

Sammantaget understryker denna avhandling betydelsen av både genetiska och kliniska riskfaktorer, såväl som deras samverkan, för prognosen vid SLE.

Acknowledgements

The work in this thesis was carried out at the Department of Medical Sciences, Faculty of Medicine, Uppsala University. During my PhD studies, I have been fortunate to have many experienced, inspiring and supportive individuals around me, and I am deeply thankful to each one of you. Especially, I would like to thank:

The *patients* and *control individuals*. My thesis would not exist without you.

Dag Leonard, my main supervisor. I am truly grateful for your guidance throughout these years, for your kindness, consistent optimism and encouragement to aim high, and for always being available and having time for my questions. Your intellect and excellent ability to see things from a new perspective, have been invaluable in making this work possible. Thank you.

Lars Rönnblom, my co-supervisor. Your guidance on each of my projects and through the world of science has been invaluable, your own work in the field an inspiration, and your humility, wealth of knowledge and knack for storytelling have made every interaction with you an enlightening experience.

All past and current members of the *Uppsala* group who have made this journey significantly more enjoyable. *Johanna Sandling*, for your intelligent comments and constructive critique on each of my projects. *Ahmed Sayadi*, for all your help with the genetic data and especially your contribution to the MR paper. *Andrei Alexsson*, for all the time you spent teaching me about statistics, genetics and especially genetic risk scores in the beginning. *Pascal Pucholt*, for your encouragement and hours of help with bioinformatics and R. *Maija-Leena Eloranta***,** for your kindness and support in anything from academic advice to administrative tasks or simple office chores. *Karin Hjorton*, for your guidance in the clinic and for opening up the opportunity for me as a junior doctor, a role I am thrilled to embrace. *Niklas Hagberg*, for your valuable contribution to, and all the discussions regarding, the STAT4 paper. *Alina Johansson*, *Cane Amcoff*, *Christian Lundtoft, Cecilia Fugmann*, *Gunnel Nordmark*, *Juliana Imgenberg-Kreutz*, *Karin Bolin, Neha Behare, Nina Oparina*, *Ioanna Giannakou, Paulina Turesson and Sule Yavuz*, for many interesting discussions and great company during breaks, dinners and conferences.

Each of the co-authors on my papers, not already mentioned. *Christoffer Sjöwall*, your valuable comments on each of my papers have led to substantial improvements of my work. *Martina Frodlund*, as a first author of one of the papers in this thesis you have had an important role for my research journey. Thank you both for a great collaboration, and for nice company during conferences and meetings. *Elisabet Svenungsson* and *Iva Gunnarsson* at *Karolinska Institute*, your feedback and suggestions have consistently improved my papers. Although not a co-author, the same goes for *Lina Diaz-Gallo* − your feedback and intriguing questions on the exclusion of the HLA region during meetings, as well as your own work, have kept me on my toes and widened my understanding of SLE genetics. *Anders Bengtsson* and *Andreas Jönsen,* Lund; *Anna Rudin*, Gothenburg; *Karoline Lerang*, Oslo; *Solbritt Rantapää-Dahlqvist*, Umeå; *Søren Jacobsen*, Copenhagen; *Anne Troldborg*, Aarhus; and *Anne Voss*, Odense, your generous sharing of data and important comments on my papers have been invaluable for the work in this thesis. *Tim Vyse* and *David Morris* at King's College, London. Tim, your critique on the organ damage paper and contribution of data for the validation cohort helped to overcome a major obstacle in my research. David, thank you for assembling and sharing the genetic data for the validation cohort.

Tove Fall and *Åsa Johansson*, for your very valuable constructive critique during my halftime seminar. Thank you Tove also for the later advise on the MR analyses.

All doctors, nurses, assistant nurses and secretaries at the *Rheumatology clinic at Akademiska Sjukhuset, Uppsala*, for all your support during my time as a junior doctor in the clinic. A special thank you to *Rezvan Kiani Dekhordi*, research nurse, for all your help with the patient samples.

Yvonne Enman. I have really appreciated both our collaboration and the time we have spent together during network meetings. Your strength, wisdom and real-life insights about living with SLE have taught me many important lessons which I will take with me throughout my medical career.

My family. *Anthony*, *Ethan* and *Alba*, your presence in my life brings me endless joy and strength. I am so grateful for your love and support. *Els-Britt*, my lifelong mentor; *Rolf*, who is always there for me and the children; and *Elizabeth***,** whose memory warmly illuminates my path. *Lewis*, your resilience has been an inspiring symbol of strength throughout my journey. *Carolina*, you occupy a cherished place in my childhood memories and in my heart. *Louise*, growing up, you were like a sister to me. I hope we get to spend more time together in the future.

All my cherished friends.

Lastly, there is one person in this acknowledgement who has impacted my journey in an inexplicable way – my late sister-in-law *Lisette*. It was a strange twist of fate when I learned that the rare disease that claimed her life was the very one I had dedicated my research to, SLE. This unexpected link in our family's story brought a newfound perspective and a personal dimension to my research. As if the universe intended for me to fight the fight that she could not. Her memory will continue to fuel my passion to unravel the mysteries of this disease.

This work was supported with grants from the Swedish Research Council for Medicine and Health, Swedish Society for Medical Research, Knut and Alice Wallenberg Foundation, Swedish Society of Medicine and Ingegerd Johansson donation, Swedish Rheumatism Foundation, King Gustaf V's 80-year Foundation, Uppsala County Council and Uppsala University Hospital (ALF), Agnes and Mac Rudberg Foundation, Åke Wiberg's foundation, Gustafsson's foundation, Selander's foundation and Gustaf Prim Foundation.

