From Department of Medicine, Solna

Karolinska Institutet, Stockholm, Sweden

IDIOPATHIC INFLAMMATORY MYOPATHIES AND CANCER: FAMILIAL RISK, GENETICS AND CONSEQUENCES

Weng Ian Che

謝詠茵

Karolinska Institutet

Stockholm 2024

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2024

ISBN 978-91-8017-174-8
Cover illustration by Weng Ian Che.

Idiopathic Inflammatory Myopathies and Cancer: Familial risk, Genetics and Consequences

Thesis for Doctoral Degree (Ph.D.)

By

Weng Ian Che

The thesis will be defended in public at Rolf Luft centrum L1 00, Karolinska Universitetssjukhuset Solna Anna Steckséns Gata 53, January 26, 2024, at 09:00.

Principal Supervisor:
Assoc. Prof. Marie Holmqvist
Karolinska Institutet
Department of Medicine, Solna
Division of Clinical Epidemiology

Co-supervisor(s):
Prof. Ingrid E Lundberg
Karolinska Institutet
Department of Medicine, Solna
Division of Rheumatology

Dr. Helga Westerlind
Karolinska Institutet
Department of Medicine, Solna
Division of Clinical Epidemiology

Dr. Karin Hellgren
Karolinska Institutet
Department of Medicine, Solna
Division of Clinical Epidemiology

Opponent:
Prof. Katja Fall
Örebro University
School of Medical Sciences
Clinical Epidemiology and Biostatistics

Examination Board:
Assoc. Prof. Anna-Karin Hultgård Ekwall
University of Gothenburg
Department of Rheumatology and Inflammation Research

Prof. Arvid Sjölander
Karolinska Institutet
Department of Medical Epidemiology and Biostatistics

Assoc. Prof. Linda Lindström
Karolinska Institutet
Department of Oncology-Pathology

Popular science summary of the thesis

Myositis, also called idiopathic inflammatory myopathies, is a rare rheumatic condition. The disease develops when our immune system wrongly identifies our muscles and sometimes other parts of the body, as foreign and want to destroy it. This causes muscle inflammation, especially in the arms and legs. Consequently, individuals with myositis can experience different levels of muscle weakness that affect their everyday tasks and quality of life. Myositis might also bring other health issues like autoimmune diseases and cancer, making lives even more challenging for those affected. Unfortunately, what causes myositis and how it leads to these complications is not fully understood. Similar to many other immune-related conditions, both genes and the environment seem to play a part in causing myositis. This thesis explores how genetics contribute to myositis and how it impacts people's lives, especially when cancer is involved.

Genetics might have a role in causing myositis, which raises two important questions: how much genetics contribute to myositis and how likely is it for family members to also get myositis. **Study I** used Swedish population register data to find answers. It compared the chance of parents, full-siblings or children having myositis among people with and without myositis. We found that people with myositis were four times more likely to have family members affected by it. Furthermore, genetics seemed to have a less crucial role in causing myositis than other nongenetic factors such as environmental stimuli.

If two diseases have similar genetic causes, they might be seen in the same families. For instance, other rheumatic diseases are often observed in relatives of people with myositis. In **Study II**, we not only saw that relatives of people with myositis were more likely to get other rheumatic diseases, but we also noticed that they were more vulnerable to other autoimmune diseases including inflammatory bowel diseases, autoimmune thyroid diseases and celiac disease (gluten intolerance).

We also looked into whether genes play a role in the development of both myositis and cancer in **Study III** and **IV**. We found that in men with dermatomyositis, which is a myositis subtype that is closely linked to cancer, there was a higher chance of their fathers, brothers or sons having cancer compared to those of men without dermatomyositis. Among various cancer subtypes, we noticed that myeloid malignancies, a type of blood cancers, and liver cancer were more likely to occur in relatives of people with myositis compared with those of people without myositis. Even though we did not find a greater risk of B-cell lymphomas among relatives of people with myositis, our genetic analyses showed shared genetic patterns between myositis and B-cell lymphomas in a few numbers of genomic regions, especially in parts of genes related to the immune system.

Myositis combined with cancer is a complicated condition that can lower life expectancy. In **Study V**, we followed people from their myositis diagnosis and looked at how the risk of dying changed once they were diagnosed with cancer. We observed that people with myositis who later got cancer had a similar risk of dying from other causes as those who did not get cancer after myositis, but they had a significantly higher risk of dying from cancer. Furthermore, within the first 10 years after being diagnosed with cancer after having myositis,

there was a 10% chance of getting another primary cancer, which further reduced survival chances. We also found that certain traits affected how people fared after a cancer diagnosis. For example, people with dermatomyositis had a higher risk of dying after being diagnosed with cancer compared to other myositis subtypes.

All these discoveries suggest that genetics play a role in myositis development and its connection to other autoimmune diseases and cancer. However, it seems that genetics are not the main factor, and environment might be more important in making people susceptible to myositis and its complications, including other autoimmune diseases and cancer. This thesis also highlights how cancer affects people with myositis. The increased risk of dying from cancer in people with myositis who later got cancer suggests that there is a need for better cancer screening and management, something that future research should focus on.

Populärvetenskaplig sammanfattning av avhandlingen

Myosit, även kallad idiopatisk inflammatorisk myopati, är en sällsynt reumatisk sjukdom. Det inträffar när vårt immunförsvar felaktigt angriper våra muskler och ibland andra delar av kroppen, vilket orsakar pågående muskelinflammation, särskilt i armar och ben. Följaktligen kan individer med myosit uppleva muskelsvaghet på olika nivåer, vilket påverkar deras vardagliga sysslor och livskvalitet. Myosit kan också medföra andra hälsoproblem som autoimmuna sjukdomar och cancer, vilket gör livet ännu mer utmanande för de drabbade. Tyvärr är det inte helt klarlagd vad som orsakar myosit med dess komplikationer. I likhet med många andra immunrelaterade tillstånd verkar både gener och miljö spela en roll i sjukdomsutvecklingen. I denna avhandling undersöker vi hur genetik bidrar till myosit och hur det påverkar människors liv, särskilt i kombination med cancer.

Genetik kan ha en roll i uppkomsten av myosit, vilket väcker två viktiga frågor: hur mycket genetiken bidrar och hur troligt är det att familjemedlemmar också får myosit. I **studie I** använde vi uppgifter från svenska folkbokföringen för att hitta svar. Vi jämförde hur stor risken var för helsyskon, föräldrar eller barn till en som hade myosit, att få myosit jämfört med dem som inte hade myosit. Vi fann att personer med myosit hade fyra gånger större sannolikhet att ha familjemedlemmar som också drabbades av myosit. Dessutom tyder vår forskning på att genetiken kanske har en mindre betydande roll i att orsaka myosit jämfört med andra icke-genetiska faktorer, som till exempel miljömässiga påverkningar.

Om två sjukdomar har liknande genetiska orsaker kan de förekomma i samma familjer. Till exempel observeras andra reumatiska sjukdomar ofta hos släktingar till personer med myosit. I **studie II** såg vi inte bara att släktingar till personer med myosit var mer benägna att få andra reumatiska sjukdomar, utan vi märkte också att de var mer mottagliga för andra autoimmuna sjukdomar inklusive inflammatoriska tarmsjukdomar, autoimmuna sköldkörtelsjukdomar och celiaki (glutenintolerans).

Vi undersökte också om gener spelar en roll i utvecklingen av både myosit och cancer i **studie III och IV**. Vi fann att hos män med dermatomyosit, som är en subtyp av myosit som är nära kopplad till cancer, fanns det en högre risk att deras fäder, bröder eller söner skulle få cancer jämfört med män utan dermatomyosit. Bland olika typer av cancer märkte vi att vissa typer av blodcancer och levercancer var vanligare hos släktingar till personer med myosit jämfört med personer utan myosit. Vi hittade inte någon större risk för blodcancer av typ lymfom bland släktingar till personer med myosit men våra genetiska analyser visade, att samma regioner i vissa gener var inblandade för myosit och B-cellslymfom, särskilt i delar av gener relaterade till immunförsvaret.

Myosit i kombination med cancer är ett komplicerat tillstånd som kan sänka den förväntade livslängden. I **studie V** följde vi människor från tiden då de fick sin myositdiagnos och tittade på hur risken att dö förändrades när de fick diagnosen cancer. Vi såg att personer medmyosit som senare fick cancer hade lika stor risk att dö av andra orsaker som de som inte fått cancer efter myosit. Däremot hade de signifikant högre risk att dö i cancer. Dessutom, inom de första 10 åren efter att ha diagnostiserats med cancer, fanns det en 10% risk att få en annan primär cancer, vilket ytterligare minskade chansen för överlevnad. Vi fann också att personer med

dermatomyosit hade högre risk att dö efter att ha diagnostiserats med cancer jämfört med andra subtyper av myosit.

Alla dessa upptäckter tyder på att genetik spelar en roll i utvecklingen av myosit och dess koppling till andra autoimmuna sjukdomar och cancer. Det verkar dock som att miljön kan vara viktigare för att göra människor mottagliga för myosit och dess komplikationer, inklusive andra autoimmuna sjukdomar och cancer. Denna avhandling belyser också hur cancer påverkar personer med myosit. Den ökade risken att dö i cancer hos personer med myosit som också fått cancer talar för att det finns ett behov av bättre screening och hantering för cancer, något som framtida forskning bör fokusera på.

論文科普摘要

肌炎，也稱特異性發炎性肌炎，是一種罕見的風濕性疾病。其病因是由於免疫系統錯誤地攻擊我們的肌肉組織甚至身體其他組織，導致手和腳部的肌肉持續性發炎。因此，肌炎患者會有不同程度的肌肉無力，令日常生活受到影響。肌炎患者還會伴隨著其他的健康問題，比如患上其他自體免疫性疾病和癌症，這些併發症都會影響患者的預後。目前我們對肌炎的病發原因以及是甚麼導致了肌炎患者容易患上上述併發症沒有一個全面的了解。我們知道的是，類似於許多其他免疫相關的疾病，基因和環境似乎在引發肌炎的過程中發揮著一定的作用。本論文探討了基因在導致肌炎和其相關併發症中的作用，以及癌症如何影響肌炎患者的預後。

如果基因能在引發肌炎中起著一定的作用，後續兩個相關的問題是：基因有多大的作用？對比普通人群的親屬，肌炎患者的親屬會有更高的風險患上肌炎嗎？利用瑞典全國人口健康數據，**研究一**發現肌炎患者的父母、兄弟姐妹或子女（一級親屬）比起非患者的一級親屬有多四倍的機會患上肌炎。但是，我們也發現了比起基因，其他非遺傳因素諸如環境刺激等可能在導致肌炎中起著更大的作用。

再有，如果兩種病有相似的遺傳因素，那麼這兩種病就容易出現在有血緣關係的人中。例如，我們常常發現肌炎患者的血緣親屬患有其他風濕性疾病。在**研究二**中，我們不僅看到肌炎患者的一級親屬比起非患者的一級親屬更有可能患上其他風濕性疾病，還發現了他們更容易患有炎症性腸病，自體免疫性甲狀腺疾病和乳糜瀉（乳糖不耐受症）。

我們也在**研究三**和**研究四**中調查了基因在肌炎和癌症的發病中起作用的可能性，結果顯示，患有皮肌炎的男性，其一級親屬患上癌症的可能性比起沒有患皮肌炎的男性的一級親屬更高，特別是骨髓性白血病和肝癌。雖然我們並未發現肌炎患者的一級親屬有更高的風險患上 B 細胞淋巴癌，但在**研究四**中，通過分析肌炎和 B 細胞淋巴癌的基因數據，我們發現了兩者在少數一些基因區域中共享遺傳信息，尤其是在與免疫相關的基因區域。

肌炎合併癌症是一種更為複雜的情況，也會導致較差的預後，降低患者的預期壽命。在**研究五**中，我們由肌炎發病開始追踪患者並觀察了他們在癌症診斷後死亡風險的變化。結果顯示，肌炎患者被診斷出癌症後，其非癌症引致的死亡風險與癌症被診斷前相當，但由癌症引致的死亡風險大大提高。此外，在患上癌症後的頭 10 年內，肌炎患者再次罹患癌症的風險大約為 10%。同時，再次罹患癌症的肌炎患者，其預期壽命會進一步下降。我們也發現了一些會影響患者預後的特徵。例如，皮肌炎患者在癌症診斷後比起有其他肌炎亞型的患者有更高的死亡風險。

綜合上述，我們的研究結果表明，基因在肌炎的發病和其關聯的併發症包括其他自體免疫性疾病和癌症的發病中起著作用。然而，基因似乎不並是主要因素，環境因素可能起著更重要的作用。我們也指出了癌症對肌炎患者預後的影響，癌症顯著地

增加了肌炎患者的死亡風險，表明有必要加強癌症的控制，這應是後續研究應該關注的重點。

Abstract

Idiopathic inflammatory myopathies (IIMs) are a group of rare rheumatic inflammatory diseases (RIDs), characterised by a diverse range of clinical, serological and histopathological characteristics, with muscle weakness as a shared hallmark. While advancements in disease management have improved the survival rates of patients with IIM, the mortality rate among patients with IIM is still higher than the general population, mainly due to association with comorbidities such as cancer. The pathogenesis of IIM, the pathological link between IIM and cancer and the impact of cancer on the survival of patients with IIM remain a subject of uncertainty. The rarity and heterogeneity inherent in IIM pose significant challenges in filing these knowledge gaps. This thesis encompasses five studies, which aimed at addressing research questions concerning the genetic contribution to IIM and its link with other autoimmune diseases and cancer, as well as the disease burden in the context of cancer in a large representative population of patients with IIM.

Study I was a population-based case-control family study including 7,615 first-degree relatives of 1,620 patients with IIM diagnosis between 1997 and 2016 and 37,309 first-degree relatives of 7,797 matched comparators without IIM. Patients with IIM were four times more likely to have at least one first-degree relative affected by IIM compared to matched comparators without IIM. The heritability of IIM, a proportion of the phenotypic variance that can be explained by additive genetic variance, was 22% in the Swedish population.

Study II, with the same study population as in **Study I**, analysed the familial associations between IIM and a variety of autoimmune diseases under a causal framework. We found shared familial factors between IIM and other RIDs, inflammatory bowel diseases, autoimmune thyroid diseases and celiac disease.

Study III, with a similar study population and analytical approach as in **Study II**, comprehensively investigated the familial co-aggregation of IIM and cancer. We did not observe a familial association between IIM and cancer overall but modification effect by sex was noted: there was a modest familial association (adjusted odds ratio=1.39) with cancer in male first-degree relatives of patients with IIM. We also found that offspring of patients with IIM were more likely to have a cancer diagnosis at age younger than 50 years compared to those of matched comparators without IIM. In the exploratory analysis by specific cancer types, findings suggest that IIM shared familial factors with myeloid malignancies and liver cancer.

Study IV explored genetic correlation between IIM and B-cell lymphomas via a cross-trait secondary analysis using summary statistics from genome-wide associations studies of IIM and four common B-cell lymphoma subtypes including diffuse large B-cell lymphoma, follicular lymphoma, chronic lymphocytic leukaemia and marginal zone lymphoma. We detected a limited number of genomic loci, predominantly within the human leukocyte antigen region, demonstrating significant genetic correlations between IIM and common B-cell lymphoma subtypes.

Study V, a cohort study, followed 1,826 patients to (first and second) cancer and death (overall and cause-specific death) events since IIM diagnosis for more than 20 years.

Compared to patients with no cancer diagnosis after IIM, patients with a first cancer diagnosis after IIM faced a greater five-year mortality (22% versus 49%). This excessive risk was due to an increased risk of death from cancer. In patients with a first cancer diagnosis after IIM, the one-year risk of having a second primary cancer was 11% and having a second cancer diagnosis slightly increased the risk of death. We also reported several prognostic factors associated with increased risks of cancer and death (overall, from cancer and from other causes).

This thesis offers useful insight into the role of genetics in IIM pathogenesis and its connections with other autoimmune diseases and cancer, as well as the impact of cancer on the survival of patients with IIM. The observed familial aggregation of IIM and familial associations between IIM and other autoimmune diseases suggest genetic involvement in the development of IIM. Family history of IIM, other RIDs, inflammatory bowel diseases, autoimmune thyroid diseases and celiac disease may serve as indicators pointing towards an IIM diagnosis. Missing heritability is suggested by the discrepancy between our family-based heritability and the SNP-based heritability, implying yet-to-be discovered genetic variants associated with IIM. The acquired knowledge of shared familial factors between IIM and other autoimmune diseases may inform future genetic studies aiming to uncover novel IIM-associated genetic variants. There is a limited shared familial/genetic susceptibility between IIM and cancer. The human leukocyte antigen region plays an important role in the limited shared genetic susceptibility between IIM and common B-cell lymphoma subtypes. IIM concomitant with cancer leads to a substantial increase in mortality, mainly due to cancer. Future research should focus on reducing cancer-related disease burden in patients with IIM.

List of scientific papers

I. Che WI, Westerlind H, Lundberg IE, Hellgren K, Kuja-Halkola R, Holmqvist M. **Familial aggregation and heritability: a nationwide family-based study of idiopathic inflammatory myopathies**. Ann Rheum Dis. 2021;80(11):1461-6.

II. Che WI, Westerlind H, Lundberg IE, Hellgren K, Kuja-Halkola R, Holmqvist ME. **Familial autoimmunity in patients with idiopathic inflammatory myopathies**. J Intern Med. 2023;293(2):200-11.

III. Che WI, Baecklund F, Hellgren K, Kuja-Halkola R, Lundberg IE, Westerlind H, et al. **Familial Co-Aggregation of Idiopathic Inflammatory Myopathies and Cancer: A Swedish Population-Based Study**. Arthritis Rheumatol. 2023;75(8):1445-55.

IV. Che WI, Öberg Sysojev A, Zhu C, Patasova K, International Lymphoma Epidemiology Consortium (InterLymph), IMACS Myositis Genetics Scientific Interest Group (MYOGEN), Ekström Smedby K, Lundberg IE, Westerlind H, Holmqvist M. **The human leukocyte antigen region as a key player in the limited shared genetic susceptibility between idiopathic inflammatory myopathies and common B-cell lymphoma subtypes**. Manuscript.

V. Che WI, Bower H, Kuja-Halkola R, Hellgren K, Lundberg IE, Westerlind H, Baecklund F, Holmqvist M. **Exploring the impact of cancer on the prognosis of patients with idiopathic inflammatory myopathies by flexible parametric multistate modelling**. Manuscript.

Scientific papers not included in the thesis

I. Che WI, Hellgren K, Lundberg IE, Holmqvist M. **Reproductive Pattern in Women with Idiopathic Inflammatory Myopathy: A Population-based Study**. J Rheumatol. 2020;47(9):1392-6.

II. Che WI, Hellgren K, Stephansson O, Lundberg IE, Holmqvist M. **Pregnancy outcomes in women with idiopathic inflammatory myopathy, before and after diagnosis-a population-based study**. Rheumatology (Oxford). 2020;59(9):2572-80.

III. Dani L, Ian Che W, Lundberg IE, Hellgren K, Holmqvist M. **Overall and site-specific cancer before and after diagnosis of idiopathic inflammatory myopathies: A nationwide study 2002-2016**. Semin Arthritis Rheum. 2021;51(1):331-7.

IV. Che WI, Lundberg IE, Holmqvist M. **Environmental Risks for Inflammatory Myopathies**. Rheum Dis Clin North Am. 2022;48(4):861 74.

Contents

1 Literature review 1

1.1 Idiopathic Inflammatory myopathies 1
1.1.1 Incidence and prevalence 1
1.1.2 Clinical subtypes 1
1.1.3 Classification criteria 4
1.1.4 Risk factors 5
1.1.5 Disease mechanisms 12
1.1.6 Mortality 13
1.2 IIM in the context of cancer 13
1.2.1 Epidemiology 13
1.2.2 Associated IIM subtypes 14
1.2.3 Associated site-specific and histological subtypes of cancer 14
1.2.4 Predictors 15
1.2.5 Pathogenesis 16
1.2.6 Cancer screening 19
1.2.7 Disease management 20
1.2.8 Prognosis 21
1.3 Specific background of Study I and II 22
1.3.1 Inherited IIM mimics 22
1.3.2 Polyautoimmunity, familial autoimmunity and shared genetic susceptibility with other autoimmune diseases 23
1.4 Specific background of Study IV 24
1.4.1 The risks of common B-cell lymphoma subtypes in IIM 24
1.4.2 Genetic variants associated with both IIM and common B-cell lymphoma subtypes 24

2 Research aims 25

3 Materials and methods 27

3.1 Study setting 27
3.2 Data sources 27
3.2.1 Swedish National Registers 27
3.2.2 The Myositis Genetics Scientific Interest Group of the International Myositis Assessment & Clinical Studies Group 29
3.2.3 The International Lymphoma Epidemiology Consortium 30
3.3 Study populations 31
3.3.1 Patients with IIM and matched comparators without IIM identified in the NPR 31
3.3.2 First-degree relatives 34
3.3.3 Defining IIM and autoimmune diseases in first-degree relatives 34
3.3.4 Population-based case-control family design, accompanied with Directed Acyclic Graph 36
3.3.5 Study I: Familial aggregation and heritability of IIM 39
3.3.6 Study II: Familial autoimmunity of IIM 40
3.3.7 Study III: Familial co-aggregation of IIM and cancer 40
3.3.8 Study IV: Genetic overlap between IIM and common B-cell lymphoma subtypes 41
3.3.9 Study V: The impact of cancer on the prognosis of patients with IIM 44
3.4 Ethical consideration 48

4 Key findings 51

4.1 Study I: Familial aggregation and heritability of IIM 51
4.2 Study II: Familial autoimmunity of IIM 53
4.3 Study III: Familial co-aggregation of IIM and cancer 55
4.4 Study IV: Genetic overlap between IIM and common B-cell lymphoma subtypes 57
4.5 Study V: The impact of cancer on the prognosis of patients with IIM 62
5 Discussion 71
5.1 Understanding the familial and genetic susceptibility of IIM with population data 71
5.2 IIM and cancer: from genetics to consequences 73
5.3 Other considerations 76
6 Conclusions 79
7 Acknowledgements 81
8 References 85

List of abbreviations

aaRS	Aminoacyl tRNA synthetase
ACR	American College of Rheumatology
AH	Ancestral haplotype
aOR	Adjusted odds ratio
ASyS	Antisynthetase syndrome
CAM	Cancer-associated myositis
CADM	Clinically amyopathic dermatomyositis
CDR	Cause of Death Register
CK	Creatine kinase
CLL	Chronic lymphocytic leukaemia
CondFDR	Conditional false discovery rate
ConjFDR	Conjunctional false discovery rate
CRT	Calreticulin
CTD	Connective tissue disease
DAG	Directed Acyclic Graph
DLBCL	Diffuse large B-cell lymphoma
DM	Dermatomyositis
EULAR	European League Against Rheumatism
FIGO	International Federation of Gynecology and Obstetrics
FL	Follicular lymphoma
HLA	Human leukocyte antigen
HMGCR	3-hydroxy-3-methylglutaryl CoA reductase
GWAS	Genome-wide association study
IBM	Inclusion body myositis
ICD	International Classification of Diseases
ICI	Immune checkpoint inhibitor
IFN	Interferon
IIM	Idiopathic inflammatory myopathy
ILD	Interstitial lung disease
IMACS	International Myositis Assessment & Clinical Studies Group
IMNM	Immune-mediated necrotising myopathy
InterLymph	International Lymphoma Epidemiology Consortium

IQR	Interquartile range
JDM	Juvenile dermatomyositis
JIIM	Juvenile idiopathic inflammatory myopathy
Jo1	Histidyl tRNA synthetase
LAVA	Local analysis of (co)variant association
LD	Linkage disequilibrium
MAA	Myositis-associated autoantibody
MCTD	Mixed connective tissue disease
MDA5	Melanoma differentiation-associated gene 5
MG	Myasthenia gravis
MGR	Multi-Generation Register
Mi2	Nucleosome remodelling deacetylase complex
MM	Multiple myeloma
MS	Multiple sclerosis
MSA	Myositis-specific autoantibody
MYOGEN	Myositis Genetics Scientific Interest Group
MZL	Marginal zone lymphoma
NHL	Non-Hodgkin lymphoma
NPC	Nasopharyngeal cancer
NPR	National Patient Register
NXP2	Nuclear matrix protein 2
OM	Overlap myositis
PC	Principal component
PDR	Prescribed Drug Register
PleioFDR	Pleiotropy-informed false discovery rate
PPV	Positive predictive value
RA	Rheumatoid arthritis
RID	Rheumatic inflammatory disease
SAE	Small ubiquitin-like modifier activatin enzyme
SCR	Swedish Cancer Register
SIR	Standardised incidence ratio
SLE	Systemic lupus erythematosus
SNOMED	Systematized Nomenclature of Medicine

SNP	Single nucleotide polymorphism
SRP	Signal recognition particle
SS	Sjögren syndrome
SSc	Systemic sclerosis
T1DM	Type 1 diabetes mellitus
TIF1γ	Transcriptional intermediary factor 1 γ
TNM	Tumour, node and metastasis
TPR	Total Population Register
UV	Ultraviolet

Introduction

Idiopathic inflammatory myopathies (IIMs) were initially documented as a disease in 1863 (1). Today, IIM is acknowledged as a rare and heterogeneous group of rheumatic inflammatory diseases (RIDs) primarily affecting proximal muscles, yet often with manifestations in multiple organs such as skin, lung, joints, heart, and gastrointestinal tract (2). Decades of research efforts have significantly progressed our understanding of IIM pathogenesis and its management. Notably, following the first genome-wide association study (GWAS) of IIM in 2013, mounting evidence has underscored the pivotal role of the human leukocyte antigen (HLA) region in the genetic susceptibility of IIM (3-7). Nevertheless, the extent of genetic contribution to IIM remains uncertain.

Cancer represents one of the common comorbidities linked to IIM (8, 9). Although the association between IIM and cancer is acknowledged, the underlying pathological connection is not fully understood. Additionally, increased mortality has been observed in patients with both IIM and cancer (10). A comprehensive exploration of the cancer-related disease burden in IIM patients has been lacking, given the rarity of the co-occurrence of IIM and cancer.

This thesis endeavors to address key research inquiries concerning the genetic underpinnings of IIM, including shared genetic susceptibility with other autoimmune diseases and cancer, as well as examining the impact of cancer on the prognosis of patients with IIM. The subsequent literature review will provide an insightful comprehension of each study encompassed within this thesis.

1 Literature review

1.1 Idiopathic Inflammatory myopathies

1.1.1 Incidence and prevalence

Idiopathic inflammatory myopathies (IIMs) are very rare rheumatic conditions, with incidence and prevalence varying across regions (11-35). A meta-analysis including 46 studies reported a pooled incidence of IIM at 8.0 per one million person-years (95% CI 7.4-8.7) and a pooled prevalence of IIM at 14.0 per 100,000 persons (95% CI 12.8-15.5) (36). Similar estimates of the incidence and prevalence of IIM have been found in other two population-based studies (34, 35). In Sweden, IIM has an incidence of 11 per one million person-years (95% CI 10-12) and a prevalence of 14.0 per 100,000 persons (95% CI 13.0-15.0) (34). For IIM overall, an increase in incidence over time has been suggested in some studies (22, 26, 31). The incidence of IIM in the UK raised to 21 per one million person-years in 2016 from 14 per one million person-years in 2007 (31). However, the incidence of IIM remained stable over time periods, 1980 to 2014 in Western Australia and 2007 to 2011 in Sweden.

The incidence and prevalence of IIM differ in sex, age and IIM subtypes. Women have approximately a two-fold higher risk of IIM, excluding inclusion body myositis (IBM), than men (31, 34). For IBM, the sex risk ratio is reversed (34, 37). Although there is a juvenile onset of IIM, the incidence rate is the highest over 50 years of age and peaks at 70-79 years of age (34). Subtype-specific data on incidence and prevalence are sparse and inconsistent. Adult dermatomyositis (DM) has been consistently reported as the most common IIM subtype, with an incidence of 3.2 per one million person-years (95% CI 2.7-3.7) and a prevalence of 3.8 per 100,000 persons (95% CI 3.5-4.3) in Sweden (31, 33, 34). The incidence and prevalence of IBM, polymyositis (PM) and juvenile DM (JDM) in Sweden is presented in **Table 1**. Notably, the higher incidence and prevalence of PM compared to DM are due to inclusion of other IIM subtypes within the PM group (34). The recognition of new IIM subtypes and updated classification criteria for IIM have resulted in a more homogeneous grouping of IIM subtypes. This evolution might account for the increased incidence of overlap myositis (OM) and other IIM, and the decreased incidence of PM observed in a recently published population-based study in Western Australia (35).

Table 1. The incidence and prevalence of IBM, PM and JDM in Sweden

	Incidence (95% CI), per one million person-years	Prevalence (95% CI), per 100,000 persons
IBM	1.1 (0.9-1.5)	1.0 (0.8-1.2)
PM	5.7 (5.1-6.4)	7.6 (7.0-8.2)
JDM	0.8 (0.5-1.1)	1.1 (0.9-1.3)

IBM: Inclusion body myositis; PM: Polymyositis; JDM: Juvenile dermatomyositis; CI: Confidence interval.

1.1.2 Clinical subtypes

Of note, with the identification of antisynthetase syndrome (ASyS), immune-mediated necrotising myopathy (IMNM) and IBM, the existence of PM as a distinct subtype of IIM is doubtful (38). Each of the subtypes has a rather distinct clinical, histopathological,

serological and prognostic pattern, but there is overlap between the phenotypes (**Figure 1** on page 4).

1.1.2.1 Dermatomyositis

DM is defined by the presence of skin manifestations and myositis primarily affecting proximal muscles. The other three less common conditions under the DM group are amyopathic DM, where muscular manifestation is absent, hypomyopathic DM, where there is laboratory evidence suggesting myositis, but no obvious muscle symptoms are presented, and DM sine dermatitis, where only myositis is presented, and muscle biopsy findings are in line with DM features. The first two conditions are defined when they last for at least six months and are often collectively called clinically amyopathic DM (CADM). High creatine kinase (CK) level is usually found in patients with DM. The typical histopathological features on muscle biopsy of DM include perivascular inflammatory infiltrate with perifascicular atrophy, membrane attack deposition on capillaries and myxovirus resistance A expression in non-necrotic fibres. Eight DM-specific autoantibodies have been discovered and they are associated with distinct clinical features; anti-nucleosome remodelling deacetylase complex (Mi2) and anti-nuclear matrix protein 2 (NXP2) autoantibodies are often linked to remarkable muscle diseases; anti-melanoma differentiation-associated gene 5 (MDA5) autoantibodies are usually amyopathic and associated with rapidly progressive interstitial lung disease (ILD); anti-small ubiquitin-like modifier activatin enzyme (SAE) autoantibodies are associated with severe skin manifestations and dysphagia; anti-transcriptional intermediary factor 1 γ (TIF1γ), anti-Sp4 and anti-CCAR1 autoantibodies are linked to cancer occurrence (2, 39).

1.1.2.2 Antisynthetase syndrome

ASyS is defined by the presence of anti-aminoacyl tRNA synthetase (aaRS) autoantibodies (2, 39). Anti-histidyl tRNA synthetase (Jo1) autoantibodies, first discovered in 1984, is the most common anti-aaRS autoantibodies (40). Seven more anti-aaRS autoantibodies including anti-PL-7, anti-PL-12, anti-EJ, anti-OK, anti-KS, anti-Zo and anti-Ha autoantibodies have been identified. In addition to proximal muscle weakness, arthritis, Raynaud's phenomenon, ILD and mechanic's hands are frequently observed in patients with ASyS. Histopathological features of ASyS include perifascicular necrosis, upregulation of major histocompatibility complex class I and II proteins in cytoplasm and scarcolemma fibres (2, 39).

1.1.2.3 Overlap myositis

Although lack of a consensus on definition, OM is broadly known as a co-occurrence of IIM and other connective tissue diseases (CTDs) including systemic sclerosis (SSc), systemic lupus erythematosus (SLE), rheumatoid arthritis (RA) and Sjögren syndrome (SS). SSc is the most common overlapping disease with IIM (5%-43%), followed by SLE (4%-21%), SS (8%-15%) and RA (3%-9%) (41-43). Moreover, in a Spanish cohort including 342 patients with OM, 33% of them had mixed CTD (MCTD), a rare disease condition comprising features of multiple CTD, a lower rate of 15% was however found in OM patients in the MYONET registry (44, 45). OM has variable clinical presentations and are associated with a higher rate of infection than other major IIM subtypes (44). Myositis-associated

autoantibodies (MAAs) are frequently found in patients with OM; anti-PM/Scl autoantibodies usually present in patients with IIM overlapping SSc; anti-Ro52 autoantibodies may occur with anti-Jo1 or anti-MDA5 autoantibodies and are associated with poor prognosis due to severe ILD; anti-Ku autoantibodies are also linked to ILD and tend to have distal muscle weakness; anti-U1RNP autoantibodies are linked to IIM overlapping with SSc, SLE and RA.

1.1.2.4 Immune-mediated necrotizing myopathy

IMNM is characterised by severe muscle manifestation, extremely high CK level and infrequent skin rash. Necrosis and regeneration of muscle fibres with sparse inflammatory infiltrate are key histopathological features. There are two IMNM specific autoantibodies; anti-3-hydroxy-3-methylglutaryl CoA reductase (HMGCR) autoantibodies, often found in patients with statin use, typically do not exhibit extramuscular manifestations. In contrast, anti-signal recognition particle (SRP) autoantibodies are more frequently associated with cardiac involvement and dysphagia (2, 39).

1.1.2.5 Inclusion body myositis

IBM is recognised by asymmetrical muscle weakness typically in the long finger flexors and quadriceps. Patients with IBM rarely have high CK level and extramuscular manifestations except dysphagia which can be observed in more than half of patients. The characteristic features of muscle biopsy are endomysial T cell infiltrates of $CD8^+$ T cells and rimmed vacuoles (2, 39). Anti-cN1A autoantibodies associated with IBM, are also seen in other connective tissue diseases but its linked clinical feature is unclear (2, 39).

1.1.2.6 Polymyositis

PM is characterized by proximal muscle weakness without skin manifestation and myositis-specific autoantibodies (MSAs). The key muscle biopsy finding is endomysial inflammatory infilitrates of $CD4^+$ and $CD8^+$ T cells in non-necrotic fibres. As classification criteria for IIM continue to develop, PM is now considered as a rare IIM subtype. Notably, it is important for researchers to consider that ASyS, IMNM and OM were frequently classified as PM in earlier studies (2, 39).

1.1.2.7 Juvenile idiopathic inflammatory myopathies

IIM diagnosed before age of 16 or 18 years is defined as juvenile form of IIM and up to 90% of juvenile IIM (JIIM) are JDM (46). JDM overall shares similarities with adult DM but differences in frequency of DM-specific autoantibodies and their associated clinical features present. Anti-TIF1γ and anti-NXP2 autoantibodies are the first and second prevalent autoantibodies found in JDM but association with malignancy has not been suggested as in adult DM. Anti-TIF1γ autoantibodies, presenting in 23-30% of JDM, are associated with photosensitivity, chronic skin manifestations including severe rash and skin ulceration, oedema and lipodystrophy. Compared to adult DM, JDM presents less extramuscular manifestations but in JDM patients with anti-NXP2 autoantibodies, gastrointestinal manifestation is common (47).

1.1.3 Classification criteria

Several diagnostic criteria have been developed to determine and classify IIM (**Figure 1**). The Bohan and Peter criteria published in 1975 is one of the first and remains widely applied diagnostic criteria nowadays to differentiate DM from PM based on presence of skin manifestations (48, 49). Other diagnostic criteria have emerged as more distinct IIM subtypes have been recognised. The Griggs criteria and the European Neuromuscular Centre criteria are commonly used criteria to classify IBM (50, 51). In 2017, the European League Against Rheumatism (EULAR)/American College of Rheumatology (ACR) proposed scored-based classification criteria for IIM. A major advantage of the EULAR/ACR criteria is that it can determine IIM without muscle biopsy data. A predetermined cut-off for the aggregated score is set to ≥ 5.5 or ≥ 6.7 when muscle biopsy data is available to reach the best balance between sensitivity and specificity. Compared to the Bohan and Peter criteria, the EULAR/ACR criteria demonstrate better performance in identifying and classifying IIM, especially for ADM. However, one major limitation of the EULAR/ACR criteria is that it cannot differentiate ASyS, IMNM and OM from PM (52).

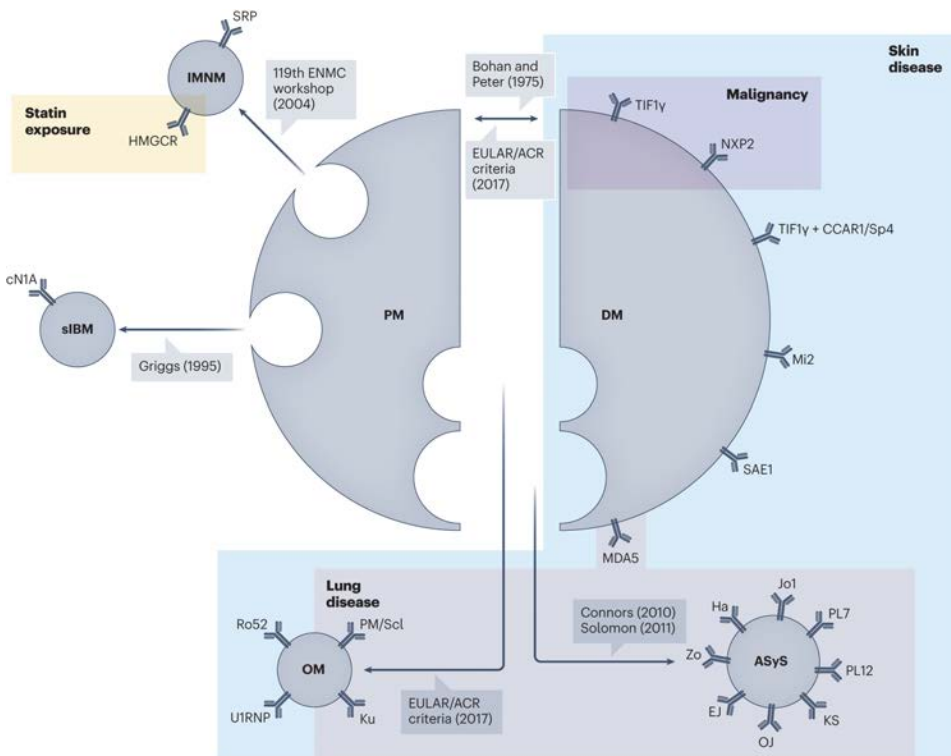


Figure 1. Changing criteria of the idiopathic inflammatory myopathies. *(reproduced with permission from Springer Nature: Khoo T, Lilleker JB, Thong BY, Leclair V, Lamb JA, Chinoy H. Epidemiology of the idiopathic inflammatory myopathies. Nat Rev Rheumatol. 2023;19(11):695-712.)*

1.1.4 Risk factors

Like many autoimmune diseases, the onset of IIM is multifactorial and involves both genetic and environmental factors (53). Numerous genetic and environmental factors associated with IIM have been identified but no single factor to date has been found to trigger IIM by itself.

1.1.4.1 Genetic factors

1.1.4.1.1 Familial component

Although IIM is not considered a hereditary disease, familial aggregation of IIM has been reported, suggesting a genetic predisposition to IIM development. Having more than one family member affected by IIM is very rare. Supportive evidence of familial IIM is mainly from case studies reporting familial JDM, IBM, IBM/PM and ASyS (54-58). There are only four small scale or population-based family studies attempting to investigate familial aggregation of IIM (59-63). One study published in 1998 reported 16 families with at least two first-degree relatives affected by IIM; of these families, six were affected by DM solely, seven were affected by PM solely, one was affected by IBM solely, one was affected by PM and DM, and one was affected by PM and IBM (63). Another study published in the same year and aimed to study familial autoimmunity in 21 pedigrees of IIM by comparing to 21 matched control probands (59). There was only one first-degree relative of patients with IIM developed DM. A family study including first- to third-degree relatives of 304 children with JDM observed that 0.04% of their relatives were affected by DM, compared with the prevalence of DM of 0.02% in the general population, corresponding to an odds ratio (OR) of 3.0 (95% CI 0.3-28.9) (60). A Danish study examined familial aggregation of a number of autoimmune diseases including 949 patients with IIM reported a familial OR of 3.9 (95%CI 0.6-27.7) for parents or siblings of patients with DM and 4.9 (95%CI 0.7-35.2) for parents only (61). Recently, a Swedish study comprising 2,668 patients with one diagnostic code suggesting IIM between 1964 and 2012 observed similar familial aggregation of DM/PM among first-degree relatives (standardised incidence ratio, SIR=4.0, 95%CI 1.3-8.4) (62). Four of the total five familial cases of DM/PM were found in siblings (SIR=7.4, 95%CI 1.9-19.1). The familial association was stronger in men (SIR=5.8, 95%CI 1.1-14.2) than in women (SIR=2.8, 95% CI 0.3-8.0). These findings suggest a familial component in IIM development. However, problems like IIM ascertainment using invalidated algorithms and wide 95% of confidence interval found in the previous studies increase the uncertainty of these findings.

It is unknown if familial IIM differs from sporadic IIM. Rider *et al*. comparing 36 familial IIM to 181 sporadic IIM in terms of clinical, serologic and immunogenetic features found no obvious differences except younger onset age of familial DM (mean±SD, 25.6±18.9 versus 39.7±13.1) and IBM (39.2±16.6 versus 53.7±10.3), lower prevalence of autoimmune diseases (37% versus 61%) in familial IIM. However, the age of onset became similar between adult familial and sporadic DM after excluding JDM cases. Regarding immunogenetic association, both familial (OR=5.5 95% CI 2.6-11.8) and sporadic (OR=9.5 95% CI 5.4-16.9) IIM were associated with *HLA-DRB1*03:01* while *DQA1* homozygosity (of *DQA1*05:01* in particular) was found as a unique genetic risk factors of familial IIM.

This study also found more concordant clinical and serological features between familial IIM than between randomly paired sporadic IIM (63).

1.1.4.1.2 Genetic variants and heritability

Many genetic variants associated with IIM have been identified in the past decade, thanks to international collaboration, advancement in sequencing techniques, and refined statistical analyses. **Table 2** (on page 8) summarises the genetic variants and alterations associated with IIM that have been reported in Caucasian populations. However, the heritability of IIM, defined as the proportion of the phenotypic variation of IIM attributable to genetic variation, remains unclear (64). Rothwell *et al*. performed a post hoc analysis to estimate how much of all single nucleotide polymorphisms (SNPs) on the ImmunoChip account for the phenotypic variation of DM and PM. They reported a SNP-based heritability of 8.3% for PM and 5.5% for DM (65). These figures might be underestimated due to selected loci and limitations associated with the analytical method (66, 67).

The HLA 8.1 ancestral haplotype (AH) on chromosome 6 has been confirmed as the major genetic factor influencing susceptibility to or protection from IIM (4-7). The 8.1 AH is more than four megabases long with alleles in extensive linkage disequilibrium (LD). This haplotype is conserved in European populations and associated with various types of autoimmune diseases (68). Genome-wide association studies (GWASs) have identified *HLA-DRB1*03:01* and *HLA-B*08:01* as the strongest independent genetic variants associated with various IIM subtypes, MSAs and MAAs (4-6, 69-71). Other HLA alleles demonstrating independent effects include *HLA-DQB1*04:02* for DM (5), *HLA-DQB1*02:01* and *HLA-C*02:02* for JDM (4, 70), and *HLA-DRB1*15* for anti-U1RNP autoantibodies (7). Although an interaction between HLA alleles has not been suggested, a combination of multiple alleles in 8.1 AH confers an even stronger risk of IIM than *HLA-DRB1*03:01* or *HLA-B*08:01* alone. The strongest combination differed slightly among clinical IIM subtypes but all of them presented *HLA-B*08:01*, *-DRB1*03:01* and *-DQB1*02:01* (4).