References

1. Herrmann, M., et al., *Impaired phagocytosis of apoptotic cell material by monocyte-derived macrophages from patients with systemic lupus erythematosus.* Arthritis Rheum, 1998. **41**(7): p. 1241-50.
2. Hooks, J.J., et al., *Immune interferon in the circulation of patients with autoimmune disease.* N Engl J Med, 1979. **301**(1): p. 5-8.
3. Ytterberg, S.R. and T.J. Schnitzer, *Serum interferon levels in patients with systemic lupus erythematosus.* Arthritis Rheum, 1982. **25**(4): p. 401-6.
4. Rawlings, D.J., et al., *Altered B cell signalling in autoimmunity.* Nat Rev Immunol, 2017. **17**(7): p. 421-436.
5. Katsuyama, T., G.C. Tsokos, and V.R. Moulton, *Aberrant T Cell Signaling and Subsets in Systemic Lupus Erythematosus.* Front Immunol, 2018. **9**: p. 1088.
6. Ghodke-Puranik, Y. and T.B. Niewold, *Immunogenetics of systemic lupus erythematosus: A comprehensive review.* J Autoimmun, 2015. **64**: p. 125-36.
7. Kuhn, A., et al., *The Diagnosis and Treatment of Systemic Lupus Erythematosus.* Dtsch Arztebl Int, 2015. **112**(25): p. 423-32.
8. Arbuckle, M.R., et al., *Development of autoantibodies before the clinical onset of systemic lupus erythematosus.* N Engl J Med, 2003. **349**(16): p. 1526-33.
9. Bengtsson, A.A. and L. Ronnblom, *Systemic lupus erythematosus: still a challenge for physicians.* J Intern Med, 2017. **281**(1): p. 52-64.
10. Danchenko, N., J.A. Satia, and M.S. Anthony, *Epidemiology of systemic lupus erythematosus: a comparison of worldwide disease burden.* Lupus, 2006. **15**(5): p. 308-18.
11. Osio-Salido, E. and H. Manapat-Reyes, *Epidemiology of systemic lupus erythematosus in Asia.* Lupus, 2010. **19**(12): p. 1365-73.
12. Pons-Estel, G.J., M.F. Ugarte-Gil, and G.S. Alarcon, *Epidemiology of systemic lupus erythematosus.* Expert Rev Clin Immunol, 2017. **13**(8): p. 799-814.
13. Morris, D.L., et al., *Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus.* Nat Genet, 2016. **48**(8): p. 940-946.
14. Symmons, D.P., *Frequency of lupus in people of African origin.* Lupus, 1995. **4**(3): p. 176-8.
15. Stojan, G. and M. Petri, *Epidemiology of systemic lupus erythematosus: an update.* Curr Opin Rheumatol, 2018. **30**(2): p. 144-150.
16. Housey, M., et al., *Incidence and prevalence of systemic lupus erythematosus among Arab and Chaldean Americans in southeastern Michigan: the Michigan Lupus Epidemiology and Surveillance Program.* Am J Public Health, 2015. **105**(5): p. e74-9.
17. Rees, F., et al., *The incidence and prevalence of systemic lupus erythematosus in the UK, 1999-2012.* Ann Rheum Dis, 2016. **75**(1): p. 136-41.

18. Voss, A., A. Green, and P. Junker, *Systemic lupus erythematosus in Denmark: clinical and epidemiological characterization of a county-based cohort.* Scand J Rheumatol, 1998. **27**(2): p. 98-105.
19. Nossent, H.C., *Systemic lupus erythematosus in the Arctic region of Norway.* J Rheumatol, 2001. **28**(3): p. 539-46.
20. Lerang, K., et al., *High incidence and prevalence of systemic lupus erythematosus in Norway.* Lupus, 2012. **21**(12): p. 1362-9.
21. Jonsson, H., et al., *Estimating the incidence of systemic lupus erythematosus in a defined population using multiple sources of retrieval.* Br J Rheumatol, 1990. **29**(3): p. 185-8.
22. Stahl-Hallengren, C., et al., *Incidence studies of systemic lupus erythematosus in Southern Sweden: increasing age, decreasing frequency of renal manifestations and good prognosis.* J Rheumatol, 2000. **27**(3): p. 685-91.
23. Simard, J.F., et al., *Systemic lupus erythematosus prevalence in Sweden in 2010: what do national registers say?* Arthritis Care Res (Hoboken), 2014. **66**(11): p. 1710-7.
24. Ingvarsson, R.F., A.A. Bengtsson, and A. Jonsen, *Variations in the epidemiology of systemic lupus erythematosus in southern Sweden.* Lupus, 2016. **25**(7): p. 772-80.
25. Tselios, K., et al., *Disease course patterns in systemic lupus erythematosus.* Lupus, 2019. **28**(1): p. 114-122.
26. Fries, J.F. and H.R. Holman, *Systemic lupus erythematosus: a clinical analysis.* Major Probl Intern Med, 1975. **6**: p. v-199.
27. Huang, X., et al., *A Contemporary Update on the Diagnosis of Systemic Lupus Erythematosus.* Clin Rev Allergy Immunol, 2022.
28. Solomon, D.H., et al., *Evidence-based guidelines for the use of immunologic tests: antinuclear antibody testing.* Arthritis Rheum, 2002. **47**(4): p. 434-44.
29. Aggarwal, R., et al., *Distinctions between diagnostic and classification criteria?* Arthritis Care Res (Hoboken), 2015. **67**(7): p. 891-7.
30. Tan, E.M., et al., *The 1982 revised criteria for the classification of systemic lupus erythematosus.* Arthritis Rheum, 1982. **25**(11): p. 1271-7.
31. Hochberg, M.C., *Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus.* Arthritis Rheum, 1997. **40**(9): p. 1725.
32. Petri, M. and L. Magder, *Classification criteria for systemic lupus erythematosus: a review.* Lupus, 2004. **13**(11): p. 829-37.
33. Petri, M., et al., *Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus.* Arthritis Rheum, 2012. **64**(8): p. 2677-86.
34. Petri, M., et al., *Predictors of organ damage in systemic lupus erythematosus: the Hopkins Lupus Cohort.* Arthritis Rheum, 2012. **64**(12): p. 4021-8.
35. Hartman, E.A.R., et al., *Performance of the 2012 Systemic Lupus International Collaborating Clinics classification criteria versus the 1997 American College of Rheumatology classification criteria in adult and juvenile systemic lupus erythematosus. A systematic review and meta-analysis.* Autoimmun Rev, 2018. **17**(3): p. 316-322.
36. Aringer, M., et al., *2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus.* Ann Rheum Dis, 2019. **78**(9): p. 1151-1159.

37. Diaz-Gallo, L.M., et al., *Four Systemic Lupus Erythematosus Subgroups, Defined by Autoantibodies Status, Differ Regarding HLA-DRB1 Genotype Associations and Immunological and Clinical Manifestations.* ACR Open Rheumatol, 2022. **4**(1): p. 27-39.
38. Idborg, H., et al., *Two subgroups in systemic lupus erythematosus with features of antiphospholipid or Sjogren's syndrome differ in molecular signatures and treatment perspectives.* Arthritis Res Ther, 2019. **21**(1): p. 62.
39. Decker, P., et al., *An updated review of anti-Ro52 (TRIM21) antibodies impact in connective tissue diseases clinical management.* Autoimmun Rev, 2022. **21**(3): p. 103013.
40. Miyakis, S., et al., *International consensus statement on an update of the classification criteria for definite antiphospholipid syndrome (APS).* J Thromb Haemost, 2006. **4**(2): p. 295-306.
41. Bertolaccini, M.L., et al., *'Non-criteria' aPL tests: report of a task force and preconference workshop at the 13th International Congress on Antiphospholipid Antibodies, Galveston, TX, USA, April 2010.* Lupus, 2011. **20**(2): p. 191-205.
42. Ruiz-Irastorza, G., et al., *High impact of antiphospholipid syndrome on irreversible organ damage and survival of patients with systemic lupus erythematosus.* Arch Intern Med, 2004. **164**(1): p. 77-82.
43. Pericleous, C., et al., *Antiphospholipid antibody levels in early systemic lupus erythematosus: are they associated with subsequent mortality and vascular events?* Rheumatology (Oxford), 2020. **59**(1): p. 146-152.
44. Smeenk, R., et al., *Antibodies to DNA in patients with systemic lupus erythematosus. Their role in the diagnosis, the follow-up and the pathogenesis of the disease.* Clin Rheumatol, 1990. **9**(1 Suppl 1): p. 100-10.
45. Tektonidou, M.G., et al., *Survival in adults and children with systemic lupus erythematosus: a systematic review and Bayesian meta-analysis of studies from 1950 to 2016.* Ann Rheum Dis, 2017. **76**(12): p. 2009-2016.
46. Jorge, A.M., et al., *Unchanging premature mortality trends in systemic lupus erythematosus: a general population-based study (1999-2014).* Rheumatology (Oxford), 2018. **57**(2): p. 337-344.
47. Bjornadal, L., et al., *Cardiovascular disease a hazard despite improved prognosis in patients with systemic lupus erythematosus: results from a Swedish population based study 1964-95.* J Rheumatol, 2004. **31**(4): p. 713-9.
48. Gustafsson, J.T., et al., *Risk factors for cardiovascular mortality in patients with systemic lupus erythematosus, a prospective cohort study.* Arthritis Res Ther, 2012. **14**(2): p. R46.
49. Ingvarsson, R.F., et al., *Good survival rates in systemic lupus erythematosus in southern Sweden, while the mortality rate remains increased compared with the population.* Lupus, 2019. **28**(12): p. 1488-1494.
50. McMahon, M. and B. Skaggs, *Pathogenesis and treatment of atherosclerosis in lupus.* Rheum Dis Clin North Am, 2014. **40**(3): p. 475-95, viii.
51. Hsu, C.Y., et al., *Cumulative immunosuppressant exposure is associated with diversified cancer risk among 14 832 patients with systemic lupus erythematosus: a nested case-control study.* Rheumatology (Oxford), 2017. **56**(4): p. 620-628.
52. Lewis, M.J. and A.S. Jawad, *The effect of ethnicity and genetic ancestry on the epidemiology, clinical features and outcome of systemic lupus erythematosus.* Rheumatology (Oxford), 2017. **56**(suppl_1): p. i67-i77.
53. Gladman, D., et al., *The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology*

damage index for systemic lupus erythematosus. Arthritis Rheum, 1996. **39**(3): p. 363-9.
54. Alarcon, G.S., et al., *Systemic lupus erythematosus in three ethnic groups. XX. Damage as a predictor of further damage.* Rheumatology (Oxford), 2004. **43**(2): p. 202-5.
55. Francula-Zaninovic, S. and I.A. Nola, *Management of Measurable Variable Cardiovascular Disease' Risk Factors.* Curr Cardiol Rev, 2018. **14**(3): p. 153-163.
56. Lu, X., et al., *Patients with systemic lupus erythematosus face a high risk of cardiovascular disease: A systematic review and Meta-analysis.* Int Immunopharmacol, 2021. **94**: p. 107466.
57. Bello, N., et al., *Cardiovascular events and risk in patients with systemic lupus erythematosus: Systematic literature review and meta-analysis.* Lupus, 2023. **32**(3): p. 325-341.
58. Schoenfeld, S.R., S. Kasturi, and K.H. Costenbader, *The epidemiology of atherosclerotic cardiovascular disease among patients with SLE: a systematic review.* Semin Arthritis Rheum, 2013. **43**(1): p. 77-95.
59. Tselios, K., et al., *All-cause, cause-specific and age-specific standardised mortality ratios of patients with systemic lupus erythematosus in Ontario, Canada over 43 years (1971-2013).* Ann Rheum Dis, 2019. **78**(6): p. 802-806.
60. Esdaile, J.M., et al., *Traditional Framingham risk factors fail to fully account for accelerated atherosclerosis in systemic lupus erythematosus.* Arthritis Rheum, 2001. **44**(10): p. 2331-7.
61. Boulos, D., et al., *Cardiovascular risk profiles in a lupus cohort: what do different calculators tell us?* Lupus Sci Med, 2017. **4**(1): p. e000212.
62. McMahon, M., R. Seto, and B.J. Skaggs, *Cardiovascular disease in systemic lupus erythematosus.* Rheumatol Immunol Res, 2021. **2**(3): p. 157-172.
63. Tyden, H., et al., *Endothelial dysfunction is associated with activation of the type I interferon system and platelets in patients with systemic lupus erythematosus.* RMD Open, 2017. **3**(2): p. e000508.
64. Hanly, J.G., et al., *The frequency and outcome of lupus nephritis: results from an international inception cohort study.* Rheumatology (Oxford), 2016. **55**(2): p. 252-62.
65. Parikh, S.V., et al., *Update on Lupus Nephritis: Core Curriculum 2020.* Am J Kidney Dis, 2020. **76**(2): p. 265-281.
66. Tanaka, Y., et al., *Systemic Lupus Erythematosus: Targeted Literature Review of the Epidemiology, Current Treatment, and Disease Burden in the Asia Pacific Region.* Arthritis Care Res (Hoboken), 2022. **74**(2): p. 187-198.
67. Mok, C.C., *Towards new avenues in the management of lupus glomerulonephritis.* Nat Rev Rheumatol, 2016. **12**(4): p. 221-34.
68. Kandane-Rathnayake, R., et al., *Longitudinal associations of active renal disease with irreversible organ damage accrual in systemic lupus erythematosus.* Lupus, 2019. **28**(14): p. 1669-1677.
69. Galindo-Izquierdo, M., et al., *Characterization of Patients With Lupus Nephritis Included in a Large Cohort From the Spanish Society of Rheumatology Registry of Patients With Systemic Lupus Erythematosus (RELESSER).* Medicine (Baltimore), 2016. **95**(9): p. e2891.
70. Churg J, B.J., Glassrock R., *Renal disease: classification and atlas of glomerular diseases.* New York: Igaku-Shoin, 1995(2 ed.): p. 151-5.
71. Chow, T.K., L.M. Looi, and P.L. Cheah, *A comparison of 1995 WHO classification with 2003 ISN/RPS classification of lupus nephritis: a single centre observation.* Malays J Pathol, 2015. **37**(3): p. 239-46.