There are also several genes in the HLA class III region associated with IIM overall, subtypes including DM, ASyS, IBM, PM and JDM, and IIM-related autoantibodies (72, 73). For example, low gene copy number of *C4A* has been suggested as risk factor of adult IIM, JDM, anti-Jo1 and PM/Scl autoantibodies, as well as MAA (73, 74).

It has been suggested that HLA associations with IIM may be driven by amino acid associations which have important implications in functionality (5, 7, 69, 71). One typical example is that the arginine at position 74 of *HLA-DRB1* showed similar associations with DM, anti-Jo1, PM/Scl and cN1A autoantibodies as *HLA-DRB1*03:01* did. Arginine at position 74 of *HLA-DRB1* locates within the peptide binding groove and may alter the functional structure of HLA DR protein (5). This structural alteration may affect antigen presentation to T lymphocytes and may eventually contribute to the breakdown of immune tolerance upon a specific exposure. More amino acid associations have recently been reported in serologically defined IIM subgroups and these associations are summarised in **Figure 2** (on page 11) (7).

More than 40 non-HLA variants, particularly immune-related loci, have been discovered to be associated with IIM although not all of the associations have reached GWAS level of

significance (3, 5, 72, 75, 76). Rare variants of the *IFI35* gene have been suggested to be a potential risk factor of IIM but the validation cohort failed to replicate this association (72). A recent study using genome-wide imputation of the ImmunoChip data suggested several novel genetic associations with IIM (75). Furthermore, a higher rate of Klinefelter syndrome (47,XXY), a X chromosome abnormality, has been found in male patients with DM/PM (1.0% versus 0.1%), IBM (4.1% versus 0.1%) and SSc (4.3% versus 0.1%), compared to healthy controls.

Table 2. The genetic alleles and alterations associated with IIM in Caucasian populations and their associations with other autoimmune diseases and common B-cell lymphoma subtypes

Genetic allele/aleration	Effect	Subtypes/autoantibodies	Autoimmune diseases	B-cell lymphomas	Ref
HLA class II					
DRB1*03:01	Risk	IIM, DM, ASyS, IBM, PM, JDM Jo1, PL-12, La, PM/Scl, Ro, cN1A	SLE (77), SS (78), MS (79), IBD (80), T1DM (81), AITD (82), CeD (83), MG (84)		(4-6, 69, 70)
DRB1*03	Risk	PM/Scl, Jo1, Jo1/Ro52			(7)
	Protective	U1RNP, Mi2, TIF1γ			
DRB1*01	Protective	ASyS Anti-Jo1, TIF1γ	SSc (85), SS (78)	FL (86)	(7, 71)
DRB1*04:01	Protective	IIM, PM	RA (87), MS (79), T1DM (81)		(69)
DRB1*04	Risk	U1RNP			(7)
DRB1*07:01	Risk	Mi2	SSc (85), IBD (80), MG (84)	CLL (88)	(71)
	Protective	AsyS Jo1			
DRB1*07	Risk	Mi2, TIF1γ			(7)
DRB1*11	Risk	Ku, HMGCR, U1RNP	SSc (85)		(7, 71)
DRB1*13	Risk	Seronegative		FL (86)	(7)
DRB1*15:01	Protective	IIM	SSc (85), MS (79), IBD (80), MG (84)		(4)
DRB1*15	Risk	U1RNP			(7)
DRB1*16	Risk	Mi2			(7)
DRB1*15/16	Protective	PM/Scl			(71)
DQA1*05:01	Risk	IIM, PM, DM, IBM, AsyS Jo1, PM/Scl, Ro	SLE (77), SSc (85), SS (78), T1DM (81), AITD (82), CeD (83), MG (84)		(4, 69, 71)
DQA1*05	Risk	PM/Scl, Jo1/Ro52			(7)
	Protective	Mi2, TIF1γ			
DQA1*01:01	Protective	AsyS Jo1, PM/Scl	IBD (80)	FL (89)	(71)
DQA1*01:03	Protective	IIM	IBD (80), MG (84)		(69)
DQA1*01	Risk	Seronegative			(7, 69, 71)
	Protective	DM, AsyS			
DQA1*02:01	Risk	Mi2	SSc (85), SS (78), IBD (80), MG (84)		(69, 71)
	Protective	IIM, PM, IBM, AsyS Jo1, Ro			
DQA1*02	Risk	Mi2, TIF1γ			(7)
DQA1*03	Risk	U1RNP			(7, 69)
	Protective	IBM			
DQB1*02:01 (JDM)	Risk	IIM, DM, PM, JDM Jo1, TIF1 (juvenile), PM/Scl	SLE (77), SS (78), IBD (80), T1DM (81), CeD (83)		(4, 6, 70)
DQB1*02:02	Risk	TIF1 (adult)	SSc (85), IBD (80)		(6)
DQB1*02	Risk	PM/Scl, Jo1/Ro52	MG (84)		(7)
	Protective	U1RNP, seronegative			
DQB1*03	Risk	U1RNP			(7)
DQB1*04:02	Risk	DM	IBD (80)		(5)
DQB1*05	Risk	TIF1γ			(7)
DQB1*06	Risk	Seronegative			(7)
	Protective	TIF1γ			
DRA1*01:03	Risk	IIM			(4)

Continued on next page

Table 2. The genetic alleles and alterations associated with IIM in Caucasian populations and their associations with other autoimmune diseases and common B-cell lymphoma subtypes (Continued)

Genetic allele/aleration	Effect	Subtypes/autoantibodies	Autoimmune diseases	B-cell lymphomas	Ref
DRA1*02:01	Risk	IIM, DM			(4)
DPB1*01:01	Risk	IIM, DM, PM Jo1			(4)
	Protective	JDM			
HLA class I					
B*08	Risk	PM/Scl			(71)
B*08:01	Risk	IIM, DM, ASyS, PM, JDM Jo1, La, Ro	SLE (77), IBD (80), MG (84)	DLBCL (90), MZL (91)	(4-6, 69, 71)
B*15	Risk	IIM,			(69)
A*01	Risk	IIM, ASyS Jo1	T1DM (92)		(69, 71)
A*01:01	Risk	PM/Scl	MG (84)		(71)
A*30	Protective	IIM			(69)
A*68	Risk	IIM, DM			(69)
C*02:02	Risk	JDM			(4, 70)
C*03:04	Risk	PL-7	IBD (80)		(71)
C*04	Protective	Jo1			(71)
C*07:01	Risk	IIM, DM, ASyS, PM, JDM Jo1, La, Ro	IBD (80), MG (84)		(4, 69, 71)
C*07:02	Protective	IIM	IBD (80)		(69)
C*14	Risk	IBM			(69)
PSMB8	Risk	ASyS			(72)
PSMB9	Risk	ASyS		CLL (93)	(72)
CCHCR1	Risk	IIM			(72)
HLA class III					
NOTCH4	Risk	IIM, DM			(72)
TNF	Risk	IIM			(72)
AGER	Risk	ASyS			(72)
C2	Protective	IIM, ASyS, PM			(72)
Low gene copy number of CA4	Risk	IIM, IBM, PM, JDM Jo1, PM/Scl, MAA			(73)
CTP21A2	Protective	ASyS			(72)
Non-HLA (chr no.)					
PTPN22 (chr1)	Risk	IIM, PM	AITD (94), IBD (95), RA (76, 96), SLE (76, 97), SSc (98), TIDM (99)		(5, 75, 76)
	Protective	IIM			
YDJC (chr22)	Risk	IIM	CeD (100), IBD (95), RA (76, 96), SLE (76, 97),SSc (76)		(5, 75, 76)
DGKQ (chr4)	Risk	IIM	RA (76), SLE (76), SSc (76, 101)		(5, 75, 76)
	Protective	IIM			
STAT4 (chr2)	Risk	IIM	RA (96), SLE (76, 97), SSc (76, 101)		(5, 75, 76)
	Protective	IIM			
MGCT4A (chr2)	Risk	IIM			(5)
PRR5L|TRAF6 (chr11)	Protective	IIM			(5)
CCL17 (chr16)	Risk	IIM			(5)
EOMES (chr3)	Protective	IIM		DLBCL (102), CLL (93)	(5)

Continued on next page

Table 2. The genetic alleles and alterations associated with IIM in Caucasian populations and their associations with other autoimmune diseases and common B-cell lymphoma subtypes (Continued)

Genetic allele/aleration	Effect	Subtypes/autoantibodies	Autoimmune diseases	B-cell lymphomas	Ref
CD28 (chr2)	Risk	IIM	RA (96)		(5)
RPL31P10 (chr12)	Protective	IIM			(5)
LOC728073|RPL 38 (chr17)	Risk	PM			(5)
UBE3B|MMAB (chr12)	Protective	PM			(5)
NAB1 (chr2)	Risk	IIM, PM	RA (76), SLE (76), SSc (101) (76)		(5, 75, 76)
FAM167A|BLK (chr8)	Risk	IIM, DM	RA (76, 96), SLE (76, 97), SSc (76, 101)		(3, 5, 75, 76)
	Protective	PM			
IL18R1 (chr2)	Protective	PM			(5)
SLC26A1|IDUA (chr4)	Risk	PM			(5)
RGS1 (chr1)	Protective	PM	CeD (100), MS (103), T1DM (99)		(5)
ROPN1L|ANKRD 33B (chr5)	Risk	DM			(5)
PTTG1|ATP10B (chr5)	Risk	DM			(5)
GSDMB (chr17)	Risk	DM	IBD (95), MS (103), RA (96), T1DM (99)		(5)
IFI35 (chr17)	Risk	IIM			(72)
PRDX3 (chr10)	Risk	IIM			(72)
SLAMF1 (chr1)	Risk	IIM			(72)
ZFAT (chr8)	Risk	IIM			(72)
PTPN6 (chr12)	Risk	IIM			(72)
CD1C (chr1)	Risk	DM			(72)
IRF5 (chr7)	Risk	IIM	RA (76), SLE (76), SSc (76)		(76)
TNFAIP3 (chr6)	Risk	IIM	SLE (76), SSc (76)		(76)
IL12RB2 (chr1)	Protective	IIM	SLE (76), SSc (76)		(76)
PHTF1-RSBN1 (chr1)	Risk	IIM	RA (76), SLE (76)		(76)
AP4B1 (chr1)	Protective	IIM	RA (76), SLE (76), SSc (76)		(76)
TNFSF4-LOC100506023 (chr1)	Risk	IIM	SLE (76), SSc (76)		(76)
NCF2 (chr1)	Risk	IIM	SLE (76)		(76)
DNASE1L3 (chr3)	Risk	IIM	SLE (76), SSc (76)		(76)
KPNA4-ARL14 (chr3)	Risk	IIM	RA (76), SLE (76), SSc (76)		(76)
TNIP1 (chr5)	Risk	IIM	RA (76), SLE (76), SSc (76)		(76)
SCT-DRD4 (chr11)	Protective	IIM	RA (76), SLE (76), SSc (76)		(76)
PTPN11 (chr12)	Risk	IIM	SLE (76), SSc (76)		(76)
IRF8 (chr16)	Protective	IIM	SLE (76), SSc (76)		(76)
TYK2 (chr19)	Protective	IIM	SLE (76), SSc (76)		(76)
PRP12 (chr19)	Risk	IIM	SLE (76), SSc (76)		(76)

Continued on next page

Table 2. The genetic alleles and alterations associated with IIM in Caucasian populations and their associations with other autoimmune diseases and common B-cell lymphoma subtypes (Continued)

Genetic allele/aleration	Effect	Subtypes/autoantibodies	Autoimmune diseases	B-cell lymphomas	Ref
SDK2 (chr17)	Risk	IIM			(75)
LINC00924 (chr15)	Protective	IIM			(75)
TEC (chr4)	Risk	IIM	RA (96)		(75)
PLCL1 (chr2)	Risk	IIM, DM	SLE (104, 105)		(3, 75)
	Protective	DM			
LTBR (chr12)	Protective	IIM			(75)
CCR5 (chr3)	Protective	IBM			(75)
CCL21 (chr9)	Protective	DM	RA (87)		(3)
47,XXY X chromosome aneuploidiesn	Risk	IIM	SLE (106), SSc (107)		(107)

IIM: Idiopathic inflammatory myopathy; DM: Dermatomyositis; ASyS: Antisynthetase syndrome; IBM: Inclusion body myositis; PM: Polymyositis: JDM: Juvenile dermatomyositis; RA: Rheumatoid arthritis; SLE: Systemic lupus erythematosus; SSc: Systemic sclerosis; SS: Sjogren Syndrome; MS: Multiple sclerosis; IBD: Inflammatory bowel diseases; T1DM: Type 1 diabetes mellitus; AITD: Autoimmune thyroid disease, CeD: Celiac disease; MG: Myasthenia gravis; DLBCL: Diffuse large B-cell lymphoma; FL: Follicular lymphoma; CLL: Chronic lymphocytic leukemia; MZL: Marginal zone lymphoma; Chr: chromosome; Ref: Reference.
Boldface indicates independent association; Italic indicates suggestive level of significance.

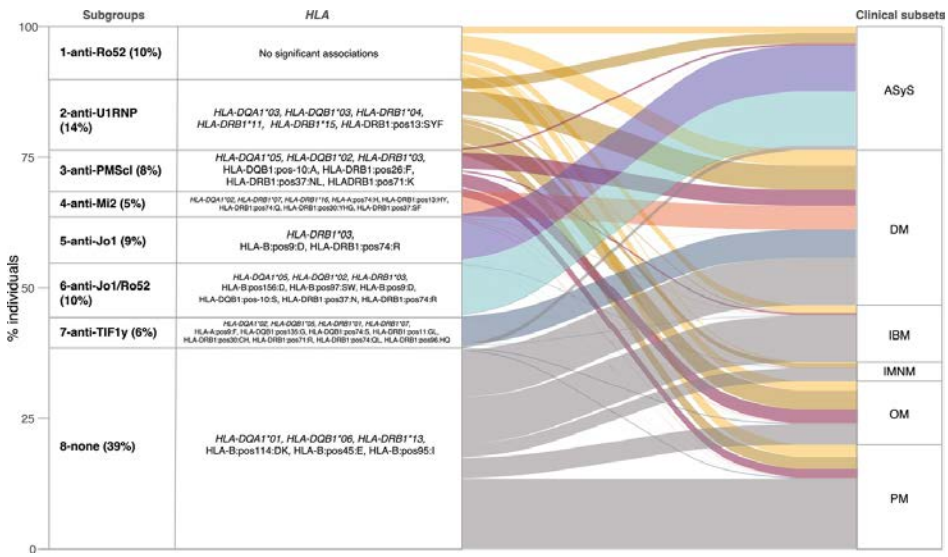


Figure 2. The subgroups of IIM based on autoantibody profiles, HLA and amino acid significant associations, and their correspondence to the clinical/pathological subsets. *IIM: Idiopathic inflammatory myopathy; HLA: Human leukocyte antigen; AsyS: Antisynthetase syndrome; IBM: Inclusion body myositis; JDM:Juvenile dermatomyositis; PM: Polymyositis; DM: Dermatomyositis; IMNM: Immune-mediated necrotising myopathy; OM: Overlap myositis. (Reproduced with permission from Elsevier: Leclair V, Galindo-Feria AS, Rothwell S, Krystufkova O, Zargar SS, Mann H, et al. Distinct HLA associations with autoantibody-defined subgroups in idiopathic inflammatory myopathies. EBioMedicine. 2023;96:104804.)*

1.1.4.2 Environmental factors

Various environmental factors have been linked to IIM but only a few of these associations have been replicated in population-based studies.

An association with ultraviolet (UV) radiation has been consistently found in DM and in serological subgroup of anti-Mi2 autoantibodies (108-112).

A higher frequency of smoking has been noted in patients with anti-Jo1 autoantibodies and *HLA-DRB1*03:01*, suggesting smoking may contribute to presence of anti-Jo1 autoantibodies through interacting with this risk allele (113).

Environmental pollutants mainly silica and dust have been suggested to trigger IIM (114-116). A Swedish cohort study found a two-fold increased risk of RIDs including SLE, SSc and DM in male workers with occupational exposure to silica compared to those unexposed (114). Other study reported an increased risk of having SSc, vasculitis or IIM in US discharge military veterans exposed to inorganic dust compared with those unexposed and a dose-dependent effect with increase in length of service was also observed (117). Furthermore, the presence of anti-Jo1 autoantibodies has been frequently reported in IIM cases exposed to occupational pollutants (116, 118-122).

Infection is considered a potential risk factor of IIM given the observed seasonal pattern of IIM onset in some MSA groups (i.e., anti-TIF1γ, Jo1 and MDA5 autoantibodies) and in groups of patients with specific HLA alleles (i.e., *HLA-DRB1*03:01* and *HLA-DQA1*05:01*) (123-127). The association between IIM and infection has been further supported by an elevated risk of adult IIM in individuals with respiratory tract or gastrointestinal infections found in a Swedish population-based study (128). Severe acute respiratory syndrome coronavirus 2 is another potential trigger of IIM (129-131). It has been observed that coronavirus diseases 19 presents similar pathophysiological features to autoimmune anti-MDA5 syndrome (130).

Pharmacological treatments may also trigger IIM and the linked medications are statins and immune checkpoint inhibitors (ICIs). The use of statins has been implicated in various IIM subtypes, exhibiting a strong association with anti-HMGCR autoantibodies in IMNM and weaker associations with other subtypes, including IMNM with anti-SRP autoantibodies, DM, IBM, and PM (132-137). It is well-known that the administration of ICIs may lead to immune-related adverse events including myositis, myasthenia gravis (MG) and myocarditis as they function to boost the immune system against tumours (138-141). Importantly, even though lymphocytic infiltration is presented in muscle biopsy of patients with ICI-associated myositis (138), characteristics like concurrence with myasthenia gravis or myocarditis, infrequent positivity of MSAs/MAAs, high mortality and resolution after steroid and intravenous immunoglobin treatments highlight a distinct pathological mechanism different from that of IIM (138, 140, 141).

1.1.5 Disease mechanisms

The pathogenesis of IIM is not fully understood but there is a growing body of evidence suggesting that IIM can be triggered by environmental risk factors in genetically predisposed individuals. The pathogenesis of IIM involves both immune-mediated and non-immune-

mediated mechanisms. The important roles of T and B cells, and interferon (IFN) pathway in the pathogenesis of IIM, particularly DM, are typical examples of immune-mediated mechanisms (142-144). Non-immune-mediated mechanisms related to IIM include endoplasmic reticulum stress response, autophagy and hypoxia (53, 145). For example, it has been suggested that accumulation of misfolded amyloid β in IBM can lead to muscle cell dysfunction by production of reactive oxygen species and this damage is further amplified via endoplasmic reticulum stress response (146).

1.1.6 Mortality

Patients with IIM still have a higher mortality compared to the general population although improvement in survival has been observed in recent years (10, 35, 147-151). The five-year survival ranges from 59.6% to 82.8% (10, 35, 147, 148). The mortality rate has been found to be the highest shortly after IIM diagnosis. Compared to the Swedish general population, there was an approximately 10-fold higher risk of death following the first year of IIM diagnosis (10). The excess mortality in patients with IIM can be attributed to commonly associated comorbidities including cancer (10, 35, 152, 153), respiratory diseases (interstitial lung disease, ILD in particular) (10, 35, 152-160), cardiovascular diseases (10, 35, 161) and infections (152, 157, 158, 162-165).

1.2 IIM in the context of cancer

1.2.1 Epidemiology

The occurrence of cancer in patients with IIM was firstly described in 1916 (166) and recognised as a comorbid condition in DM and PM by Bohan and Peter in 1975 (48, 49). The association between IIM and cancer has become well-established after the publications of well-conducted population-based studies released in late 20th to early 21st century (8, 167, 168). Cancer may occur concurrently with or any time before or after IIM diagnosis, but it is frequently observed within three years before and after IIM diagnosis, which is a common timeframe to define cancer-associated myositis (CAM) (169). However, there is no consensus on the definition of CAM. Some investigators extended the timeframe to five years after IIM diagnosis (167, 170, 171) or included no time restriction (172). This discrepancy together with differences in study design and year, IIM ascertainment, IIM subtypes being studied and ethnic background lead to a significant variation in frequency of cancer measured in patients with IIM. The prevalence of cancer in patients with IIM from two tertiary centers in Riyadh, Saudi Arabia was 6.7% while it was 78% in a hospital-based study including DM or ADM patients positive to anti-TIF1γ autoantibodies (173, 174). In population-based studies conducted in Nordic countries, Scotland, Korea and Taiwan, the overall prevalence of cancer in patients with IIM varied from 11% to 24% (8, 9, 168, 175-177). Incidence of cancer after IIM diagnosis was commonly reported as a proportion or a rate ratio in previous studies, with the risks varying from 10% to 17% (167, 178-181), and the rate ratios ranging from 2 to 7.7 for patients with DM and 1.3 to 2.1 for patients with PM (8, 20, 168, 175). Our previous work reported that the incidence rate of cancer three months after IIM diagnosis was 21.1/1000 person years, which was 1.7 times higher than the general population (9).

The risk of cancer in IIM is time dependent. The risk can rise as soon as three years before and usually peak within the year of IIM diagnosis, when it can be up to 25-fold higher than the general population, and then decrease gradually but can remain significantly elevated more than 10 years after IIM diagnosis (8, 9, 20, 178, 180).

1.2.2 Associated IIM subtypes

The association with cancer has been observed in several IIM subtypes. The strongest association is consistently found in DM, followed by PM. In a meta-analysis study including five large-scale observational studies, the pooled SIR was 4.7 (95% CI 3.3-6.5) and 1.8 (95%CI 1.4-2.3) for DM and PM, respectively (180). Moreover, patients with DM are also more likely to have cancer history compared to patients with other IIM subtypes (8, 9). In the Swedish population, the OR of cancer one year before to within 90 days after IIM diagnosis was 12.1 (95%CI 7.1-20.8) and 2.3 (95%CI 1.4-3.5) for DM and other IIM, respectively (9). There is also evidence supporting an association with cancer for CADM (179, 182), IMNM (183, 184) and IBM (178, 185, 186), though with conflicting findings (175, 187). The increased risk of cancer in CADM and IMNM may be attributed to the presence of anti-TIF1γ autoantibodies and anti-HMGCR autoantibodies, respectively (182-184). There is little evidence indicating an increased risk of cancer in patients with JDM, ASyS and OM (44, 45, 176, 188).