72. Jackson, M., et al., *The genetic basis of disease.* Essays Biochem, 2018. **62**(5): p. 643-723.
73. Condo, I., *Rare Monogenic Diseases: Molecular Pathophysiology and Novel Therapies.* Int J Mol Sci, 2022. **23**(12).
74. Jorde, L.B.e.a., *Medical genetics, 4th ed.* 2010: Elsevier Mosby. 350.
75. Mayhew, A.J. and D. Meyre, *Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities.* Curr Genomics, 2017. **18**(4): p. 332-340.
76. Dawn Teare, M. and J.H. Barrett, *Genetic linkage studies.* Lancet, 2005. **366**(9490): p. 1036-44.
77. Brian D. Modena, A.D., Parth Patel, Varshini Sathish,, *Chapter 18 - Leveraging genomics to uncover the genetic, environmental and age-related factors leading to asthma*. Genomic and Precision Medicine (Third Edition). 2019: Academic Press.
78. Jahad Alghamdi, S.P., *Chapter 12 - Fundamentals of Complex Trait Genetics and Association Studies*. Handbook of Pharmacogenomics and Stratified Medicine. 2014: Academic Press.
79. Tam, V., et al., *Benefits and limitations of genome-wide association studies.* Nat Rev Genet, 2019. **20**(8): p. 467-484.
80. MacArthur, J., et al., *The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog).* Nucleic Acids Res, 2017. **45**(D1): p. D896-D901.
81. McCarthy, M.I., et al., *Genome-wide association studies for complex traits: consensus, uncertainty and challenges.* Nat Rev Genet, 2008. **9**(5): p. 356-69.
82. Iles, M.M., *What can genome-wide association studies tell us about the genetics of common disease?* PLoS Genet, 2008. **4**(2): p. e33.
83. Wray, N.R., M.E. Goddard, and P.M. Visscher, *Prediction of individual genetic risk to disease from genome-wide association studies.* Genome Res, 2007. **17**(10): p. 1520-8.
84. Wray, N.R., et al., *The genetic interpretation of area under the ROC curve in genomic profiling.* PLoS Genet, 2010. **6**(2): p. e1000864.
85. Janssens, A.C., et al., *Predictive testing for complex diseases using multiple genes: fact or fiction?* Genet Med, 2006. **8**(7): p. 395-400.
86. Nelson, M.R., et al., *The support of human genetic evidence for approved drug indications.* Nat Genet, 2015. **47**(8): p. 856-60.
87. King, E.A., J.W. Davis, and J.F. Degner, *Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval.* PLoS Genet, 2019. **15**(12): p. e1008489.
88. Demkova, K., D.L. Morris, and T.J. Vyse, *Genetics of SLE: does this explain susceptibility and severity across racial groups?* Rheumatology (Oxford), 2023. **62**(Suppl 1): p. i15-i21.
89. Igo, R.P., Jr., T.G. Kinzy, and J.N. Cooke Bailey, *Genetic Risk Scores.* Curr Protoc Hum Genet, 2019. **104**(1): p. e95.
90. Mars, N., et al., *Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers.* Nat Med, 2020. **26**(4): p. 549-557.
91. Khera, A.V., et al., *Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations.* Nat Genet, 2018. **50**(9): p. 1219-1224.

92. Mahajan, A., et al., *Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps.* Nat Genet, 2018. **50**(11): p. 1505-1513.
93. Inouye, M., et al., *Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention.* J Am Coll Cardiol, 2018. **72**(16): p. 1883-1893.
94. Riveros-Mckay, F., et al., *Integrated Polygenic Tool Substantially Enhances Coronary Artery Disease Prediction.* Circ Genom Precis Med, 2021. **14**(2): p. e003304.
95. Lambert, S.A., G. Abraham, and M. Inouye, *Towards clinical utility of polygenic risk scores.* Hum Mol Genet, 2019. **28**(R2): p. R133-R142.
96. Klarin, D., et al., *Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease.* Nat Genet, 2019. **51**(11): p. 1574-1579.
97. Collister, J.A., X. Liu, and L. Clifton, *Calculating Polygenic Risk Scores (PRS) in UK Biobank: A Practical Guide for Epidemiologists.* Front Genet, 2022. **13**: p. 818574.
98. Duncan, L., et al., *Analysis of polygenic risk score usage and performance in diverse human populations.* Nat Commun, 2019. **10**(1): p. 3328.
99. Alarcon-Segovia, D., et al., *Familial aggregation of systemic lupus erythematosus, rheumatoid arthritis, and other autoimmune diseases in 1,177 lupus patients from the GLADEL cohort.* Arthritis Rheum, 2005. **52**(4): p. 1138-47.
100. Deapen, D., et al., *A revised estimate of twin concordance in systemic lupus erythematosus.* Arthritis Rheum, 1992. **35**(3): p. 311-8.
101. Chen, L., D.L. Morris, and T.J. Vyse, *Genetic advances in systemic lupus erythematosus: an update.* Curr Opin Rheumatol, 2017. **29**(5): p. 423-433.
102. Cheng, M.H. and M.S. Anderson, *Monogenic autoimmunity.* Annu Rev Immunol, 2012. **30**: p. 393-427.
103. Bentham, J., et al., *Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus.* Nat Genet, 2015. **47**(12): p. 1457-1464.
104. Demirci, F.Y., et al., *Identification of a New Susceptibility Locus for Systemic Lupus Erythematosus on Chromosome 12 in Individuals of European Ancestry.* Arthritis Rheumatol, 2016. **68**(1): p. 174-83.
105. Graham, R.R., et al., *Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus.* Nat Genet, 2008. **40**(9): p. 1059-61.
106. Han, J.W., et al., *Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus.* Nat Genet, 2009. **41**(11): p. 1234-7.
107. Hom, G., et al., *Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX.* N Engl J Med, 2008. **358**(9): p. 900-9.
108. Julia, A., et al., *Genome-wide association study meta-analysis identifies five new loci for systemic lupus erythematosus.* Arthritis Res Ther, 2018. **20**(1): p. 100.
109. Kozyrev, S.V., et al., *Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus.* Nat Genet, 2008. **40**(2): p. 211-6.
110. Langefeld, C.D., et al., *Transancestral mapping and genetic load in systemic lupus erythematosus.* Nat Commun, 2017. **8**: p. 16021.
111. Lee, H.S., et al., *Ethnic specificity of lupus-associated loci identified in a genome-wide association study in Korean women.* Ann Rheum Dis, 2014. **73**(6): p. 1240-5.

112. Lessard, C.J., et al., *Identification of a Systemic Lupus Erythematosus Risk Locus Spanning ATG16L2, FCHSD2, and P2RY2 in Koreans.* Arthritis Rheumatol, 2016. **68**(5): p. 1197-1209.
113. Molineros, J.E., et al., *Confirmation of five novel susceptibility loci for systemic lupus erythematosus (SLE) and integrated network analysis of 82 SLE susceptibility loci.* Hum Mol Genet, 2017. **26**(6): p. 1205-1216.
114. Morris, D.L., 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. **91**(5): p. 778-93.
115. Okada, Y., et al., *A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus eyrthematosus in Japanese.* PLoS Genet, 2012. **8**(1): p. e1002455.
116. Song, Q., et al., *Genome-wide association study on Northern Chinese identifies KLF2, DOT1L and STAB2 associated with systemic lupus erythematosus.* Rheumatology (Oxford), 2021. **60**(9): p. 4407-4417.
117. Sun, C., et al., *High-density genotyping of immune-related loci identifies new SLE risk variants in individuals with Asian ancestry.* Nat Genet, 2016. **48**(3): p. 323-30.
118. Tangtanatakul, P., et al., *Meta-analysis of genome-wide association study identifies FBN2 as a novel locus associated with systemic lupus erythematosus in Thai population.* Arthritis Res Ther, 2020. **22**(1): p. 185.
119. Wang, Y.F., et al., *Multi-ancestral GWAS identifies shared and Asian-specific loci for SLE and links type III interferon signaling and lysosomal function to the disease.* Arthritis Rheumatol, 2021.
120. Wang, Y.F., et al., *Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups.* Nat Commun, 2021. **12**(1): p. 772.
121. Zhang, F., et al., *Independent Replication on Genome-Wide Association Study Signals Identifies IRF3 as a Novel Locus for Systemic Lupus Erythematosus.* Front Genet, 2020. **11**: p. 600.
122. Zhang, H., et al., *Meta-analysis of GWAS on both Chinese and European populations identifies GPR173 as a novel X chromosome susceptibility gene for SLE.* Arthritis Res Ther, 2018. **20**(1): p. 92.
123. Zhang, J., et al., *Gene-based meta-analysis of genome-wide association study data identifies independent single-nucleotide polymorphisms in ANXA6 as being associated with systemic lupus erythematosus in Asian populations.* Arthritis Rheumatol, 2015. **67**(11): p. 2966-77.
124. Harley, I.T.W. and A.H. Sawalha, *Systemic lupus erythematosus as a genetic disease.* Clin Immunol, 2022. **236**: p. 108953.
125. Hagberg, N., C. Lundtoft, and L. Ronnblom, *Immunogenetics in systemic lupus erythematosus: Transitioning from genetic associations to cellular effects.* Scand J Immunol, 2020. **92**(4): p. e12894.
126. Gullstrand, B., et al., *Complement classical pathway components are all important in clearance of apoptotic and secondary necrotic cells.* Clin Exp Immunol, 2009. **156**(2): p. 303-11.
127. Yang, F., et al., *Programmed Cell Death Pathways in the Pathogenesis of Systemic Lupus Erythematosus.* J Immunol Res, 2019. **2019**: p. 3638562.
128. Tsokos, G.C., et al., *New insights into the immunopathogenesis of systemic lupus erythematosus.* Nat Rev Rheumatol, 2016. **12**(12): p. 716-730.
129. Schejbel, L., et al., *Molecular basis of hereditary C1q deficiency--revisited: identification of several novel disease-causing mutations.* Genes Immun, 2011. **12**(8): p. 626-34.