1.2.3 Associated site-specific and histological subtypes of cancer

Various types of cancer have been reported in patients with IIM and the associated subtypes of cancer seem to be in line with the population cancer risk. For example, an elevated risk of nasopharyngeal cancer (NPC) is frequently observed in patients with IIM from Asian countries including China (189-193) and Singapore (194) while patients from Korea are more likely to have NPC and stomach cancer (195, 196). Patients with IIM from Taiwan have been found to be associated with an increased risk of brain cancer, but this finding was based on cases less than five. Breast, prostate, ovarian, stomach, colorectal, kidney, bladder and skin cancers are commonly seen in patients with IIM from Nordic countries (8, 9, 167, 197), Scotland (168), Hungary (198), Australia (185) and the US (199) where these cancer types are more prevalent. Lung and haematological cancers are associated with IIM in both Asian and Western populations (8, 189, 196).

Most of the associated cancer subtypes are more likely to occur around the time of IIM diagnosis. For instance, elevated risks of lung, stomach, colorectal and ovarian cancers, as well as haematological malignancies are usually found close to the time of IIM diagnosis and seldomly observed more than 5 years after (8, 9, 168, 200). Specifically, in our previous work, we found increased risks of lung, colorectal and ovarian cancers only before IIM diagnosis while increased risks of oropharyngeal, cervical and skin cancers were observed only 90 days after IIM diagnosis (9). However, a Nordic study found that an increased risk of colorectal cancer could last up to more than 5 years following DM diagnosis (8).

There is no consistent pattern of the associated cancer subtypes found between IIM subgroups. However, increased risks of lung, ovarian, cervical, stomach and colorectal cancers are more likely observed in patients with DM than in patients with PM in Western countries (8, 167, 168). Among patients with anti-TIF1γ autoantibodies positive DM from the UK, breast and

ovarian cancers, and lymphoma are overrepresented (201). A single center study in China also observed similar findings (202) while another Chinese single center study found that NPC, breast and lung cancers were overrepresented (203). Furthermore, when considering cancer association by histology, DM is associated with squamous cell carcinoma, adenocarcinoma and haematological malignancies while PM is only associated with haematological malignancies (8, 9).

1.2.4 Predictors

Patients affected by IIM and cancer are associated with specific features. Meta-analysis studies investigating predicators of cancer by comparing IIM patients with cancer and those without showed that old age at IIM diagnosis, male sex, rapid disease onset of IIM, skin manifestations especially cutaneous necrosis, dysphagia, lower levels of CK and lactate dehydrogenase, elevated levels of C-reactive protein and erythrocyte sedimentation rate (≥ 35 mm/h) and presence of anti-TIF1γ autoantibodies were associated with an increased risk of cancer, while ILD, Raynaud's phenomenon, arthritis, presence of anti-Jo1, anti-EJ or any anti-aaRS autoantibodies were protective factors of cancer (204-206). Of these factors, old age, male sex, dysphagia, ILD, anti-TIF1γ autoantibodies and anti-aaRS autoantibodies were suggested to be established predictors of cancer in a meta-analysis including the largest number of studies (n=67) (206). Importantly, patients with IIM diagnosed at young age and female patients are also at a higher risk of developing cancer compared to the general population, although old age and male sex may confer a higher risk (9, 176).

There are more predictors that have been suggested to increase risk of cancer in IIM; however, these associations remain to be verified. There is evidence from multivariate analyses suggesting low baseline level of complement 4, neutrophil to lymphocyte ratio > 5.5, low level of alanine transaminase and absence of fever as risk factors of cancer in patients with IIM (171, 193, 207). Tumour marker CA125 has also been suggested to be associated with cancer in IIM but its predictability is not promising (206, 208, 209). Furthermore, associations between anti-HMGCR, anti-NXP2, anti-SAE and anti-calreticulin (CRT) autoantibodies, MSA negative, obesity and cancer have been observed (183, 184, 188, 202, 210-214), although conflicting findings for some of these factors exit (193, 204, 206, 208).

Little is known if there is genetic variant associated with cancer in patients with IIM. Limaye et al. observed higher frequency of *HLA-A28* in patients with IIM and cancer than those without cancer (11% versus 2%) (185). HLA-Cw ^{90}AGSHTLQWM^{98} protein binding motif sequence has been reported to positively associated with CAM (69). The *HLA-DRB1*03:01* allele, which is strongly associated with IIM is however not associated with cancer in patients with IIM (69, 198).

1.2.4.1 Utility of MSAs in cancer prediction

The predictability of anti-TIF1γ autoantibodies and negativity of other IIM-related autoantibodies has been examined. A meta-analysis study reported a pooled sensitivity of 78% (95% CI 45-94%) and a specificity of 89% (95% CI 82-93%) when using the presence of anti-TIF1γ autoantibodies to predict cancer in adult patients with DM or ADM (215). A UK study with a large patient cohort found that the combination of anti-TIF1γ autoantibodies

positivity and absence of anti-Jo1, anti-Ku, anti-PM/Scl, anti-U1RNP and anti-U3RNP autoantibodies detected via hospital-based routine immunology testing had high sensitivity (93.8%) and negative predictive value (99.2%) but low specificity (44.7%) and PPV (9.3%) (216).

1.2.4.2 Cancer predictors in patients with anti-TIF1γ autoantibodies

Not all patients with anti-TIF1γ autoantibodies positive DM experience cancer although anti-TIF1γ autoantibodies are strong risk factor of cancer. Among patients with anti-TIF1γ autoantibodies positive DM, up to 62% of them could be free of cancer (201, 203). Identifying factors associated with cancer in patients with anti-TIF1γ autoantibodies positive DM is important to inform cancer assessment. A Chinese study including 87 patients with anti-TIF1γ autoantibodies identified 14 clinical variables that could result in good performance of cancer predication based on different machine learning algorithms (area under the receiver operating characteristic curve ranged from 88% to 100% in the training samples and from 70% to 91% in the testing samples) (203). Of these clinical variables, disease duration had the highest importance in cancer prediction, followed by percentage of lymphocytes, percentage of neutrophils, neutrophil-to-lymphocyte ratio (NLR), sex, C-reactive protein, shawl sign, arthritis/arthralgia, V-neck sign, anti-PM/Scl75 autoantibodies, deterioration of general condition, Raynaud's phenomenon, fever and anti-Jo1 autoantibodies. Furthermore, several novel autoantibodies or biomarkers have been reported to be associated with cancer in patients with anti-TIF1γ autoantibodies. Anti-Sp4 and anti-CCAR1 autoantibodies concurrent with anti-TIF1γ autoantibodies in patients with DM have been shown to be associated with a reduced risk of cancer (217, 218). A significant higher serum level of galectin-9, an immune checkpoint protein, has been observed in DM patients with newly onset and untreated cancer than in DM patients with treated cancer, DM patients without cancer and cancer patients without IIM (219). Anti-TIF1γ autoantibodies in combination with the presence of serum galectin-9 showed good predictability of cancer (area under the receiver operating characteristic curve=89%, 95% CI 80%-98%) in patients with DM (219).

1.2.5 Pathogenesis

Little is known the pathological link between IIM and cancer, but some hypotheses have been suggested based on evidence from mainly epidemiological studies. The mainstream proposed hypotheses are cancer-induced autoimmunity and autoimmunity-induced cancer. These two hypotheses, together with shared pathogenesis as the third hypothesis are introduced in the following sections.

1.2.5.1 Cancer-induced autoimmunity

IIM in some cases, particularly DM, can be considered as a paraneoplastic symptom triggered by the internal malignancy in an individual. Evidence supporting this proposed hypothesis includes firstly close temporal relationship between IIM and cancer and association between active disease activity of IIM and cancer (180, 196, 204). Concordant disease courses of IIM and cancer is frequently observed, for example, IIM symptoms can be resolved after successful cancer treatments or surgeries (198, 220-224). Furthermore, DM patients with

cancer associated MSAs including anti-TIF1γ, anti-NXP2 and anti-SAE1 autoantibodies usually have cancer detected close to IIM onset (188, 201, 225). Two studies also found that titre of anti-TIF1γ autoantibodies were correlated with the disease courses of DM and cancer, and patients had a lower level of anti-TIF1γ autoantibodies or became anti-TIF1γ autoantibodies negative after control of cancer (172, 226). Importantly, there is increasing evidence supporting the important role of TIF1γ protein in the development of paraneoplastic DM. TIF1γ protein has been found to overexpress in tumour cells, muscle and skin tissues of DM patients with cancer (227) and in regenerating muscle cells (228). Features of muscle regeneration have been found in patients with newly diagnosed cancer (229). Furthermore, somatic mutations and loss of heterozygosity of *TIF1* genes have been found in tumours of patients with cancer-associated DM positive to anti-TIF1γ autoantibodies (230, 231). Together, this evidence leads to a hypothesis where anti-tumour immunity against neoantigen (mutated self-protein produced during tumorigenesis) may cross-react with the wild-type self-protein expressed in regenerating muscle tissues, leading to muscle damage and onset of IIM (227, 232).

Cancer may also indirectly lead to IIM development via anti-cancer treatment. Hormone therapy aromatase inhibitors, one of the primary treatments for breast cancer, have been suggested to be associated with ASyS and other autoimmune diseases including lupus, SSc (233-235). This suggested association is however based on findings from case reports. As abovementioned, onset or flare of IIM has been observed in patients treated with ICIs (138-141). However, the pathogenesis of ICI-associated myositis remains unclear. It is possible that ICIs may contribute to IIM development via enhancing T cells activation which has an important role in the pathogenesis of IIM (143, 236).

1.2.5.2 Autoimmunity-induced cancer

The long-term risk of cancer observed in patients with IIM may suggest that in some cases, especially IIM subtypes other than DM, the autoimmune conditions may lead to cancer (9). IIM is a chronic condition, and its development involves chronic inflammation (53). Although little is currently known if the underlying chronic inflammation of IIM may contribute to the onset of cancer years following IIM diagnosis, it is well-known that chronic inflammation favours formation of tumour microenvironment (237). Long-term use of immunosuppressive treatment in patients with IIM may also pose risks of developing cancer, experiencing cancer recurrence and second cancer but evidence is lacking in patients with IIM. Relevant knowledge is mainly learnt from studies of patients with other autoimmune diseases or patients experienced solid organ transplantation. Glucocorticoids are the first-line treatment for IIM and its long-term use is needed for patients with moderate to severe IIM (53). A meta-analysis found no higher risk of malignancy in kidney transplant patients using steroids versus those discontinuing steroids after five years of follow-up (238). However, corticosteroid in combination with cyclophosphamide has been reported to be associated with an increased risk of cancer in patients with SLE (239). Moreover, azathioprine, the first line nonsteroid treatment for IIM, has been found to be associated with skin cancer (240, 241).

1.2.5.3 Shared pathogenesis

Adopting a broader perspective and moving beyond direct causality, IIM and cancer may have common pathological pathways given that immune system plays a central role in the developments of both conditions. Song *et al*. conducted a meta-analysis of gene expression data of IIM and controls derived from muscle and skin tissues from public databases (145). The authors identified top 10 up-regulated differentially expressed genes associated with IIM and these genes were enriched in IFN signalling pathway, unfolded-protein responses and protein secretion. It is already known that IFN signalling pathway has an important pathogenic role in DM, ASyS and IBM and its role in CAM has also been proposed in a review paper written by Selva-O'Callaghan *et al*. (144, 232). Unfolded-protein response signalling pathway has been recognised as a regulator of tumorigenesis (242). Moreover, the top up-regulated gene *TSMB10* reported in the abovementioned meta-analysis study has been found to overexpress in breast and liver cancers (145, 243, 244).

1.2.5.3.1 Familial component

Studying familial co-aggregation of IIM and cancer is an useful approach to investigate if IIM and cancer share familial (genetic and environmental) factors. A study conducted in 1997 found no significant difference in frequency of family history of cancer between patients with newly onset JDM, patients with juvenile RA and healthy controls (245).There are also a few studies examining familial co-aggregation of IIM and haematological malignancies but none of them detected statistically significant associations. The earliest study published in 2000 failed to observe any cases of lymphoma and leukaemia in 511 offspring of 348 parents with DM/PM (246). Engels *et al.* later reported that patients with non-Hodgkin lymphoma (NHL) were more likely to have family history of DM compared to controls (0.9% versus 0%, p=0.02) (247). Landgren *et al.* using Nordic nationwide register data to examine if family history of DM/PM was associated with chronic lymphocytic leukaemia (CLL), HL and multiple myeloma (MM) among first-degree relatives (248-250). The OR was 0.9 (95%CI 0.5-1.6) for CLL, 1.5 (95%CI 0.8-2.8) for HL and 1.0 (95%CI 0.7-1.5) for MM. A Swedish population-based study with more recent data observed a similar finding for MM and an insignificant negative association for HL but the estimation was based on only one HL case. This study also found no significant familial association between DM/PM and NHL overall (SIR=0.8, 95%CI 0.4-1.4), as well as for diffuse large B-cell lymphoma (DLBCL) (SIR=2.1, 95%CI 0.8-4.3) and follicular lymphoma (FL) (SIR=1.2, 95%CI 0.2-3.6) (251). These findings overall do not support that familial component plays an role in the co-occurrence of IIM and haematological malignancies. However, given the limitations such as using only inpatient data to define IIM and defining IIM based on only one diagnostic visit in the abovementioned studies (248-251), further investigation is needed. Moreover, prior to **study III**, no research had investigated familial co-aggregation between IIM and solid cancers.

1.2.5.3.2 Shared genetic susceptibility

No study has published to date to investigate shared genetic susceptibility between IIM and cancer. Nevertheless, there are genes or genetic variants associated with IIM and cancer reported in independent GWASs of these diseases. For example, though with different associated variants, HLA class I genes associated with IIM are also related to various types

of cancer including lung and cervical cancers, and haematological malignancies (4, 90, 91, 252-255). Furthermore, several IIM-associated genes outside the HLA region such as *GSDMB*, *STAT4*, *PTPN22*, *PHTF1-RSBN1* are linked with diverse types of cancer (5, 75, 76, 256-261). Furthermore, Song *et al*. identified several expression quantitative trait loci potentially associated with IIM via in silico analysis, one of which is mapped to the *AGER* gene (145). Genetic variant of the *AGER* gene has been found to be associated with increased neutrophil-lymphocyte ratio and mortality in lung cancer (262). Notably, higher neutrophil-lymphocyte ratio has been found as a risk factor of cancer in patients with IIM (203, 207). These findings suggest shared pathogenesis between IIM and cancer from a genetic perspective.

1.2.5.3.3 Shared environmental factors

There are no confirmed environmental risk factors associated with IIM in the context of cancer. However, suggested risk factors of IIM including smoking, UV radiation and infections are well established risk factors of various cancer types (263, 264). With data from the MYONET registry, prevalence of smoking was higher in IIM patients with cancer (51%) than those without (37%) (45). Yet, an Australian study including 80 patients with IIM and cancer found similar frequency of cancer between never smokers (9.2%) and current or former smokers (10.8%) (185).

1.2.6 Cancer screening

In 2022, the first risk-based cancer screening approach for IIM, formulated based on available evidence of risk factors of cancer and clinical experiences, was introduced and it was further refined in a recent publication (265, 266). This approach primarily targets patients who have adult-onset IIM diagnosis within a three-year window of time and risk levels of cancer are categorized as low, moderate or high based on IIM subtypes, clinical and serological factors (**Figure 3**). It is recommended to perform age- and sex-appropriate cancer screening as per local guidelines to all patients with IIM regardless of risk level of cancer. Furthermore, patients at low risk group (one moderate risk factor or ≥ one low risk factor) should undergo basic cancer screening at IIM diagnosis. At moderate risk group (≥ 2 moderate risk factors or one high risk factor), patients should undertake basic and enhanced cancer screening panels at IIM diagnosis. At high risk group (≥ two high risk factors), patients should undergo basic and enhanced cancer screening panels at IIM diagnosis plus annual basic cancer screening for three years based on the overall risk assessment. Moreover, in cases where basic and enhanced cancer screening panels reveal no signs of cancer in patients at high risk group, clinicians may consider using ^{18}F-fluoro-deoxy-glucose positron emission tomography-computed tomography as an additional cancer screening method, particularly when there are suspicions of lymphoma. Importantly, no study thus far has evaluated the benefits gained from this risk-based cancer screening approach.

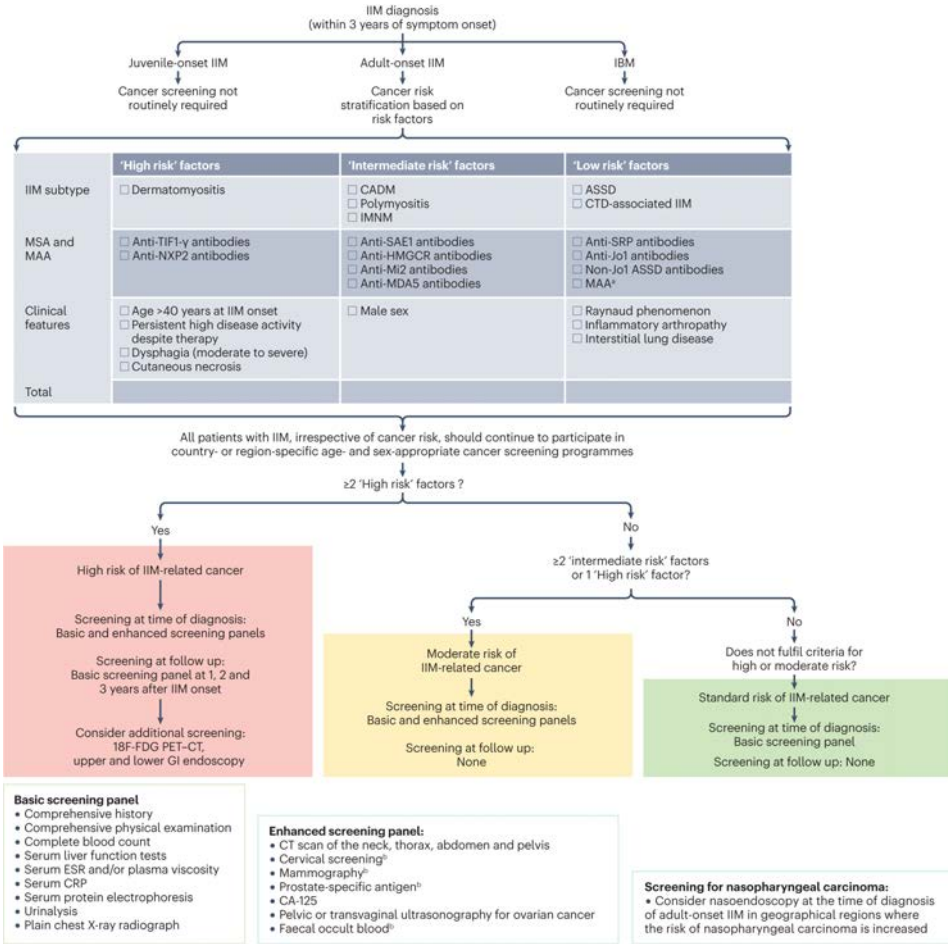


Figure 3. Risk stratification and frequency of screening for IIM-related cancer. *ASSD: Antisynthetase syndrome; CADM: Clinically amyopathic dermatomyositis; HMGCR: 3-hydroxy 3-methylutaryl coenzyme A reductase; IBM: Inclusion body myositis; IMNM: Immune-mediated necrotizing myopathy; MDA5: Melanoma differentiation-associated gene 5; NXP2: Nuclear matrix protein 2; RNP: Ribonucleoprotein; SAE1: Small ubiquitin-like modifier-1 activating enzyme; SRP: Signal recognition particle; TIF1γ: Transcription intermediary factor 1 γ. [a]Anti-PM/Scl, anti-Ku, anti-RNP, anti-SSA/Ro, anti-SSB/La autoantibodies. [b]If not already part of country/region-specific age- and sex-appropriate cancer screening programmes. (reproduced with permission from Springer Nature Oldroyd AGS, Callen JP, Chinoy H, Chung L, Fiorentino D, Gordon P, et al. International Guideline for Idiopathic Inflammatory Myopathy-Associated Cancer Screening: an International Myositis Assessment and Clinical Studies Group (IMACS) initiative. Nat Rev Rheumatol. 2023;19(12):805-17.)*

1.2.7 Disease management

Disease management for patients with IIM and cancer is challenging as effective strategy is lacking to maintain a balance between IIM and cancer treatment. The current pharmacological treatment to control IIM in patients with CAM is a combination of prednisone (1 mg/kg/day), cyclosporine (up to 5 mg/kg/day) or tacrolimus (0.06 mg/kg/day) and intravenous immunoglobulins (2 g per kg every 4-6 weeks). Moreover, IIM medications can interact with many antineoplastic agents and therefore there should be additional attention to drug-drug interaction when prescribing treatments to patients affected by IIM and cancer. Similar to the health management advices given to the general population,

patients with IIM are recommended to avoid smoking, alcohol consumption and UV exposure, to have physically active life and equilibrated Mediterranean diet, and to follow vaccination programmes against infections (267).