130. Deng, Y. and B.P. Tsao, *Updates in Lupus Genetics.* Curr Rheumatol Rep, 2017. **19**(11): p. 68.
131. Ma, K., et al., *Multiple Functions of B Cells in the Pathogenesis of Systemic Lupus Erythematosus.* Int J Mol Sci, 2019. **20**(23).
132. Tenbrock, K. and T. Rauen, *T cell dysregulation in SLE.* Clin Immunol, 2022. **239**: p. 109031.
133. Li, H., et al., *Abnormalities of T cells in systemic lupus erythematosus: new insights in pathogenesis and therapeutic strategies.* J Autoimmun, 2022. **132**: p. 102870.
134. Ronnblom, L. and D. Leonard, *Interferon pathway in SLE: one key to unlocking the mystery of the disease.* Lupus Sci Med, 2019. **6**(1): p. e000270.
135. Ronnblom, L.E., G.V. Alm, and K.E. Oberg, *Possible induction of systemic lupus erythematosus by interferon-alpha treatment in a patient with a malignant carcinoid tumour.* J Intern Med, 1990. **227**(3): p. 207-10.
136. Niewold, T.B., *Interferon alpha as a primary pathogenic factor in human lupus.* J Interferon Cytokine Res, 2011. **31**(12): p. 887-92.
137. Niewold, T.B., et al., *High serum IFN-alpha activity is a heritable risk factor for systemic lupus erythematosus.* Genes Immun, 2007. **8**(6): p. 492-502.
138. Ghodke-Puranik, Y. and T.B. Niewold, *Genetics of the type I interferon pathway in systemic lupus erythematosus.* Int J Clin Rheumtol, 2013. **8**(6).
139. Gallucci, S., S. Meka, and A.M. Gamero, *Abnormalities of the type I interferon signaling pathway in lupus autoimmunity.* Cytokine, 2021. **146**: p. 155633.
140. Aragam, K.G., et al., *Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants.* Nat Genet, 2022. **54**(12): p. 1803-1815.
141. Brandes, N., N. Linial, and M. Linial, *Genetic association studies of alterations in protein function expose recessive effects on cancer predisposition.* Sci Rep, 2021. **11**(1): p. 14901.
142. Mishra, A., et al., *Stroke genetics informs drug discovery and risk prediction across ancestries.* Nature, 2022. **611**(7934): p. 115-123.
143. Pividori, M., et al., *Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide studies.* Lancet Respir Med, 2019. **7**(6): p. 509-522.
144. Wuttke, M., et al., *A catalog of genetic loci associated with kidney function from analyses of a million individuals.* Nat Genet, 2019. **51**(6): p. 957-972.
145. Saunders, G.R.B., et al., *Genetic diversity fuels gene discovery for tobacco and alcohol use.* Nature, 2022. **612**(7941): p. 720-724.
146. Ishigaki, K., et al., *Multi-ancestry genome-wide association analyses identify novel genetic mechanisms in rheumatoid arthritis.* Nat Genet, 2022. **54**(11): p. 1640-1651.
147. Cordell, H.J., et al., *An international genome-wide meta-analysis of primary biliary cholangitis: Novel risk loci and candidate drugs.* J Hepatol, 2021. **75**(3): p. 572-581.
148. Lopez-Isac, E., et al., *Combined genetic analysis of juvenile idiopathic arthritis clinical subtypes identifies novel risk loci, target genes and key regulatory mechanisms.* Ann Rheum Dis, 2021. **80**(3): p. 321-328.
149. Li, Y., et al., *A genome-wide association study in Han Chinese identifies a susceptibility locus for primary Sjogren's syndrome at 7q11.23.* Nat Genet, 2013. **45**(11): p. 1361-5.
150. International Multiple Sclerosis Genetics, C., *Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility.* Science, 2019. **365**(6460).

151. Liu, J.Z., et al., *Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations.* Nat Genet, 2015. **47**(9): p. 979-986.
152. Dudding, T., et al., *Genome wide analysis for mouth ulcers identifies associations at immune regulatory loci.* Nat Commun, 2019. **10**(1): p. 1052.
153. Taylor, K.E., et al., *Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus.* PLoS Genet, 2008. **4**(5): p. e1000084.
154. Bolin, K., et al., *Association of STAT4 polymorphism with severe renal insufficiency in lupus nephritis.* PLoS One, 2013. **8**(12): p. e84450.
155. Svenungsson, E., et al., *A STAT4 risk allele is associated with ischaemic cerebrovascular events and anti-phospholipid antibodies in systemic lupus erythematosus.* Ann Rheum Dis, 2010. **69**(5): p. 834-40.
156. Restivo, V., et al., *Systematic review and meta-analysis of cardiovascular risk in rheumatological disease: Symptomatic and non-symptomatic events in rheumatoid arthritis and systemic lupus erythematosus.* Autoimmun Rev, 2022. **21**(1): p. 102925.
157. Sigurdsson, S., et al., *A risk haplotype of STAT4 for systemic lupus erythematosus is over-expressed, correlates with anti-dsDNA and shows additive effects with two risk alleles of IRF5.* Hum Mol Genet, 2008. **17**(18): p. 2868-76.
158. Song, K., et al., *An update on genetic susceptibility in lupus nephritis.* Clin Immunol, 2020. **210**: p. 108272.
159. Jarvinen, T.M., et al., *Polymorphisms of the ITGAM gene confer higher risk of discoid cutaneous than of systemic lupus erythematosus.* PLoS One, 2010. **5**(12): p. e14212.
160. Kim-Howard, X., et al., *ITGAM coding variant (rs1143679) influences the risk of renal disease, discoid rash and immunological manifestations in patients with systemic lupus erythematosus with European ancestry.* Ann Rheum Dis, 2010. **69**(7): p. 1329-32.
161. Leonard, D., et al., *Coronary heart disease in systemic lupus erythematosus is associated with interferon regulatory factor-8 gene variants.* Circ Cardiovasc Genet, 2013. **6**(3): p. 255-63.
162. Li, X., et al., *Occupational risk factors for systemic lupus erythematosus: a nationwide study based on hospitalizations in Sweden.* J Rheumatol, 2012. **39**(4): p. 743-51.
163. Pollard, K.M., *Silica, Silicosis, and Autoimmunity.* Front Immunol, 2016. **7**: p. 97.
164. Parks, C.G. and G.S. Cooper, *Occupational exposures and risk of systemic lupus erythematosus: a review of the evidence and exposure assessment methods in population- and clinic-based studies.* Lupus, 2006. **15**(11): p. 728-36.
165. Cozier, Y.C., et al., *Relationship of Cigarette Smoking and Alcohol Consumption to Incidence of Systemic Lupus Erythematosus in a Prospective Cohort Study of Black Women.* Arthritis Care Res (Hoboken), 2019. **71**(5): p. 671-677.
166. Barbhaiya, M., et al., *Cigarette smoking and the risk of systemic lupus erythematosus, overall and by anti-double stranded DNA antibody subtype, in the Nurses' Health Study cohorts.* Ann Rheum Dis, 2018. **77**(2): p. 196-202.
167. Jiang, F., S. Li, and C. Jia, *Smoking and the risk of systemic lupus erythematosus: an updated systematic review and cumulative meta-analysis.* Clin Rheumatol, 2015. **34**(11): p. 1885-92.
168. He, Y. and A.H. Sawalha, *Drug-induced lupus erythematosus: an update on drugs and mechanisms.* Curr Opin Rheumatol, 2018. **30**(5): p. 490-497.