1.2.8 Prognosis

1.2.8.1 Cancer recurrence and second cancer

Our understanding of cancer recurrence in patients with IIM is primarily based on case reports that often present IIM as a paraneoplastic syndrome associated with cancer relapse (170, 268-281). In some of these cases, IIM was detected when a previously diagnosed primary cancer reoccurred (170, 268-270, 273, 274, 280, 281). In other cases, IIM was presented as the first manifestation that led to a diagnosis of a primary cancer during cancer examination, subsequently, cancer recurrence was detected when IIM flared during follow-up (271, 272, 275-279). Prospective epidemiological studies investigating the risk of cancer recurrence in patients with IIM who later get cancer is lacking. In a systematic review of 48 case reports including 110 patients with DM/PM and ovarian cancer, 31 patients (28.2%) experienced cancer recurrence within a median follow-up of 24.5 months (ranges 5-210 months) (282). Furthermore, about two-third of these cases had DM as the first manifestation (282). Similar findings were reported in another systematic review of 27 case studies of patients with DM and colorectal cancer, with a frequency of cancer recurrence of 19.2% after a median follow-up of 7.9 months (ranges 2-21 months) (283). Little is currently known what factor is associated with cancer recurrence in IIM. The majority of abovementioned cases reported cancer recurrence in patients with DM but we cannot draw any conclusive interpretation as these reported cases could be highly selective (170, 268, 269, 271-283). Another potential factor associated with cancer recurrence in IIM is anti-CRT autoantibodies. In a Chinese single center study including 469 patients with IIM, the positivity rate of anti-CRT autoantibodies was significantly higher in IIM patients with cancer recurrence (62.5%) than those with cancer in remission (18.2%) (214). Moreover, the frequency of cancer recurrence in patients with anti-CRT autoantibodies was also higher than in patients without anti-CRT autoantibodies (214). There is also a study reporting increased level of anti-TIF1γ autoantibodies in DM patients with cancer recurrence, but findings were based on a few number of patients ($n < 5$) (226).

Besides cancer recurrence, patients with IIM may also experience multiple primary cancers. However, previous epidemiological studies have predominantly focused on the risk of ever cancer diagnosed within three years of IIM diagnosis or the risk of cancer after IIM diagnosis, leading to restricted understanding of the risk of a second cancer in IIM. According to the descriptive data from population-based studies, as well as single and multicentre cohorts, the frequency of IIM patients with more than one primary cancer among all included IIM patients ranged from 0.2% to 3.2% (168, 170, 176, 226). Among IIM subtypes, the frequency of having a second cancer in patients with PM varied from 0.3% to 0.7%, versus 0.1% to 0.4% in patients with DM (168, 176).

1.2.8.2 Mortality

Cancer is one of the major comorbidities associated with IIM and it worsens patients' prognosis. Patients affected by IIM and cancer have poorer survival than the general population, cancer patients without IIM and patients with primary IIM. Compared to the general population, patients with IIM have a four-fold higher risk of death from cancer than the Swedish general population (10). Increased mortality has also been found in cancer patients with CTD compared to those without CTD (284-286). The survival rate of lung cancer patients with DM/PM was much lower than that of lung cancer patients without any rheumatic diseases (285). A Swedish study also found an increased mortality in lung cancer patients with DM/PM compared to the cancer-specific mortality rate in the general population (SIR=4.2, 95% CI 3.0-5.6) (284). Furthermore, compared to patients with primary IIM, two cohort studies reported lower one- and five-year overall survival rates in patients with CAM (193, 198). In a Chinese large multicentre cohort, the one- and five-year survival rates of CAM were 84% and 76% versus 94% and 92% for primary IIM (193).

Study investigating mortality between DM and PM patients with cancer is scarce and findings are inconsistent. Wakata *et al.* reported a one-year survival rate of 80% in DM patients with cancer versus 100% in primary DM while it was 50% in PM patients with cancer versus 100% in patients with primary PM (287). A more recent study with larger sample size however found lower one-year survival rate in DM patients with cancer (49%) compared to 69% in PM patients with cancer while at five-year survival, a more reduced survival rate was observed in PM compared to DM (15% versus 28%) (288).

The length of time interval between IIM and cancer diagnoses may also have an impact on mortality. It has been found that patients with cancer concurrent with active IIM had shorter mean survival time (50.2 months, 95%CI 28.2-71.6 months) than patients with cancer diagnosed during inactive myositis (219.1, 95%CI 165.5-272.7 months), corresponding to a hazard ratio of 4.3 (95%CI 1.5-12.7) (196). Similar findings were observed in two other studies when comparing patients with CAM to those with non-CAM (177, 188). Other factors potentially associated with an increased mortality in patients with IIM and cancer include male sex, advanced cancer stage and cutaneous necrosis (196, 210, 289). For example, the cumulative one-year survival rate was the lowest for male patients with advanced cancer (69%) compared to patients with either of risk factor (93%) and patients with no risk factor (100%) (210).

Malignancy and respiratory diseases are common causes of death reported in patient with IIM and cancer, but evidence is mainly from descriptive data. Two studies found cancer as the major cause of death in patients with IIM and cancer (193, 196, 289-293), while other studies reported ILD as the primary cause of death (215, 287) and one study found even contribution between ILD and cancer (294).

1.3 Specific background of Study I and II

1.3.1 Inherited IIM mimics

Inherited muscular diseases such as metabolic myopathies (i.e., McArdle disease and Pompe disease) and muscular dystrophies (i.e., fascioscapulohumeral muscular dystrophy and

dysferlinopathies) are common IIM mimics. Given the shared clinical and histopathological features with IIM, particularly PM, misdiagnosing inherited muscular diseases as IIM is not uncommon. Misdiagnosing of inherited muscular diseases as PM usually happen when distinct disease features are not noticed and there is a late disease onset accompanied with proximal muscle weakness, elevated CK level and inflammatory infiltrate on muscle biopsy, characteristics that fulfil the Bohan and Peter criteria for PM (295-298). In a French study including 40 patients with genetically confirmed dysferlinopathies, 10 of the patients (25%) were misdiagnosed as PM in early diagnostic work-up. However, in a screening of non-inflammatory myopathies in 3,170 patients included in the MYONET registry, only 103 patients were excluded due to indications of IIM mimics or other reasons irrevalent to IIM mimics, corresponding to a PPV of 97% (45).

In Sweden, considerable measures have been implemented to reduce the risk of misdiagnosing inherited muscular diseases as IIM. These efforts involve high awareness of IIM mimics among rheumatologists, neurologists and neuropathologists, alongside thorough diagnostic evaluation and vigilant monitoring. In Sweden, patients with IIM generally have at least two follow-up visits within the year of initial diagnosis at specialised clinics for prognostic assessment. The first follow-up visit usually occurs within one to three months after the initial diagnosis. The number of visits increases and the time interval between follow-up visits is shortened when disease is severe or there is no beneficial effect from immunosuppressive treatments, which is a typical feature of inherited muscular diseases. Once no treatment response is found, patients are usually subject to a new diagnostic workup and a second muscle biopsy is often performed together with magnetic resonance imaging. Therefore, even if a misdiagnosis occurs, it is unlikely that the misdiagnosis remained uncorrected for an extended period in the Swedish healthcare system.

1.3.2 Polyautoimmunity, familial autoimmunity and shared genetic susceptibility with other autoimmune diseases

Autoimmune diseases have a common origin and therefore, co-occurrence of more than one autoimmune diseases in an individual is frequently observed and usually defined as polyautoimmunity (299). Polyautoimmunity has been observed in more than one third of patients with an autoimmune disease (299). The frequency of polyautoimmunity in patients with IIM can be reflected by the frequency of OM, ranging from 12% to 28% (45, 300). IIM may not only overlap with MCTD, SSc, SLE, SS and RA but also other autoimmune diseases including celiac disease (CeD), psoriasis, inflammatory bowel diseases (IBD), autoimmune thyroid diseases (AITD) and MG (60, 300-302). An increased prevalence of CeD has been observed in patients with adult IIM compared to the general population (4.5% versus 0.5%) (300). A population-based study in Taiwan found that patients with ulcerative colitis had a six-fold high risk of having DM than the general population (301). An increased risk of PM in UC was however not found. Furthermore, a strong association between IIM and MG has been reported (OR=21.0, 95%CI 5.8-75.3) (302).

IIM may cluster within families along with other autoimmune diseases. These autoimmune diseases include RA, SLE, SSc, Type 1 diabetes mellitus (T1DM) and AITD (59, 60, 62, 303-305), but conflicting findings exist (59, 60, 245, 306-308). There are also studies examining familial associations between IIM, and SS, multiple sclerosis (MS), IBD, CeD

and MG but found no significant associations, probably due to small number of IIM cases (59, 60, 245, 309-311). Knowledge of familial association between two diseases is helpful to understand the aetiology of both diseases and to guide genetic studies exploring genetic correlation between diseases.

Several attempts have been made to study genetic overlap between IIM and various autoimmune diseases and have successfully discovered numerous genetic variants associated with both IIM, RA, SLE and SSc (3, 5, 76). Moreover, insights drawn from independent genetic studies of IIM and other autoimmune diseases, including SS, T1DM, MS, IBD, CeD, AITD and MG, reveal share genetic variants with IIM (3, 5, 76-85, 87, 92, 94-101, 103-105). **Table 2** presents the genetic variants that have been reported to be associated with IIM and these autoimmune diseases. Among these autoimmune diseases, SLE and SSc share the highest number of genetic variants with IIM.

1.4 Specific background of Study IV

1.4.1 The risks of common B-cell lymphoma subtypes in IIM

Among haematological malignancies, only NHL has been found to be associated with both DM and PM and its association with IIM has been observed in patients from Western and Asian countries (8, 196). Further investigation has discovered positive associations between IIM and DLBCL and FL (312). Other study however observed no statistically significant association between IIM and DLBCL, FL, CLL and marginal zone lymphoma (MZL) (313).

1.4.2 Genetic variants associated with both IIM and common B-cell lymphoma subtypes

B-cell lymphomas are a group of malignancies with highly variable genetic alterations, pathological pathways, clinical courses, treatment strategies and responses. The most common subtypes, derived from different stages of B lymphocyte maturation, are DLBCL, FL, CLL and MZL (314).

The genetics of B-cell lymphomas involve both germline and somatic alterations, with the latter playing a more established role in the development of B-cell lymphomas (315). The introduction, however, addresses the germline alterations of B-cell lymphomas, aligning with the study aim of **Study IV**. The SNP-based heritability of DLBCL, FL, CLL and MZL is 9%, 16%, 24% and 8%, respectively (316). Only a few genetic variants associated with IIM are also linked to the common B-cell lymphoma subtypes (**Table 2**). The shared genetic variants are *HLA-B*08:01* and *EOMES* for DLBCL (90, 102), *HLA-DRB1*01*, *HLA-DRB1*13* and *HLA-DQA1*01:01* for FL (86, 89), *HLA-DRB1*07:01*, *PSMB9* and *EOMES* for CLL (88, 93), and *HLA-B*08:01* for MZL (91).

2 Research aims

This PhD thesis comprising five studies aims to improve our understanding of IIM and its associations with other autoimmune diseases and cancer from a genetic perspective, as well as to inform disease burden due to cancer events in patients with IIM. The study-specific aims are presented as follows:

Study I: Familial aggregation and heritability of IIM

Our knowledge of how likely IIM may aggregate within families and what extent genetics may contribute to IIM development is limited. **Study I** aimed to quantify the familial aggregation of IIM among first-degree relatives and the heritability of IIM using nationwide register data in Sweden.

Study II: Familial autoimmunity in patients with IIM

Autoimmune diseases have a common immunogenetic nature and clustering of multiple autoimmune diseases is frequently observed. However, evidence supporting familial associations between IIM and other autoimmune diseases is scarce and inconsistent. **Study II** aimed to investigate familial associations of a wide range of autoimmune diseases with IIM in comparison to the general population.

Study III: Familial co-aggregation of IIM and cancer

Study IV: Genetic overlap between IIM and common B-cell lymphoma subtypes

It is not known if IIM shares genetic susceptibility with cancer. **Study III** aimed to address this knowledge gap by exploring familial associations between IIM and a wide variety of cancer types in comparison to the general population. **Study IV** further addressed the research question and focused on exploring the extent of genetic overlap between IIM and four common B-cell lymphoma subtypes using GWAS summary statistics.

Study V: Impact of cancer on the prognosis of patients with IIM

Due to the rarity of the comorbid condition of IIM and cancer, our understanding of how cancer may affect the prognosis of patients with IIM, and the relevant prognostic factors remains limited. **Study V** aimed to examine the incidence of cancer events and cause-specific mortality by following patients since IIM diagnosis, as well as to determine factors associated with cancer and death events.

3 Materials and methods

3.1 Study setting

Sweden, a Nordic country with a land area of 407,284 km^2, has 10.6 million inhabitants in 2023. The Swedish healthcare system is largely funded by tax incomes (80%), making it is accessible and affordable to all residents with Swedish personal identity number (317).

The Swedish healthcare is delivered via primary and secondary healthcare sectors that are managed and operated by regions (n=21) and municipalities (n=290) under the healthcare policy determined by the state. Generally, primary care provided by general practitioners, psychologists or physiotherapists is the first point of contact to physical healthcare. An individual will be referred to secondary health care sectors when there is a needed (317).

Sweden has a well-established registration system to prospectively collect demographic and healthcare data at national scale. The unique identification number assigned to each legal resident enables linkage between multiple national registers, making the registration system a valuable and powerful asset in epidemiological research. The following section will introduce the registers used in this thesis in details.

3.2 Data sources

3.2.1 Swedish National Registers

3.2.1.1 The National Patient Register

The National Patient Register (NPR) prospectively collects data on inpatient visits since 1964 and outpatients visits since 1997. The coverage has reached a nationwide level of nearly 100% since 1987 for inpatient visits and approximately 80% since 2001 for outpatient visits. The lower coverage of outpatient visits is due to unregistered visits to private sectors. The NPR records information on date of visit, date of discharge (for inpatient visit), main and contributory diagnoses labelled with the International Classification of Diseases (ICD) codes, as well as visiting clinic and hospital. The NPR applies different ICD code versions in different time periods: ICD-7 during 1964 to 1968, ICD-8 during 1969-1986, ICD-9 during 1987-1996 and ICD-10 from 1997 to present (318). In Sweden, patients with IIM are treated and followed by hospital-based specialists at specific inpatient or (mostly) outpatient clinics. Hence, the NPR has a good coverage of patients with IIM.

3.2.1.2 The Swedish Cancer Register

All histologically or cytologically confirmed malignant tumours should be reported by clinicians and pathologists to the Swedish Cancer Register (SCR) since 1958. The SCR has an overall coverage of 96% and a validity of nearly 100%. However, there is a large variation in underreporting among site-specific cancers. An investigation of completeness of the SCR against the inpatient data of the NPR found underreporting rate varied from 1% (skin) to 20% (other sites) across cancer sites and further increased for age at diagnosis over 70 years. Lack of histology or cytology verification is the major reason of underreporting. In the SCR, each

primary cancer of an individual is registered separately. Annual examination is performed by the six regional cancer registers to prevent multiple registrations of a primary cancer. Unfortunately, cancer relapse is not captured by the SCR (319).

The SCR holds a wide range of medical data including date of diagnosis and diagnostic codes for cancer site, histological type and cancer stage. Each cancer is coded by ICD code since 1958 and Systematized Nomenclature of Medicine (SNOMED) code since 1993, with different versions in different time periods: ICD-7 (since 1958), ICD-9 (since 1987), ICD for Oncology second edition (ICD-O/2) (since 1993), SNOMEDO10 (since 1993), SNOMED3 (since 2005), ICD-O/3 (since 2005). In 2004, the tumour, node and metastasis (TNM) staging system and the International Federation of Gynecology and Obstetrics (FIGO) staging system were introduced for cancer staging (319).

3.2.1.3 The Cause of Death Register

The Cause of Death Register (CDR) records all deaths since 1952 and holds data on death date, underlying and contributing causes of death. Underlying and contributing causes of death are specified based on the medical death certificate issued by the patients' usual physician or the last-contacted physician. Different ICD code versions are used to define cause of death: ICD-7 (1958-1968), ICD-8 (1969-1986), ICD-9 (1987-1996) and ICD-10 (1997 and onwards). The completeness of CDR is virtually 100% and 96% of individuals in the CDR have information on underlying cause of death (320). In 2015, the missingness of death is 0.9%. The quality of the CDR depends on the validity of death certificate and is overall high. For deaths with malignant neoplasm as the cause of death in the CDR, 90% of them agree with the medical record (321). Therefore, when studying prevalent cancer, researchers may use data from the CDR to capture additional cases that were not registered in the SCR.

3.2.1.4 The Total Population Register

The TPR founded in 1968 holds birth and death records and other basic demographic data of nearly the entire Swedish population, as well as data on immigration and emigration, with a coverage of 95% and 91%, respectively. The TPR is the source of comparators in many epidemiological studies (322).

In the TPR, emigration is defined as individuals staying abroad for a year or more except those who work for the Swedish state or who are in prison or who died suddenly. Moreover, individuals moving to other Nordic countries will remain in the Swedish TPR till they are registered in the population register of their new country of residence (322).

3.2.1.5 The Multi-Generation Register

The Multi-Generation Register (MGR) was established using data derived from census data and further developed into the present state after several reformations done between 1991 and 2013 and became part of the TPR. The MGR holds parental information of individuals who were born in 1932 or later and registered in Sweden since January 1 1961. Information om biological and adoptive parents is available. This thesis considered only biological parents. The coverage of the MGR varies with country of birth. For individuals born in

Sweden in 1947, 97% and 95% of them have information on mother and father, respectively, and the coverage of both parents is complete for those born in 1972 and onwards. The coverage of parental information was much lower in those born outside Sweden in 1947 (27% for mother and 22% for father) and it remained below 85% for those born in 2007 (323). The quality of parental information is high in the MGR given that the frequency of misattributed paternity is 1% in Sweden (324).

3.2.1.6 The Prescribed Drug Register

The Prescribed Drug Register (PDR) funded in mid-2005 holds data on all dispensations of prescribed drugs including drug name, prescription and dispensation dates and dosage. The Anatomical Therapeutic Chemical code is used to label and classify the prescribed drug. Drugs that purchase over the counter or administrated at hospital are not included in the PDR (325).

3.2.2 The Myositis Genetics Scientific Interest Group of the International Myositis Assessment & Clinical Studies Group

The International Myositis Assessment & Clinical Studies Group (IMACS) is an international consortium of clinicians, researchers, patients and pharma representatives with interest in IIM research. The IMACS has a vision to improve quality of life of patients affected by IIM through conducting collaborative research in various topics including aetiological research, clinical trials and developing new diagnostic criteria and disease management guidelines.

The Myositis Genetics Scientific Interest Group (MYOGEN) is one of the focus groups under the IMACS and aims to investigate genetic profile of IIM as whole and its major subtypes via diverse approaches such as exploratory GWAS, and subsequent fine mapping and functional studies.

The MYOGEN consortium holds genetic data of up to 2,954 patients with IIM and 15,651 geographically matched controls from the US and 13 European countries. The Bohan and Peter criteria was used to define DM, JDM and PM (48, 49). Patients with IBM were diagnosed based on the Griggs or the European Neuromuscular Centre criteria (51). The age cut-off for JIIM was < 16 years old for all cases except those cases from the USA where the age cut-off was < 18 years old. Patients were also screened for IIM mimics and those with indication of IIM minics were excluded. Geographically matched controls are selected from different consortia (3-6).

The MYOGEN consortium contains several sets of data derived from the initial GWAS and the ImmunoChip study. The original GWAS data contains genotyping data for the whole genome of 1,710 patients with IIM and 4,724 controls. Three years after the initial GWAS of IIM published in 2013, samples were collected from 2,954 patients with IIM and 15,651 controls and genotyped using the Immunochip (covering 186 established loci related to 12 autoimmune diseases) (3-6). Recently, a new imputation using the Trans-Omics for Precision Medicine Imputation Server against the Trans-Omics for Precision Medicine reference panel version R2 on human genome assembly GRCh38 was performed with these two data sets (**Table 3**) (326). All included patients were European descents. Before imputation, SNPs

were filtered on a call rate < 95% or a p-value < 10^{-5} for controls and a p-value < 10^{-9} for cases in the Hardy-Weinberg equilibrium test. Furthermore, individuals with a genotyping missing rate > 5% or with pi-hat > 0.2 determined by the identity-by-descent coefficient were excluded. In this quality control step, duplicates among these two data sets were identified (pi-hat=1) and only the individuals in the ImmunoChip data set were kept, making the two data sets independent to each other. Imputation was performed separately for IIM subgroups including DM, PM, JDM and anti-Jo1 autoantibodies in each data set. After imputation, SNPs with minor allele frequency < 0.01 or low imputation quality (R^2 < 0.6) were removed. Principal component (PC) analysis with PLINK for the GWAS data and Ancestry Inference using principal component analysis and spatial analysis for the ImmunoChip data were performed to correct for population stratification. SNP associations were tested using logistic regression where the first five PCs were adjusted in the models.

Table 3. The number of patients with IIM, controls and SNPs, and the used Illumina arrays in the GWAS and the ImmunoChip data

	No. Case	No. Control	No. typed SNP	No. SNP after imputation	Illumina chip
GWAS					
IIM	936	4,336	223,661	8,668,073	Human1M-Duo v3 Human660W-Quad v1 HumanHap550 HumanCNV370-Duo v1 Human610-Quad v1
DM	402				
PM	255				
JDM	260				
Anti-Jo1	97				
ImmunoChip					
IIM	2,592	7,486	142,220	1,201,876	Immunoarray
DM	870				
PM	923				
JDM	476				
Anti-Jo1	336				

IIM: Idiopathic inflammatory myopathy; SNP: Single nucleotide polymorphism; GWAS: Genome-wide association study; DM: Dermatomyositis; PM: Polymyositis; JDM: Juvenile dermatomyositis; Jo1: Histidyl tRNA synthetase.

3.2.3 The International Lymphoma Epidemiology Consortium

The International Lymphoma Epidemiology Consortium (InterLymph), formed in 2001, is a collaborative group of investigators conducting epidemiological research of NHL. The InterLymph holds genetic data of patients with pathologically confirmed B-cell lymphomas (subtypes including DLBCL, FL, CLL and MZL) and sex- and age matched controls from 22 studies: nine cohort studies, eight population-based case-control studies and five clinic or hospital-based case-control studies (86, 90, 91, 93). B-cell lymphomas were classified based on the World Health Organisation classification (2008). All included individuals were European descents.

The genetic data of all B-cell lymphoma subtypes except MZL were from multiple study groups. **Table 4** summaries the number of patients and controls, the number of typed SNPs after quality control and the used Illumina chips in each study group, as well as the number of quality-controlled SNPs after imputation for each B-cell lymphoma subtype. Information on quality control criteria prior to imputation is presented in previous studies (86, 90, 91, 93, 327). Imputation was performed using IMPUTE2 with the 1000 Genomes Project version 3 as the reference panel. Only common SNPs with minor allele frequency > 0.01 and with imputation INFO score > 0.3 were included. The HLA region was imputed using SNP2HLA

based on the reference panel from the Type 1 Diabetes Genetics Consortium. Logistic regression models with SNPTEST version 2 were used to test SNP associations, adjusting for age, sex and significant principal components. Meta-analyses were done with the imputed SNPs using fixed effects inverse variance models. SNPs were aligned to human genome assembly GRCh37 (86, 90, 91, 93, 327).

Table 4. The number of patients, controls and SNPs, and the used Illumina arrays in the each study group for each B-cell lymphoma subtype

	No. Case	No. Control	No. typed and quality controlled SNP	No. SNP after imputation	Illumina chip
			DLBCL		
NCI GWAS	2,661	6,221	611,844	9,116,853	Illumina OmniExpress, Illumina Omni2.5
MAYO GWAS	393	172	523,949		Illumina HumanHap 660W
UCSF2 GWAS	254	748	290,523		Illumina HumanCNV370-Duo
GELA/EPIC GWAS	549	525	513,264		Illumina HumanHap 610K for cases and Illumina HumanHap 660W or 610K for controls
Total	3587	7666			
			FL		
NCI GWAS	2,142	6,221	611,844	9,078,855	Illumina OmniExpress
SCALE GWAS	376	791	298,045		Illumina HumanHap 317K
UCSF2 GWAS	210	746	290,523		Illumina HumanCNV370-Duo
UCSF1 GWAS	119	349	614,320		Illumina OmniExpress
Total	2847	8107			
			CLL		
NCI GWAS	2,179	6,221	608,811	9,098,434	Illumina OmniExpress, Illumina Omni2.5
UTAH GWAS	321	405	512,171		Illumina HumanHap 610K
GEC GWAS	387	294	687,578		Affymetrix 6.0
UCSF2 GWAS	213	747	290,523		Illumina HumanCNV370-Duo
Total	3100	7667			
			MZL		
NCI GWAS	825	6,221	611,856	8,478,065	Illumina OmniExpress, Illumina Omni2.5

SNP: Single nucleotide polymorphism; DLBCL: Diffuse large B-cell lymphoma; FL: Follicular lymphoma; CLL: Chronic lymphocytic leukaemia; MZL: Marginal zone lymphoma; GWAS: Genome-wide association study; NCI: National Cancer Institute; MAYO: Mayo Clinic study; UCSF2: San Francisco bay area non-Hodgkin lymphoma 1 case-control study; GELA: Groupe d'Etude des Lymphomes de l'Adulte; EPIC: European prospective investigation into cancer, chronic diseases, nutrition and lifestyles; SCALE: Scandinavian lymphoma etiology FL case-control study; UCSF1: San Francisco bay area non-Hodgkin lymphoma 2 case-control study; UTAH: Utah CLL study; GEC: Genetic epidemiology of CLL consortium.

3.3 Study populations

3.3.1 Patients with IIM and matched comparators without IIM identified in the NPR

Prevalent (**Study I-III**) and incident (**Study V**) cases of IIM were identified using modified register-based algorithms, formulated based on clinical experience in diagnostic workup for IIM and verified to be robust (34). In the study of Svensson *et al.,* the original base case definition used to define incident IIM required a first ever inpatient or outpatient visit indicating IIM, followed by at least one subsequent visit occurring within one to 12 months after the first IIM visit in the NPR or the Swedish Rheumatology Qaulity Register between 2007 and 2011. Similarly, the original base case definition for prevalent IIM in 2012 required ≥ 2 inpatient or outpatient visits indicating IIM in the NPR or a Swedish Rheumatology

Qaulity Register registration. In these definitions, ICD-10 codes M33.0, M33.1, M33.2, G72.4, M60.8 and M60.9 from internal medicine, rheumatology, dermatology, neurology and paediatrics clinics were used to define IIM cases. The authors found this algorithm to be robust against liberal (required ≥ 1 visit) and strict (additionally required an indication of IIM treatment within 12 months after the first visit) case definitions. Moreover, the ICD-10 codes relevant to IIM were validated by using subtype data from the Swedish Rheumatology Qaulity Register, where diagnosis is entered by clinicians, as the gold standard. The ICD-10 code M33.0, M33.1 and G72.4 has a PPV ≥ 90% to define JDM, DM and IBM, respectively. The PPV is 83% for M33.2 and 57% for M33.9, indicating a risk of subtype misclassification using these codes.

To align with the available data sources and research aims of this thesis, a couple of modifications were made to the abovementioned definitions: 1) using NPR data only due to unavailability of Swedish Rheumatology Qaulity Register data; 2) extending time window to cover more IIM cases; 3) dropping ICD-10 codes M60.8 and M60.9 to increase specificity since these codes are usually used for spurious cases not definite case of IIM; 4) including additional specialist clinics in **Study III** and **V** to increase coverage since patients were treated in clinic other than the usual ones at some hospitals; 5) grouping cases with M33.2, M33.9 or G72.4 as other IIM as these codes are not subtype specific. The register-based algorithm used to define IIM in each of the **Study I-III** and **V** is presented in **Table 5**.

In **Study I-III**, each patient was matched to up to five individuals without IIM randomly selected from the TPR on sex, birth year and residential area at the index date. In **Study I** and **II**, individuals without IIM had no first-degree relatives alive in 1987 or born before 1932 or outside Sweden were excluded. In **Study III**, individuals without IIM had no first-degree relatives alive in 1958 or born before 1932 or outside Sweden were excluded. In these studies, patients with IIM and the matched individuals without IIM were collectively referred as index individuals.

Table 5. The register-based algorithms used to define and classify IIM in Study I-III and V

	Study I and II	Study III	Study V
Type of algorithm	Prevalence	Prevalence	Incidence
Inpatient visit only	1997-2000: ≥ 1 inpatient visit with IIM as main diagnosis	1997-2000: ≥ 1 inpatient visit with IIM as main diagnosis	1998-2001: a first ever inpatient visit with IIM as main diagnosis
Inpatients and outpatient visits	2001-2016: ≥ 2 inpatient or outpatient visits with IIM as main or contributory diagnosis	2001-2016: ≥ 2 inpatient or outpatient visits with IIM as main or contributory diagnosis	2002-2020: a first ever inpatient or outpatient visit indicating IIM and ≥ 1 subsequent visit happening within one to 12 months after the first visit of IIM
ICD 10 codes	M33 and G72.4	M33 and G72.4	M33 and G72.4 *ICD codes of the previous versions to exclude prevalent cases: ICD-9: 710D and 710E ICD-8: 716,00 and 716,10 ICD-7: 710,00, 710,01 and 726,30
Specific clinics	Internal medicine, rheumatology, dermatology, neurology and paediatrics	All Swedish hospitals except the twos specified below: internal medicine, rheumatology, dermatology, neurology and paediatrics Östersunds sjukhus: rehabilitation Länssjukhuset Ryhov: orthopaedic clinic	All Swedish hospitals except the twos specified below: internal medicine, rheumatology, dermatology, neurology and paediatrics Östersunds sjukhus: rehabilitation Länssjukhuset Ryhov: orthopaedic clinic
Exclusion criteria	With contributory diagnosis indicating IIM only or born before 1932 or born outside Sweden or without first-degree relatives who were alive in 1987	With contributory diagnosis indicating IIM only or born before 1932 or born outside Sweden or without first-degree relatives who were alive in 1958	With contributory diagnosis indicating IIM only or age at diagnosis < 18 years
Index date	1997-2000: Discharge date 2001-2016: Discharge or outpatient visit date of the second visit	1997-2000: Discharge date 2001-2016: Discharge or outpatient visit date of the second visit	1998-2001: Discharge date 2002-2020: Discharge or outpatient visit date of the earliest eligible subsequent visit
IIM subtype classification	DM (M33.0 and M33.1 and age at diagnosis > 18 years) Other IIM (M33.2, M33.9 and G72.4 and age at diagnosis > 18 years) JIIM (age at diagnosis ≤ 18 years)	DM (M33.0 and M33.1 and age at diagnosis > 18 years) Other IIM (M33.2, M33.9 and G72.4 and age at diagnosis > 18 years) JIIM (age at diagnosis ≤ 18 years)	DM (M33.0 and M33.1 and age at diagnosis ≥ 18 years) Other IIM (M33.2, M33.9 and G72.4 and age at diagnosis ≥ 18 years)

IIM: Idiopathic inflammatory myopathy; DM: Dermatomyositis; JIIM: Juvenile IIM; ICD: International Classification of Diseases.

3.3.2 First-degree relatives

Given that members within a family, to various extent, share genetics and are exposed to similar living environment, they are a good source of data to explore genetic and environmental effects on disease development. In this thesis, first-degree relatives including parents, full-siblings and offspring who share 50% of genetic similarity were used. An assumption was made concerning the degree of shared environmental factor among different types of first-degree relatives: full-siblings were presumed to have a higher degree of shared environmental factors than parent-offspring relationship. In **Study I-III**, first-degree relatives of the corresponding index individuals were identified from the MGR.

3.3.3 Defining IIM and autoimmune diseases in first-degree relatives

The NPR data overall has a good validity to define autoimmune diseases. The PPV of using ≥ 1 visit in the NPR for disease ascertainment was above 80% for RA (90% for inpatient visit and 85% for outpatient visit), SLE (97%), MS (93%), IBD (88%), T1DM (95%), CeD (86%) and MG (83%) (302, 328-333). A lower PPV was observed for IIM (60%), SSc (68%) and SS (77%) (12, 334, 335). To increase number of cases, in the main analysis in **Study I** and **II**, an autoimmune disease was defined by ≥ 1 visit with main diagnosis indicating the studied autoimmune diseases between 1987 and 2017 in the NPR. Furthermore, strict definitions requiring ≥ 2 visits indicating the studied autoimmune disease were used in the sensitivity analyses to test the robustness of the main findings. **Table 6** shows the main and strict definitions used to define IIM and other autoimmune diseases including RA, other RIDs (SLE, SSc, SS and other CTD), MS, IBD, T1DM, AITD, CeD and MG among first-degree relatives in **Study I** and **II**.

Table 6. The definitions used to define idiopathic inflammatory myopathies and other autoimmune diseases in first-degree relatives

Autoimmune diseases	Time period	ICD codes	Types of diagnosis	Main definition	Strict definition
Idiopathic inflammatory myopathies	1987-2017	710D, 710E, M33.0, M33.1, M33.2, M33.9, G72.4	Main/contributory diagnosis	≥ 1 main diagnosis in the NPR	≥ 1 main diagnosis in inpatient register between 1987 and 2000 or ≥ 2 diagnoses (≥ 1 main diagnosis) in the NPR between 2001 and 2017 in internal medicine, rheumatology, neurology, dermatology or paediatric clinic
Rheumatoid arthritis	1987-2017	714A-C, 714W, M05, M06.0, M06.2, M06.3, M06.8, M06.9	Main/contributory diagnosis	≥ 1 main diagnosis in the NPR	≥ 1 main diagnosis in inpatient register between 1987 and 2000 or ≥ 2 diagnoses (≥ 1 main diagnosis) in the NPR between 2001 and 2017 in internal medicine, rheumatology or paediatric clinic
Other rheumatic inflammatory diseases	1987-2017	136B, 710A, 710B, 710C, 710W, 710X, 725, M32, M34.0, M34.1, M34.8, M34.9, M35	Main diagnosis	≥ 1 main diagnosis in the NPR	-
Multiple sclerosis	1987-2017	340, G35	Main/contributory diagnosis	≥ 1 main diagnosis in the NPR	≥ 1 main diagnosis in inpatient register between 1987 and 2000 or ≥ 2 diagnoses (≥ 1 main diagnosis) in the NPR between 2001 and 2017 in internal medicine or neurology or paediatrics clinic
Inflammatory bowel diseases	1987-2017	555, 556, K50-51	Main/contributory diagnosis	≥ 1 main diagnosis in the NPR	≥ 1 main diagnosis in inpatient register between 1987 and 2000 or ≥ 2 diagnoses (≥ 1 main diagnosis) in the NPR between 2001 and 2017 in internal medicine, gastroenterology, rheumatology, surgical care, gastrointestinal care or paediatrics clinic

Continued on next page

Table 6. The definitions used to define idiopathic inflammatory myopathies and other autoimmune diseases in first-degree relatives (Continued)

Autoimmune diseases	Time period	ICD codes	Types of diagnosis	Main definition	Strict definition
Type 1 diabetes mellitus	1987-2017	250, E10	Main/contributory diagnosis	≥ 1 main diagnosis in the NPR. The diagnosis had to be made ≤ 30 years of age as the ICD-9 code 250 cannot distinguish Type 1 diabetes mellitus from Type 2 diabetes mellitus and the ICD-10 code E10 might be used for Type 2 diabetes mellitus as it can developed into insulin dependency.	≥ 1 main diagnosis in inpatient register between 1987 and 2000 or ≥ 2 diagnoses (≥ 1 main diagnosis) in the NPR between 2001 and 2017 in endocrinology, internal medicine or paediatrics clinic. The diagnosis had to be made ≤ 30 years of age as the ICD-9 code 250 cannot distinguish Type 1 diabetes mellitus from Type 2 diabetes mellitus and the ICD-10 code E10 might be used for Type 2 diabetes mellitus as it can developed into insulin dependency.
Autoimmune thyroid diseases	1987-2017	242, 242A, 242B, 242D, 242E, 244X, 245C, E038, E039, E063, E050, E051, E052, E053, E058, E059, O905	Main diagnosis	≥ 1 main diagnosis in the NPR or ≥ 1 filling of a prescription of thyroid hormone substitution therapy (H03AA01 and H03AA02) between 2005-2017, with no history of thyroid cancer or a prescription of iodine-containing drugs (C01BD01, N05AN01, L03AB01/L03AB04/L03AB05)	-
Celiac disease	1987-2017	579A, K90.0	Main/contributory diagnosis	≥ 1 main diagnosis in the NPR	≥ 1 main diagnosis in inpatient register between 1987 and 2000 or ≥ 2 diagnoses (≥ 1 main diagnosis) in the NPR between 2001 and 2017 in internal medicine, gastroenterology, gastrointestinal care, paediatrics clinic
Myasthenia gravis	1987-2017	358A, G70.0	Main/contributory diagnosis	≥ 1 main diagnosis in the NPR	≥ 1 main diagnosis in inpatient register between 1987 and 2000 or ≥ 2 diagnoses (≥ 1 main diagnosis) in the NPR between 2001 and 2017 in internal medicine, neurology or paediatrics clinic

NPR: National Patient Register.

3.3.4 Population-based case-control family design, accompanied with Directed Acyclic Graph

A population-based case-control family study can be conceived as a family study using case-control sampling at population level. This design has advantages of increasing study efficiency (case-control sampling from population register enables cost-effective inclusion of a desired number of patients with a rare disease, along with controls and their relatives),

minimizing systematic bias (use of prospectively collected data from register to define exposure can largely reduce recall bias compared to data collection via interview or questionnaires) and having better generalizability (population-level sampling can reduce risk of biased selection). Given the strengths of this study design, it was used to estimate the familial aggregation of IIM (**Study I**), and the familial co-aggregation of IIM and other autoimmune diseases (**Study II**) and cancer (**Study III**).

In traditional case-control family study investigating familial co-aggregation of two diseases, estimate is often interpreted as familial risk or association. Causal relationships between one disease in relatives and another disease in index individuals, as well as other involved factors are rarely specified, making causal interpretation difficult. Hudson *et al.* has proposed a structural approach to tackle this limitation by incorporating Directed Acyclic Graph (DAG) to specify causal relationships between variables within a relative pair under different causal scenarios (336, 337). DAG, a causal diagram, uses directed edges to present causal relationships between variables and does not allow cyclic causal relationship (a variable causes itself directly or indirectly through other variables). In a DAG, an open path (absence of common effect (collider), that is, a variable has more than one cause) represents a statistical association between two variables while a closed path (presence of common effect) indicates that such statistical association should be absent. A closed path can be opened after adjusting for a collider (338).

In this structural approach, causal relationships between variables within a relative pair are illustrated in **Figure 4**. Shared familial factors (C) are a latent (unmeasured) variable that has a central role in these DAGs. Other variables presented in the DAGs include the four disease variables indicating IIM status and a secondary disease status (an autoimmune disease or cancer) in index individuals (IIM_1 and SD_1) and in any type of relatives (IIM_2 and SD_2). First-degree relative is used in the subsequent description to maintain consistency with the thesis. There are four edges pointing to the four diseases variables from C, representing potential causal relationships between shared familial factors and diseases clustering within relatives. Additional four latent variables represent unique individual factors (i.e., sex and age) that may lead to IIM and a secondary disease in index individuals (U_1) and first-degree relatives (U_2) and disease-specific factors that are not shared within families (U_{IIM} and U_{SD}). It is assumed that the latent variables have additive effects and should be independent to each other when the four disease variables are not adjusted (336).

When no causal relationship between IIM and a secondary disease is assumed, an observed statistical association can be directly interpreted as being due to shared familial factors (**Figure 4a**). It is because only the paths $IIM_{1/2} \leftarrow C \rightarrow SD_{2/1}$ are open. The following introduction focuses on the paths linking IIM_1 and SD_2 since this thesis studied the familial association between IIM in index individuals and an autoimmune disease or cancer in first-degree relatives.

Interpretation of findings from this structural approach can become complicated when there are direct effects between IIM and a secondary disease as additional paths without C are open, as illustrated in **Figure 4b** and **4c**. In **Figure 4b**, assuming IIM can be caused by a secondary disease, two additional paths ($IIM_1 \leftarrow SD_1 \leftarrow C \rightarrow SD_2$ and $IIM_1 \leftarrow SD_1 \leftarrow U_{SD} \rightarrow SD_2$) are open, and the latter does not include C, meaning that there can be a statistical association even in

the absence of shared familial factors. The latter path can be blocked by adjusting for the secondary disease in index individuals, but this adjustment can open other new paths ($IIM_1 \leftarrow U_1 \rightarrow SD_1 \leftarrow U_{SD} \rightarrow SD_2$, $IIM_1 \leftarrow U_1 \rightarrow SD_1 \leftarrow C \rightarrow SD_2$ and $IIM_1 \leftarrow C \rightarrow SD_1 \leftarrow U_{SD} \rightarrow SD_2$) since SD_1 is a collider. Adjustment of a collider often leads to downward bias by introducing negative association between variables, here IIM_1 and SD_2 (339). Similar consequence can be seen in the scenario assuming a direct effect from IIM to a secondary disease (**Figure 4c**). Therefore, the authors of this approach suggest that researchers could interpret the adjusted estimate as a lower limit of familial association when assuming a causal relationship between two diseases. Moreover, if statistical significance remains after adjustment, it is a strong evidence suggesting the presence of familial factors shared between IIM and a secondary disease (336).

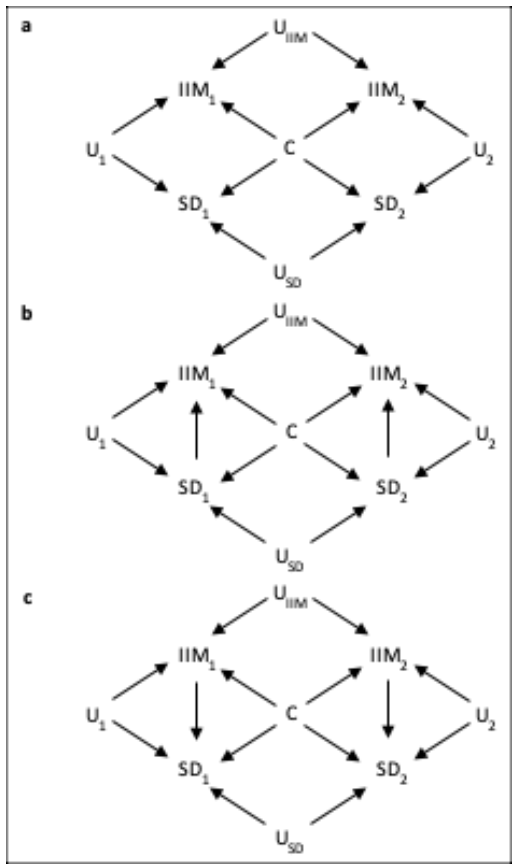


Figure 4. The proposed underlying mechanisms between IIM and a secondary disease (SD, an autoimmune disease or cancer in this thesis) illustrated by Directed Acyclic Graphs. *IIMj and SDj represent IIM and a SD, respectively, for individual j in a given pair; j = 1, 2. U_{IIM} represents common causes for IIM_1 and IIM_2, and U_{SD} represents common causes for SD_1 and SD_2. Uj represents unique individual common causes for IIMj and SDj that may vary within the pair, and C represents shared family factors for IIMj and SDj that are constant within the pair. a. No causal relationship; b. Causal relationship from a SD to IIM; c. Causal relationship from IIM to a SD. IIM: Idiopathic inflammatory myopathy.*

Key features in terms of study design and statistical analyses for each study are presented in the following sections. For details, please refer to the corresponding publication/manuscript.

3.3.5 Study I: Familial aggregation and heritability of IIM

Study I was a population-based case-control family study. Patients with prevalent IIM, matched comparators without IIM and their first-degree relatives were identified as described in section 3.3. We defined IIM in first-degree relatives based on the main and strict definitions presented in **Table 6**.

Given the matching design, logistic regression conditioning on matching stratum was used to compare the odds of having ≥ one first-degree relative affected by IIM (exposure) in patients with IIM (outcome) with that in matched comparators without IIM. To increase statistical efficiency, each first-degree relative was paired with the corresponding index individual and each relative pair was treated as an independent unit. Same regression model was used but standard error was adjusted by robust sandwich variance estimator to account for data dependence in this data format. Furthermore, in the conditional logistic regression model, sex and birth year of first-degree relatives were adjusted in addition to the matching factors. Parents, full-siblings and offspring share varying degrees of similarity in environmental factors. We performed a subgroup analysis by types of kinship to explore difference in familial aggregation of IIM between these relationships.