169. Zamansky, G.B., *Sunlight-induced pathogenesis in systemic lupus erythematosus.* J Invest Dermatol, 1985. **85**(3): p. 179-80.
170. Jog, N.R. and J.A. James, *Epstein Barr Virus and Autoimmune Responses in Systemic Lupus Erythematosus.* Front Immunol, 2020. **11**: p. 623944.
171. Hahn, J., et al., *Vitamin D and marine omega 3 fatty acid supplementation and incident autoimmune disease: VITAL randomized controlled trial.* BMJ, 2022. **376**: p. e066452.
172. Sanchez-Guerrero, J., et al., *Past use of oral contraceptives and the risk of developing systemic lupus erythematosus.* Arthritis Rheum, 1997. **40**(5): p. 804-8.
173. Ha, E., S.C. Bae, and K. Kim, *Recent advances in understanding the genetic basis of systemic lupus erythematosus.* Semin Immunopathol, 2022. **44**(1): p. 29-46.
174. Ortona, E., et al., *Sex-based differences in autoimmune diseases.* Ann Ist Super Sanita, 2016. **52**(2): p. 205-12.
175. Armstrong, D.L., et al., *GWAS identifies novel SLE susceptibility genes and explains the association of the HLA region.* Genes Immun, 2014. **15**(6): p. 347-54.
176. National Kidney, F., *K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification.* Am J Kidney Dis, 2002. **39**(2 Suppl 1): p. S1-266.
177. Kurki, M.I., et al., *FinnGen provides genetic insights from a well-phenotyped isolated population.* Nature, 2023. **613**(7944): p. 508-+.
178. Cortes, A. and M.A. Brown, *Promise and pitfalls of the Immunochip.* Arthritis Res Ther, 2011. **13**(1): p. 101.
179. Sanderson, E., et al., *Mendelian randomization.* Nature Reviews Methods Primers, 2022. **2**(1).
180. Smith, G.D. and S. Ebrahim, *'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?* Int J Epidemiol, 2003. **32**(1): p. 1-22.
181. Cunninghame Graham, D.S., et al., *Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus.* PLoS Genet, 2011. **7**(10): p. e1002341.
182. Gateva, V., et al., *A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus.* Nat Genet, 2009. **41**(11): p. 1228-33.
183. Hughes, T., et al., *Fine-mapping and transethnic genotyping establish IL2/IL21 genetic association with lupus and localize this genetic effect to IL21.* Arthritis Rheum, 2011. **63**(6): p. 1689-97.
184. International Consortium for Systemic Lupus Erythematosus, G., et al., *Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci.* Nat Genet, 2008. **40**(2): p. 204-10.
185. Kim, K., et al., *Variation in the ICAM1-ICAM4-ICAM5 locus is associated with systemic lupus erythematosus susceptibility in multiple ancestries.* Ann Rheum Dis, 2012. **71**(11): p. 1809-14.
186. Leffers, H.C.B., et al., *Established risk loci for systemic lupus erythematosus at NCF2, STAT4, TNPO3, IRF5 and ITGAM associate with distinct clinical manifestations: A Danish genome-wide association study.* Joint Bone Spine, 2022. **89**(4): p. 105357.

187. Lessard, C.J., et al., *Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study.* Am J Hum Genet, 2012. **90**(4): p. 648-60.
188. Lessard, C.J., et al., *Identification of a systemic lupus erythematosus susceptibility locus at 11p13 between PDHX and CD44 in a multiethnic study.* Am J Hum Genet, 2011. **88**(1): p. 83-91.
189. Molineros, J.E., et al., *Admixture mapping in lupus identifies multiple functional variants within IFIH1 associated with apoptosis, inflammation, and autoantibody production.* PLoS Genet, 2013. **9**(2): p. e1003222.
190. Sheng, Y.J., et al., *Association analyses confirm five susceptibility loci for systemic lupus erythematosus in the Han Chinese population.* Arthritis Res Ther, 2015. **17**(1): p. 85.
191. *FinnGen R5 Phenotype Data*. 2021 [cited 2023 2023-03-16]; Available from: r5.risteys.finngen.fi.
192. Organization, W.H., *ICD-10 : international statistical classification of diseases and related health problems : tenth revision, 2nd ed*, W.H. Organization, Editor. 2004.
193. Hagberg, N., et al., *The STAT4 SLE risk allele rs7574865[T] is associated with increased IL-12-induced IFN-gamma production in T cells from patients with SLE.* Ann Rheum Dis, 2018. **77**(7): p. 1070-1077.
194. Andersson, T., et al., *Calculating measures of biological interaction.* Eur J Epidemiol, 2005. **20**(7): p. 575-9.
195. Mathur, M.B. and T.J. VanderWeele, *R Function for Additive Interaction Measures.* Epidemiology, 2018. **29**(1): p. e5-e6.
196. Hemani, G., K. Tilling, and G. Davey Smith, *Orienting the causal relationship between imprecisely measured traits using GWAS summary data.* PLoS Genet, 2017. **13**(11): p. e1007081.
197. Hemani, G., et al., *The MR-Base platform supports systematic causal inference across the human phenome.* Elife, 2018. **7**.
198. Chambers, S.A., et al., *Damage and mortality in a group of British patients with systemic lupus erythematosus followed up for over 10 years.* Rheumatology (Oxford), 2009. **48**(6): p. 673-5.
199. Goncalves, M.J., et al., *Characterization of damage in Portuguese lupus patients: analysis of a national lupus registry.* Lupus, 2015. **24**(3): p. 256-62.
200. Legge, A., et al., *Prediction of Damage Accrual in Systemic Lupus Erythematosus Using the Systemic Lupus International Collaborating Clinics Frailty Index.* Arthritis Rheumatol, 2020. **72**(4): p. 658-666.
201. Bruce, I.N., et al., *Factors associated with damage accrual in patients with systemic lupus erythematosus: results from the Systemic Lupus International Collaborating Clinics (SLICC) Inception Cohort.* Ann Rheum Dis, 2015. **74**(9): p. 1706-13.
202. Conti, F., et al., *The chronic damage in systemic lupus erythematosus is driven by flares, glucocorticoids and antiphospholipid antibodies: results from a monocentric cohort.* Lupus, 2016. **25**(7): p. 719-26.
203. Gianfrancesco, M.A., et al., *Genetic risk and longitudinal disease activity in systemic lupus erythematosus using targeted maximum likelihood estimation.* Genes Immun, 2016. **17**(6): p. 358-62.
204. Alonso-Perez, E., et al., *Lack of replication of higher genetic risk load in men than in women with systemic lupus erythematosus.* Arthritis Res Ther, 2014. **16**(3): p. R128.

205. Hughes, T., et al., *Analysis of autosomal genes reveals gene-sex interactions and higher total genetic risk in men with systemic lupus erythematosus.* Ann Rheum Dis, 2012. **71**(5): p. 694-9.
206. Taylor, K.E., et al., *Risk alleles for systemic lupus erythematosus in a large case-control collection and associations with clinical subphenotypes.* PLoS Genet, 2011. **7**(2): p. e1001311.
207. Habe, K., et al., *Presence of Antiphospholipid Antibodies as a Risk Factor for Thrombotic Events in Patients with Connective Tissue Diseases and Idiopathic Thrombocytopenic Purpura.* Intern Med, 2016. **55**(6): p. 589-95.
208. Hermansen, M.L., et al., *The risk of cardiovascular morbidity and cardiovascular mortality in systemic lupus erythematosus and lupus nephritis: a Danish nationwide population-based cohort study.* Rheumatology (Oxford), 2017. **56**(5): p. 709-715.
209. Uzrail, A.H., A.M. Assaf, and S.S. Abdalla, *Correlations of Expression Levels of a Panel of Genes (IRF5, STAT4, TNFSF4, MECP2, and TLR7) and Cytokine Levels (IL-2, IL-6, IL-10, IL-12, IFN-gamma, and TNF-alpha) with Systemic Lupus Erythematosus Outcomes in Jordanian Patients.* Biomed Res Int, 2019. **2019**: p. 1703842.
210. Di Stefano, A., et al., *STAT4 activation in smokers and patients with chronic obstructive pulmonary disease.* Eur Respir J, 2004. **24**(1): p. 78-85.
211. Yang, C., et al., *STAT4: an immunoregulator contributing to diverse human diseases.* Int J Biol Sci, 2020. **16**(9): p. 1575-1585.
212. Furumoto, Y., et al., *Tofacitinib Ameliorates Murine Lupus and Its Associated Vascular Dysfunction.* Arthritis Rheumatol, 2017. **69**(1): p. 148-160.
213. Egeberg, A., et al., *Clustering of autoimmune diseases in patients with rosacea.* J Am Acad Dermatol, 2016. **74**(4): p. 667-72 e1.
214. Salo, O.P., *SLE-like deposition of immunoglobulins in the skin in rosacea. A clinical and immunofluorescent study.* Ann Clin Res, 1970. **2**(1): p. 28-31.
215. Chang, A.L.S., et al., *Assessment of the genetic basis of rosacea by genome-wide association study.* J Invest Dermatol, 2015. **135**(6): p. 1548-1555.
216. Chasset, F., et al., *Rare diseases that mimic Systemic Lupus Erythematosus (Lupus mimickers).* Joint Bone Spine, 2019. **86**(2): p. 165-171.
217. Lu, H., et al., *Detection of Genetic Overlap Between Rheumatoid Arthritis and Systemic Lupus Erythematosus Using GWAS Summary Statistics.* Front Genet, 2021. **12**: p. 656545.
218. Chen, L., et al., *Genome-wide assessment of genetic risk for systemic lupus erythematosus and disease severity.* Hum Mol Genet, 2020. **29**(10): p. 1745-1756.
219. Hu, X., et al., *The JAK/STAT signaling pathway: from bench to clinic.* Signal Transduct Target Ther, 2021. **6**(1): p. 402.
220. Janssens, A.C., et al., *The impact of genotype frequencies on the clinical validity of genomic profiling for predicting common chronic diseases.* Genet Med, 2007. **9**(8): p. 528-35.
221. Thompson, D.J., et al., *UK Biobank release and systematic evaluation of optimised polygenic risk scores for 53 diseases and quantitative traits.* medRxiv, 2022: p. 2022.06.16.22276246.
222. Cui, J., et al., *Risk prediction models for incident systemic lupus erythematosus among women in the Nurses' health study cohorts using genetics, family history, and lifestyle and environmental factors.* Semin Arthritis Rheum, 2023. **58**: p. 152143.
223. Knevel, R., et al., *Using genetics to prioritize diagnoses for rheumatology outpatients with inflammatory arthritis.* Sci Transl Med, 2020. **12**(545).

224. Sandling, J.K., et al., *Molecular pathways in patients with systemic lupus erythematosus revealed by gene-centred DNA sequencing.* Ann Rheum Dis, 2021. **80**(1): p. 109-117.
225. Lu, T., et al., *Polygenic risk score as a possible tool for identifying familial monogenic causes of complex diseases.* Genet Med, 2022. **24**(7): p. 1545-1555.
226. Imgenberg-Kreuz, J., et al., *DNA Methylation-Based Interferon Scores Associate With Sub-Phenotypes in Primary Sjogren's Syndrome.* Front Immunol, 2021. **12**: p. 702037.
227. Bjork, A., et al., *Protein and DNA methylation-based scores as surrogate markers for interferon system activation in patients with primary Sjogren's syndrome.* RMD Open, 2020. **6**(1).

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1958

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-506789

ACTA UNIVERSITATIS
UPSALIENSIS
2023