In quantitative genetics, heritability is defined as a proportion of the phenotypic variance that is explained by genetic variance. Heritability can be defined as broad- or narrow-sense heritability depending on the types of genetic effect included in the estimation. Broad-sense heritability accounts for all genetic effects including additive, dominant and epistatic effects while only additive genetic effect is included in narrow-sense heritability. Among these genetic effects, only additive genetic effect is inheritable from parents to children. Together with ample amount of evidence suggesting a major role of additive genetic effect in complex diseases, we assumed that genetic variance of IIM was solely due to additive genetic effect and thus we estimated the narrow-sense heritability of IIM. We also assumed that there was no assertive mating and familial resemblance was due to genetics only. We used tetrachoric correlation to estimate the heritability of IIM. This analysis involved using a 2X2 contingency table, presenting the concordant and discordant IIM status within relative pairs of patients with IIM and matched comparators without IIM, and it was used to estimate the correlation coefficient in disease liability of IIM (tendency of having IIM) between these relative pairs. The counts of exposed and unexposed relative pairs of matched comparators without IIM were corrected by the prevalence of IIM in Sweden (0.014%) and the estimated OR of familial association of IIM to account for the matched case-control sampling used in the study (340). The heritability of IIM was obtained by dividing the correlation coefficient by the degree of relatedness, 0.5 for first-degree relatives. To test if the heritability estimate was robust to the change of the prevalence of IIM, we repeated the analyses by correcting the count numbers with a range of prevalence values between 0.004% and 0.024%.

To evaluate the potential impact of variation in IIM ascertainment, we did all analyses anew. This involved defining IIM in first-degree relative using the strict definition, and including only index individuals with first-degree relatives alive in 2001, the year when both inpatient and outpatient data became nationwide available.

3.3.6 Study II: Familial autoimmunity of IIM

Study II had the same study design and study population as **Study I**. Briefly, in the main analyses, we used the main definitions presented in **Table 6** to define IIM and a wide range of autoimmune diseases in first-degree relatives of index individuals. Conditional logistic regression was used to estimate the familial association between IIM (outcome) and each autoimmune disease by the number of first-degree relatives affected by the autoimmune disease (exposure, ≥ 1 and ≥ 2 affected first-degree relatives) and by treating each first-degree relative pair as an independent unit. In the latter analysis, conditional logistic regression models were accompanied by causal DAGs with different assumptions on direct effects between IIM and an autoimmune disease, as illustrated in **Figure 4**. Subgroup analyses by types of kinship were performed. In the main analyses, we assumed no direct effects. This is a reasonable assumption given the common genetic architecture among many autoimmune diseases. However, to test the robustness of our main findings, we conducted sensitivity analyses by additionally adjusting for firstly autoimmune disease in index individuals and secondly IIM in first-degree relatives. We performed two additional sensitivity analyses where strict definitions were used to define IIM and other autoimmune diseases in first-degree relatives and only first-degree relatives alive in 2001 were included. In the models for relative pairs, robust sandwich variance estimator was used to control standard errors, and sex and birth year of first-degree relatives were adjusted.

3.3.7 Study III: Familial co-aggregation of IIM and cancer

Study III was a population-based case-control family study with a study population including patients with IIM, matched comparators without IIM and their first-degree relatives identified according to the definitions presented in section 3.3.

We defined cancer in index individuals and their first-degree relatives based on a diagnosis in the SCR or a death record with cancer as the underlying cause of death in the CDR between 1958 and 2016. Cancers were classified into different subtypes that are associated with IIM or common in the Swedish population based on the ICD and SNOMED codes (8, 9, 168, 189, 196, 197). In order to have a more accurate and precise classification, we prioritised the latest version of ICD and SNOMED codes. For individual with more than one primary cancer diagnosis in the SCR, only the first diagnosis was classified.

We defined IIM in first-degree relatives using the same case definition but extending the time frame from 1987 to 2016. ICD-9 codes 710D and 710E were used to define IIM diagnosed between 1987 and 1996.

We formatted the family data into index individual-relative pairs. Conditional logistic regression models based on the defined DAGs shown in **Figure 4** were used to estimate the familial co-aggregation of IIM in index individuals (outcome) and cancer overall in first-degree relatives (exposure). We stratified the analyses by types of kinship, IIM subtypes and sex-concordance of relative pairs to examine potential effect modifications. We corrected standard errors with robust sandwich variance estimator and additionally adjusted for sex and birth year of first-degree relatives in the models. However, sex of first-degree relatives was not adjusted in the subgroup analyses by sex-concordance of relative pairs.

There were two exploratory analyses investigating if the familial association varied with age at cancer diagnosis (< 50 years) and cancer types. Familial associations between IIM and specific cancer types were analysed overall, by types of kinship and IIM subtypes and by sex concordance of relative pairs. In the analyses of sex-specific cancer types, only first-degree relative pairs with the corresponding sex were included and sex was not adjusted. No exploratory analysis by cancer types was performed for JIIM due to small number of cancer cases.

Multiple testing was an issue in the main and the exploratory analyses by specific cancer types. We used Benjamini and Hochberg (BH) procedure to control the false discovery rate (FDR) at 0.05 in both analyses (341).

Another concern in **Study III** was that assuming no direct effects between IIM and cancer in the main analyses was unlikely true as evidence in the literature suggests a bidirectional relationship between IIM and cancer. To investigate if direct effects between IIM and cancer had an impact on the estimation, we performed sensitivity analyses by adjusting for cancer in index individuals (when assuming a causal effect from cancer to IIM) and IIM in first-degree relatives (when assuming a causal effect from IIM to cancer).

3.3.8 Study IV: Genetic overlap between IIM and common B-cell lymphoma subtypes

Study IV was a cross-trait genetic study where summary statistics from independent GWASs of IIM, DLBCL, FL, CLL and MZL were used to explore the genetic correlations between these diseases. To enhance comprehension, the major concepts central to this study, along with the study design and statistical methods, are presented in separate sections.

3.3.8.1 The value of GWAS summary statistics

GWAS, with a capability to test associations of hundreds of thousands to millions of genetic variants (usually common SNPs with minor allele frequency > 5%) across the genome with a phenotype, has a pivotal role in understanding the genetics of complex diseases. Since the first GWAS published in 2005, more than 50,000 disease-associated genetic variants have been reported (342). Furthermore, summary statistics from GWAS, defined as aggregated level of genetic variant association data such as p value and effect estimates, have diverse applications including but not limited to estimations of single-trait heritability and cross-trait genetic correlation (343-346).

One major concern about the use of GWAS summary statistics for cross-trait genetic correlation analysis is the diversity of data formats between studies. The common format and quality problems include inconsistent human genome assembly, presence of ambiguous SNPs, SNPs with null effect size or low imputation quality, unaligned reference allele and unaccommodated effect estimate to the used software (343). Although several movements have been advocated to store GWAS summary statistics in standard format (i.e., variant call format) that is ready for analysis, many summary statistics, particularly those published in the past, are still subject to problem of heterogeneity in data formats (347). Keeping this potential problem in mind, our GWAS summary statistics from the disease consortia underwent multiple steps of quality control and standardisation before statitsitcal analyses (348-351) (**Table 7**). The quality controlled and harmonised data were presented in human

genome assembly GRCh37. We paired each B-cell lymphoma subtype to the GWAS DM, the GWAS PM, the ImmunoChip DM and the ImmunoChip PM data and kept only the overlapping SNPs in each pairing, resulting in a total of eight GWAS and eight ImmunoChip disease pairs.

Table 7. The number of cases, controls and SNPs included after quality control and data alignment for IIM and B-cell lymphoma subtypes

		GWAS DM	GWAS PM	Immuno Chip DM	Immuno Chip PM	DLBCL	FL	CLL	MZL
Cases		402	255	870	923	3,587	2,847	3,100	825
Controls		4,336	4,336	7,486	7,486	7,666	8,107	7,667	6,221
Stable biallelic SNPs	Imputed SNPs	8,668,073	8,668,073	1,201,875	1,201,875	9,116,853	9,078,855	9,098,434	8,478,065
	Single-study group	-	-	-	-	498,560	460,723	487,417	-
	SNP duplicates	-	-	-	-	2	8	3	-
	Missing odds ratio	2[a]	79[a]	-	-	-	-	-	-
	Multiallelic	573,612	573,686	79,739	79,739	-	-	-	-
	Conversion unstable SNPs	1,878	1,878	48	48	-	-	-	-
LiftOver to GRCh37	Failed to be converted	14,036	14,036	1,115	1,115	-	-	-	-
QC and Alignment	Missing on reference genome	50,963	50,945	38,455	38,455	49,253	49,005	48,787	48,121
	Ambiguous SNPs	1,241,828	1,241,813	164,677	164,677	1,324,362	1,324,362	1,323,363	1,302,798
	Zero effect value	-	-	-	-	-	-	-	3,764
	Not matched to reference genome	3,015	3,015	49	49	-	-	-	-
Qced SNPs	After QC and alignment	**6,782,631**	**6,782,590**	**917,536**	**917,536**	**7,244,776**	**7,244,757**	**7,238,864**	**7,123,382**

SNP: Single nucleotide polymorphism; GWAS: Genome-wide association study; DM: Dermatomyositis; PM: Polymyositis; DLBCL: Diffuse large B-cell lymphoma; FL: Follicular lymphoma; CLL: Chronic lymphocytic leukaemia; MZL: Marginal zone lymphoma; QC: Quality control.
a SNPs with missing OR were multiallelic.

3.3.8.2 The role of linkage disequilibrium in genetic correlation analysis

Nearly all methods developed to estimate genetic correlation using GWAS summary statistics reply on LD, a common concept in population genetics describing the nonrandom association of alleles at two or more loci due to limited recombination (342). This nonrandom association can be quantified as a measure called LD score. In GWAS, a significant SNP association does not necessarily have causal meaning. Instead, it represents an average effect of all SNPs in LD with the tested SNP (352). SNPs in LD with a causal variant tend to have higher test statistics in GWAS and thus have higher likelihood of being detected. Importantly, those SNPs also tend to have a higher LD score. The relationship between LD score and GWAS test statistics is the key concept of the LD score regression, the first method used for

genetic correlation analysis using GWAS summary statistics (344, 345, 352). In single-trait LD score regression, one can estimate (narrow-sense) heritability by regressing test statistics of SNPs on the LD scores estimated with a matched population reference panel. In cross-trait LD score regression, genetic covariance between two diseases is estimated by regressing the product of z statistics pairs (a measure with implication of effect direction) of SNPs of the two diseases on the LD scores estimated with a matched population reference panel (344).

3.3.8.3 Local genetic correction and its estimation using LAVA

Genetic correlation can be measured at global and local levels. Global genetic correlation is an overall estimate of all SNPs across the genome. Heterogeneity in genetic correlations across the genome, usually the case in complex diseases, is not captured by global genetic correlation. Moreover, global genetic correlation may also overlook genetic correlations between two diseases if there is comparable number of correlations in opposite direction, which may be averaged out in an overall estimate. Another problem is that current methods such as LD score regression usually fails to compute global genetic correlations for diseases with low heritability including IIM, DLBCL, FL and MZL (316, 344). The primary challenge encountered in LD score regression for IIM lies in the exclusion of the HLA region, owing to its complex LD structure, which the method recommend for optimal analysis. Considering the central role of the HLA region in the genetic susceptibility of IIM, estimating genetic correlations with other diseases accounting for the influence of the HLA region into account offers more comprehensive understanding.

We fixed this problem by measuring local genetic correlation using local analysis of (co)variant association (LAVA) instead (346, 349, 353). Compared to global genetic correlation, local genetic correlation estimates genetic correlation between two diseases at each partitioned genomic region (loci). LAVA, like LD score regression, utilizes LD score and z statistics from GWAS to compute genetic correlations between multiple diseases. We partitioned the whole genome into 2,495 non-overlapping loci, of which 21 loci corresponded to the HLA region. In the LAVA analyses, summary statistics data of each disease pair were fitted in a multiple logistic regression model. Local heritability (at observed scale) was first computed for the 2,495 loci for both diseases in each disease pair (univariate analysis). Subsequently, only loci showing significant local heritability for both diseases in each disease pair were tested for local genetic correlation (bivariate analysis) (346). We adjusted p-values with Bonferroni correction to account for multiple testing.

3.3.8.4 Identification of jointly associated SNP using pleioFDR

To complement the LAVA analyses, we also applied pleiotropy-informed false discovery rate (pleioFDR) method to identify SNP jointly associated with both diseases in each disease pair (354). PleioFDR has two components: conditional FDR (CondFDR) and conjunctional FDR (conjFDR). The key idea of this method is that many common SNPs are associated with multiple diseases (pleiotropy) and for related diseases, SNPs associated with one disease have higher probability of being associated with the other disease than other SNPs with null associations. That is, pleioFDR is a method used to correct p-values of SNPs of a disease at a less stringent level than Bonferroni correction through leveraging pleiotropy with another

disease. Significant genomic enrichment observed in pleioFDR analysis suggests shared genetic susceptibility between two diseases.

We first examined the genomic enrichment in a fold-enrichment plot for each disease in a disease pair by strata of the $-\log_{10}$ of the nominal p-values of SNPs of an IIM/B-cell lymphoma subtype defined based on the p-values of the corresponding SNPs of a B-cell lymphoma/IIM subtype ($-\log_{10}(p) \geq 1$, $-\log_{10}(p) \geq 2$ and $-\log_{10}(p) \geq 3$). The fold enrichment was computed as the ratio between the cumulative distribution of SNPs in a given $-\log_{10}(p)$ stratum and the cumulative distribution of all SNPs (354). An observation with the strongest fold enrichment found in the stratum of the lowest p-values of SNPs and weaker enrichment in strata with higher p-values suggests shared genetic susceptibility between two diseases.

To identify SNPs jointly associated with both diseases which LAVA does not tell, we first used CondFDR to corrects nominal p-values of SNPs of an IIM/B-cell lymphoma subtype by conditioning on the p-values of the corresponding SNPs of the paired B-cell lymphoma/IIM subtype. A conservative conjFDR of a SNP was obtained by maximising the $\text{condFDR}_{\text{IIM|B-cell lymphoma}}$ and $\text{condFDR}_{\text{B-cell lymphoma|IIM}}$ of that SNP. A significant jointly associated SNP had corrected p-value < 0.05. Lead jointly associated SNPs/loci were defined by clumping.

To avoid artificial genetic enrichment, we adjusted the nominal p-value of each SNP by using intergenic inflation control and excluded SNPs located in the HLA and the chromosome 8 inversion regions from fitting the conditional empirical cumulative distribution function (354).

3.3.9 Study V: The impact of cancer on the prognosis of patients with IIM

Study V was a cohort study following patients to cancer and death events since IIM diagnosis. Patients with incident IIM were identified in the NPR based on the algorithm presented in **Table 5**. An overview of the study design and statistical methods are described below in separate sections.

3.3.9.1 Cancer ascertainment and classification

Via linkage to the SCR, we identified all cancer diagnoses of patients with IIM between 1958 and 2020. For each patient, cancer diagnoses happened before the index date were defined as *cancer before IIM* and those occurred at or after the index date were defined as *cancer after IIM*. Furthermore, events of *first cancer after IIM* (defined as the earliest cancer diagnosis among cases of *cancer after IIM)* and *second cancer after IIM* (defined as a primary cancer subsequent to the *first cancer after IIM*) were of interest. Similar to the approach used in **Study III**, we classified cancer cases by tumour sites according to the ICD codes and, from 1993, according to SNOMED codes. We also categorized cancer cases diagnosed after 2003 into different stages according to TNM and FIGO staging systems. Benign tumours and duplicated diagnoses in the SCR were removed.

3.3.9.2 Death ascertainment

We retrieved death information including date of death and underlying cause of death in the CDR. Death records were categorized into deaths from cancer and deaths from other causes according to the ICD-10 chapter (**Table 8**) (10).

Table 8. Defining causes of death based on the ICD-10 codes, adapted from (10).

Idiopathic inflammatory myopathies	M33.0, M33.1, M33.2, M33.9, G72.4
Certain infections and parasitic diseases	A and B chapter
In situ neoplasms, Benign neoplasms, Neoplasms of uncertain or unknown behaviour, Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	D chapter
Malignant neoplasms	C chapter
Mental and behavioral disorders	F chapter
Endocrine, nutritional and metabolic disease	E chapter
Diseases of the nervous system	G chapter
Diseases of eye, adnexa, ear and mastoid process	H chapter
Diseases of the circulatory system	I chapter
Diseases of the respiratory system*	J chapter
Diseases of the digestive system	K chapter
Diseases of skin and subcutaneous tissue	L chapter
Diseases of the musculoskeletal system and connective tissue	M chapter
Diseases of the genitourinary system	N chapter
Symptoms signs and abnormal clinical lab findings	R and Q chapter
Provisional assignment of new diseases of uncertain aetiology or emergency use or Resistance to antimicrobial and antineoplastic drugs	U chapter
External causes of morbidity and mortality	V, W, X, Y chapters

ICD: International Classification of Diseases.

3.3.9.3 Multistate model: states, transitions and time scales

In **Study V**, patients' transitions to *first cancer after IIM*, *second cancer after IIM*, death from cancer and death from other causes since IIM diagnosis were of interest. Each of these transitions could be considered as a survival model and studied independently. However, this approach did not explicitly capture the changes on risk estimation between these events, i.e., the risks of death from cancer or other causes before and after *first cancer after IIM*. To understand these transitions in a better way, we applied multistate modelling where each event of interest was represented as a discrete state and a forwarding movement to a downstream event under a continuous time scale was taken as a transition. Considering co-occurrence of IIM and cancer is a rare condition, inclusion of all events of interest in one multistate model may lead to very small number of events in a state which increases statistical uncertainty. To avoid this issue, we constructed two multistate models with different research focuses, with one model focusing on the cancer-specific mortality before and after *first cancer after IIM* (first cancer multistate model, **Figure 5a**) and another model focusing on

the risk of having a *second cancer after IIM* and the following risk of death (second cancer multistate model, **Figure 5b**).

In the first cancer multistate model, patients' trajectories included transitions from "Incident IIM" state to "First cancer" state (T1), "Death from cancer" state (T2) and "Death from other causes" state (T3), as well as transitions from "First cancer" state to "Death from cancer" state (T4) and "Death from other causes" state (T5). Patients with *cancer before IIM* were included in the initial "Incident IIM" state and those died from cancer underwent T2.

In the second cancer multistate model, the five transitions were transitions from "Incident IIM" state to "First cancer" state (T1) and "Death" state (T2), transitions from "First cancer" state to "Second cancer" state (T3) and "Death" state (T4), and transition from "Second cancer" state to "Death" state (T5).

One important assumption related to multistate model is Markovian property, which asserts that the past history affects a future state solely through the present state. This assumption does not always hold, and violation can lead to biased estimates (355, 356). Time spent in the previous state is a common factor that can lead to violation of Markov assumption. This might also be the case in our study given that a strong risk of death from cancer has been observed within the first year of IIM diagnosis in patients with IIM compared to the general population (10). Assuming duration dependence was the only cause that led to violation of Markov assumption, we could solve this problem by relaxing the Markov assumption via having a Semi-Markov multistate model where follow-up time was reset when patients enter a new state and Markov assumption remains valid in this setting. In this study, we used time since IIM diagnosis as the time scale for the transitions before the "First cancer" state and reset the time scale to time since *first/second cancer after IIM* when patients entering the "First cancer" state and "Second cancer" state.

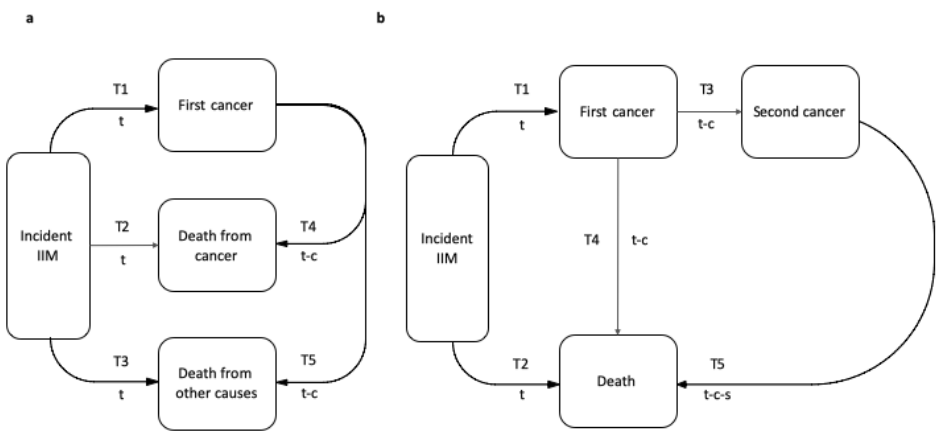


Figure 5. The multistate models consist of states and transitions to different states since IIM diagnosis, with time scale in each transition specified. *a. The first cancer multistate model included transitions to first cancer, death from cancer and death from other causes since IIM diagnosis; b. The second cancer multistate model included transitions to first cancer, second cancer and death since IIM diagnosis. The time scales for transitions were specified as t, time since incident IIM; c, time at first cancer; s, time at second cancer; t-c, time since first cancer; t-c-s, time since second cancer. IIM: Idiopathic inflammatory myopathy.*

3.3.9.4 Covariates

We explored the impacts of a number of covariates on the cancer and death events of interest. **Table 9** shows the definitions of these covariates.

Table 9. Definitions of covariates included in the multistate models

Covariate	Definition	Format	Inclusion in the models
Age at IIM diagnosis	> or ≤ 60 years	Binary	First cancer multistate model: T1-T3 Second cancer multistate model: T1-T2
Sex	Men and women	Binary	All transitions
Cancer before IIM	Yes/no	Binary	First cancer multistate model: all transitions except T2 Second cancer multistate model: all transitions
IIM subtypes	DM or other IIM	Binary	All transitions
Calendar period at IIM diagnosis	1998-2008 or 2009-2020	Binary	First cancer multistate model: T1-T3 Second cancer multistate model: T1-T2
Age at first cancer diagnosis after IIM	> or ≤ 70 years	Binary	First cancer multistate model: T4-T5 Second cancer multistate model: T3-T5
Calendar period at first cancer diagnosis after IIM	1998-2008 or 2009-2020	Binary	First cancer multistate model: T4-T5 Second cancer multistate model: T3-T5
Time from IIM to first cancer	≤ 1 year, (1-3] years or > 3 years	Categorical	First cancer multistate model: T4-T5 Second cancer multistate model: T3-T4
Time from first cancer to second cancer after IIM	≤ 1 year, (1-3] years or > 3 years	Categorical	Second cancer multistate model: T5

IIM: Idiopathic inflammatory myopathy; DM: Dermatomyositis; T: Transition.

3.3.9.5 Nonparametric estimations: state occupation probability and cumulative incidence

Nonparametric estimates can facilitate a better understanding of data and inform model fitness by comparing with parametric estimates. State occupation probability, representing a probability of being in a specific state, may be favourable to clinicians as it can be directly interpreted as a probability or risk. In this study, we estimated state occupation probabilities of each event of interest in both multistate models using the Aalen-Johansen estimator. The probability of being in "First cancer" or "Second cancer" state estimated from the multistate models were affected by the probability of death and thus it could decrease during follow-up. Considering cumulative incidences of *first cancer after IIM* (since IIM diagnosis) and *second cancer after IIM* (since the diagnosis of *first cancer after IIM*) might be of interest to some readers, we also estimated them using the Aalen-Johansen estimator in two separate simple survival models with *first cancer after IIM* and *second cancer after IIM* as the failure event, respectively.

3.3.9.6 Flexible parametric modelling

Flexible parametric modelling, first proposed by Royston and Parmar in 2002, offers several advantages over the Cox regression model (357). First, in flexible parametric model, baseline hazard is modelled and smoothed by restricted cubic spline, thus allowing easy prediction of clinically relevant measures (i.e., state occupation probability and length of stay in a given

state by time t, as well as the absolute and relative differences of these estimates). Time-dependent effect of covariate can also be easily incorporated using spline function (358, 359).

In the two multistate models, we used flexible parametric methods to estimate the transition-specific baseline hazards. Patients were followed till an event of interest (state), emigration or the end of follow-up (2020-12-31), whichever came first. We first fitted crude multistate models and then multivariable multistate models to explore the impacts of the covariates specified in **Table 9** on the cancer and death events. Time-dependent effect was incorporated into the transition for covariates that violated the proportional hazard assumption. With the transition-specific baseline hazards estimated from the models, we predicted the state occupation probability for each state and length of stay of a non-fatal state, using simulation approach. We repeated the simulation with a sample size of 10,000 for 200 times (359).

We found that there were 20 patients who had no cancer diagnosis in the SCR but died from cancer (based on death records in the CDR), probably due to underreporting. These patients might be misclassified as patients without *first cancer after IIM*. Though it was just a small proportion of patients with IIM, we reran the crude parametric multistate models by excluding these 20 patients to examine if it would make a difference.

3.4 Ethical consideration

Public trust, dedication to striving for public benefit through research, and a commitment to personal integrity are fundamental to the success of register-based research, significantly advancing medical and social sciences. Adhering strictly to the General Data Protection Regulation and relevant national legislations in Sweden, we ensured good research practices in all studies presented in this thesis.

First, we obtained ethical permits from the Swedish Ethical Review Authority for conducting **Study I-V**. Motivated by knowledge gaps in myositis research and the unmet needs of patients, our studies aimed to enhance our understanding of the pathogenesis and disease burden of IIM, crucial for improving disease management. The data utilized in this thesis were pseudonymized, de-identified, and aggregated (in **Study IV**), with thorough considerations given to potential risks associated with breaches of personal data. Every stage, from planning to ethical and data application, through execution to publication, was meticulously approached to minimize risks and potential harm to individuals, society, and the environment.

We strictly requested only the data required to answer the research questions, with all data transfers from authorities safeguarded—either encrypted on flash drives or via a secure file transfer platform at university with shell protection. Data storage occurred on department servers with shell protection, and access was granted solely to authorized researchers within our team. Moreover, all team members underwent training on accessing and utilizing the data appropriately.

Our focus on high-quality data and appropriate methodologies ensured the integrity of all analyses. We were acutely aware of methodological limitations and interpreted findings cautiously through comprehensive author discussions to avoid misleading conclusions that might adversely affect individuals.

Given the rarity of IIM, even with de-identified data, patients could be potentially identified through sparse data presentations. We apologized that we did not give thorough considerations concerning this potential risk in **Study I-III** and presented sparse data with number of cases less than five. We have made corresponding adjustment in **Study V** where data were consistently presented at the group level with at least five cases to protect anonymity.

All studies were supported by the Swedish Research Council, necessitating the free accessibility of our findings. **Study I-V** have been published with open access.

4 Key findings

The key findings of each study are outlined below. For more detailed information, please refer to the respective publications and manuscripts.

4.1 Study I: Familial aggregation and heritability of IIM

Population and family structure

We identified 1,620 patients with IIM. Their median birth year and median age at diagnosis was 1949 (interquartile range, IQR 1941-1964) and 57 years (IQR 44-66), respectively. Among these patients, 59% were women. Other IIM was the most prevalent IIM subtype (n=991, 61%), followed by DM (n=501, 31%) and JIIM (n=128, 8%). Due to the matching design, we found similar demographic characteristics in the 7,797 matched comparators without IIM. Furthermore, patients with IIM shared similar family structure with the matched comparators without IIM (**Table 10**). The mean number of first-degree relatives per family unit was 4.7 among patients with IIM, compared to 4.8 in matched comparator without IIM.

Table 10. Family structures of patients with IIM and individuals without IIM, and demographic characteristics of their first-degree relatives

	Patients with IIM		Individuals without IIM	
Any first-degree relatives, n, mean±SD	7,615	4.7±2.2	37,309	4.8±2.1
Parents	2,306	1.4±0.8	11,414	1.5±0.7
Women, n (%)	1,253 (54)		6,314 (55)	
Birth year, median (IQR)	1926 (1916-1943)		1926 (1916-1943)	
Full siblings	2,464	1.5±1.5	11,685	1.5±1.5
Women, n (%)	1,238 (50)		5,863 (50)	
Birth year, median (IQR)	1951 (1943-1962)		1950 (1943-1963)	
Offspring	2,845	1.8±1.3	14,210	1.8±1.3
Women, n (%)	1,335 (47)		6,960 (49)	
Birth year, median (IQR)	1975 (1966-1987)		1975 (1967-1989)	

IIM: Idiopathic inflammatory myopathy; IQR: Interquartile range.

Familial aggregation

There were 13 patients with IIM who had at least one first-degree relatives affected by IIM, versus 16 in matched comparators without IIM, corresponding to an adjusted OR (aOR) of 4.3 (95%CI 2.0-9.3) (**Table 11**). We found no visits indicating muscular dystrophies or metabolic myopathies in the 13 cases of familial IIM. We observed weaker but still significant familial associations of IIM in the analyses treating first-degree relative pairs as independent units. The aOR was 2.6 (95%CI 1.8-3.8) for any first-degree relative pairs and 2.5 (95%CI 1.6-4.0) for full sibling pairs. We noted consistent findings in the sensitivity

analyses using the strict definition of IIM and including only first-degree relatives alive in 2001.

Table 11. The aORs of having first-degree relatives affected by IIM in patients with IIM compared to individuals without IIM

	Patients with IIM, n/N (%)	Individuals without IIM, n/N (%)	aOR[a] (95% CI)	aOR[b] (95% CI)
≥ 1 relative	13/1620 (0.80)	16/7797 (0.21)	4.32 (2.00-9.34)	-
Any first-degree relatives	13/7615 (0.17)	16/37309 (0.04)	2.61 (1.80-3.78)	2.61 (1.80-3.79)
Parents	< 5	< 5	-	-
Full siblings	9/2464 (0.37)	10/11685 (0.09)	2.54 (1.62-3.99)	2.53 (1.62-3.96)
Offspring	< 5	< 5	-	-

aOR: Adjusted odds ratio; IIM: Idiopathic inflammatory myopathy; n: Number of exposed cases; N: Total number of individuals/relative pairs included; ≥ 1 relative: Comparison between patients with IIM and individuals without IIM; Any first-degree relatives, parents, full siblings and offspring: Comparison between relative pairs of patients with IIM and relative pairs of individuals without IIM.
a Controlled for sex, birth year and residential area of index persons.
b Controlled for sex, birth year and residential area of index persons, sex and birth year of first-degree relatives.

Heritability

With a prevalence of IIM of 0.014%, the family-based heritability of IIM among any first-degree relatives was 22% (95%CI 12%-31%) and it was 24% (95%CI 13%-37%) among full siblings. These estimates were robust to the changes of the prevalence of IIM (**Figure 6**).

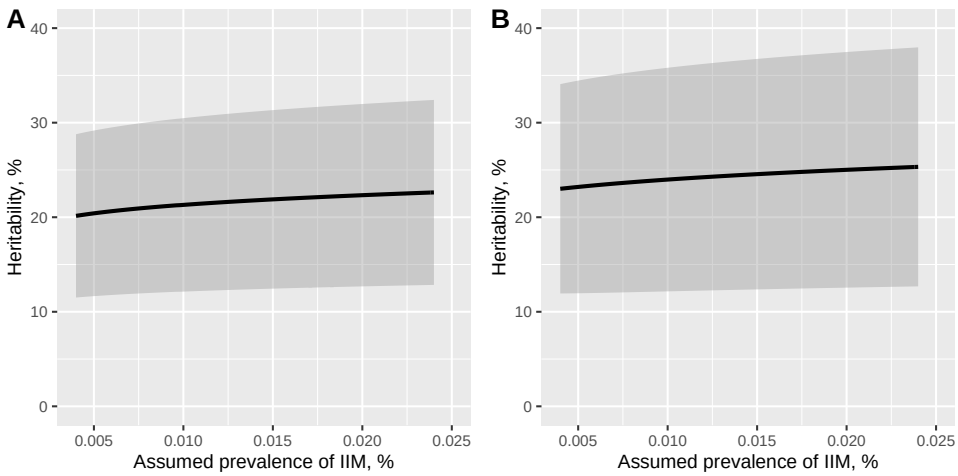


Figure 6. The point estimates with 95% confidence intervals (grey shading) of the heritability of IIM using the prevalence of IIM ranging from 0.004% to 0.024%, with a fixed interval of 0.0002% for A. any first-degree relatives and B. full siblings. *IIM: Idiopathic inflammatory myopathy.*

4.2 Study II: Familial autoimmunity of IIM

Study II had the same study population as in **Study I**. For information on patient characteristics and family structure among patients with IIM and matched comparators without IIM, please read section 4.1 and **Table 10**.

Prevalence of autoimmune diseases

All autoimmune diseases except MS were significantly more prevalent in patients with IIM compared to matched comparators without IIM (**Table 12**).

Table 12. The prevalence of each autoimmune disease in patients with IIM and individuals without IIM

	Patients with IIM (n=1,620)	Individuals without IIM (n=7,797)
Rheumatoid arthritis[a]	182 (11)	116 (1)
Rheumatic inflammatory diseases[a]	352 (22)	63 (1)
Multiple sclerosis	5 (0.3)	25 (0.3)
Inflammatory bowel diseases[a]	35 (2)	108 (1)
Type 1 diabetes mellitus[a]	9 (1)	14 (0.2)
Autoimmune thyroid diseases[a]	215 (13)	696 (9)
Celiac diseases[a]	27 (2)	35 (0.5)
Myasthenia gravis[a]	10 (1)	< 5

IIM: Idiopathic inflammatory myopathy.
a *p* values from Chi-square test < 0.05

Familial association

In the analyses by the number of affected first-degree relatives, we observed significant familial associations only for other RIDs and CeD (**Figure 7**). Patients with IIM were more likely to have ≥ 1 (aOR=1.4 95%CI 1.1-1.8) and ≥ 2 first-degree relatives (aOR=2.4 95%CI 0.9-6.5) affected by other RIDs than matched comparators without IIM although only the association with ≥ 1 first-degree relative affected was statistically significant. A tendency of dose-dependent effect was also observed for CeD but only the familial association with ≥ 2 first-degree relatives affected was statistically significant (aOR=3.6 95%CI 1.3-9.9). We observed additional familial associations with IIM for IBD (aOR=1.2 95%CI 1.0-1.4) and AITD (aOR=1.1 95%CI 1.0-1.2) when treating first-degree relative pairs as independent units. Furthermore, in subgroup analyses by types of kinship, we observed significant familial associations with IIM in parents and offspring for other RIDs and IBD, in parents and full siblings for CeD, and in offspring only for MS and AITD. Analyses with adjustments of an studied autoimmune disease in index individuals or IIM in first-degree relatives, use of strict definitions and inclusion of first-degree relatives alive in 2001 only showed no different results.

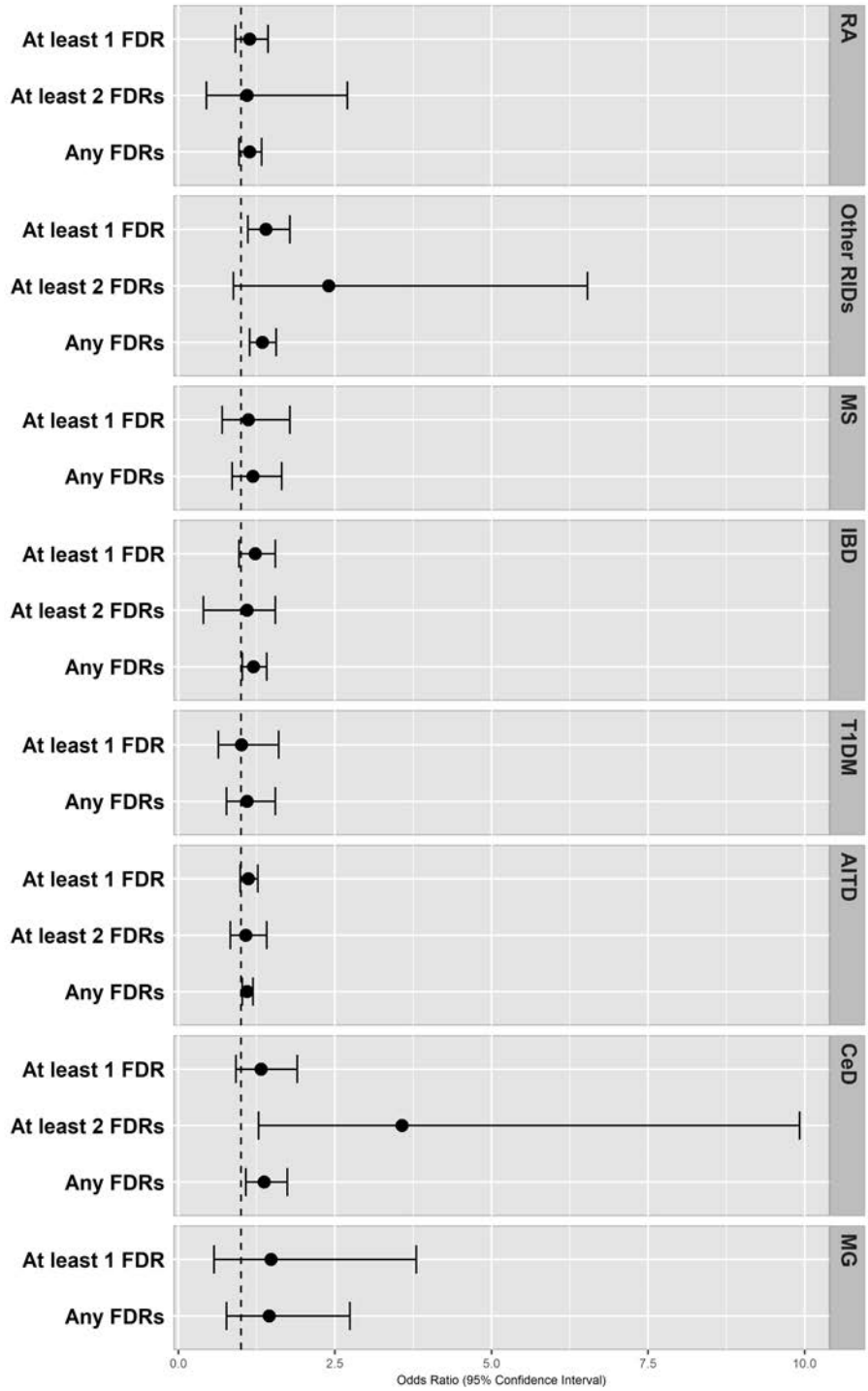


Figure 7. The adjusted odds ratios of familial associations between idiopathic inflammatory myopathies and different autoimmune diseases by the number of affected first-degree relatives and in any first-degree relative pairs. *FDRs: First-degree relatives; RA: Rheumatoid arthritis; RIDs: Rheumatic inflammatory diseases; MS: Multiple sclerosis; IBD: Inflammatory bowel diseases; T1DM: Type 1 diabetes mellitus; AITD: Autoimmune thyroid diseases; CeD: Celiac disease: MG: Myasthenia gravis.*

4.3 Study III: Familial co-aggregation of IIM and cancer

Using a modified definition of IIM, we identified slightly more patients with IIM (n=1,639) and matched comparators with IIM (n=7,878) (**Table 13**). Basic characteristics and family structure were similar to **Study I** and **II**. We found a higher frequency of lifetime cancer in patients with IIM compared to matched comparator without IIM (27% versus 20%). The prevalence of lifetime IIM was higher in first-degree relatives of patients with IIM than in those of matched comparators without IIM.

Table 13. Family structures of patients with IIM and individuals without IIM, and characteristics of their first-degree relatives

	Patients with IIM (n=1,639)	Individuals without IIM (n=7,878)
Women, n (%)	959 (59%)	4,675 (59%)
Birth year, median (IQR)	1949 (1941-1963)	1949 (1941-1963)
Living in Southern Sweden, n (%)	1,350 (82%)	6,485 (82%)
Age at inclusion, median (IQR)	58 (45-66)	58 (45-66)
IIM subtypes, n (%)		
Juvenile IIM[a]	127 (8%)	-
Dermatomyositis[b]	506 (31%)	-
Other IIM[c]	1,006 (61%)	-
Lifetime cancer, n (%)	436 (27%)	1,553 (20%)
Parents, n (median (IQR))	3,066 (2 (2-2))	14,918 (2 (2-2))
Women, n (%)	1,564 (51%)	7,565 (51%)
Birth year, median (IQR)	1920 (1910-1936)	1920 (1912-1937)
Lifetime IIM, n (%)	< 5	< 5
Full siblings, n (median IQR)	2,494 (1 (0-2))	11,819 (1 (0-2))
Women, n (%)	1,249 (50%)	5,917 (50%)
Birth year, median (IQR)	1951 (1943-1962)	1950 (1943-1963)
Lifetime IIM, n (%)	7 (0%)	8 (0%)
Offspring, n (median (IQR))	2,900 (2 (1-3))	14,390 (2 (1-3))
Women, n (%)	1,355 (47%)	7,033 (49%)
Birth year, median (IQR)	1975 (1966-1987)	1975 (1967-1988)
Lifetime IIM, n (%)	< 5	-

IIM: Idiopathic inflammatory myopathy; IQR: Interquartile range.
a Age at diagnosis ≤ 18 years of age.
b With diagnostic code M33.1 and age at diagnosis > 18 years of age.
c With diagnostic code M33.2, M33.9 or G72.4 and age at diagnosis > 18 years of age.

Familial co-aggregation of IIM and cancer

Among 8,460 first-degree relatives of patients with IIM, 1,598 (18.9%) had a lifetime cancer diagnosis compared to 7,370 (17.9%) in first-degree relatives of matched comparators without IIM, resulting an aOR of 1.04 (95%CI 0.99-1.09). We found similar estimates in subgroup analyses by IIM subtypes and types of kinship.

In the analyses exploring if familial co-aggregation of IIM and cancer differed by sex concordance of first-degree relative pairs and IIM subtypes, we observed statistically significant association only in male first-degree relative pairs of patients with DM (aOR=1.4 95% CI 1.2-1.7), after adjusting for multiple testing (**Table 14**). Moreover, similar point estimates were found in father-son pairs and brother pairs of patients with DM but these associations did not reach statistical significance after correcting for multiple testing.

In the analyses by age at cancer diagnosis, we observed a familial co-aggregation of IIM and cancer diagnosed at 50 years of age or olders (aOR=1.1 95%CI 1.0-1.3) in full siblings and a familial co-aggregation of IIM and cancer diagnosed before 50 years of age in offspring (aOR=1.2 95%CI 1.0-1.3).

Table 14. The familial co-aggregation of IIM and cancer overall by sex concordance of any first-degree relative pairs and IIM subtypes

	Patients with IIM, n (%)	Individuals without IIM, n (%)	aOR (95% CI)[a]
Female concordant pairs			
Any IIM	504 (20.6%)	2,392 (19.5%)	1.06 (0.97-1.16)
DM	177 (21.6%)	699 (20.1%)	1.05 (0.91-1.22)
Juvenile IIM	10 (5.8%)	51 (6.7%)	0.88 (0.47-1.64)
Other IIM	317 (21.9%)	1,237 (20.1%)	1.07 (0.96-1.19)
Male concordant pairs			
Any IIM	328 (18.2%)	1,345 (16.2%)	1.17 (1.05-1.30)
DM	103 (20.1%)	301 (15.4%)	1.39 (1.15-1.68)[b]
Juvenile IIM	< 5	6 (1.9%)	-
Other IIM	223 (18.1%)	893 (17.1%)	1.09 (0.96-1.25)
Discordant pairs			
Any IIM	766 (18.2%)	3,633 (17.7%)	0.98 (0.92-1.06)
DM	232 (17.4%)	990 (17.7%)	0.94 (0.82-1.07)
Juvenile IIM	8 (3.3%)	46 (4.2%)	0.79 (0.39-1.60)
Other IIM	526 (19.9%)	2,152 (19.1%)	1.01 (0.93-1.10)

IIM: Idiopathic inflammatory myopathy; aOR: Adjusted Odds ratio; CI: Confidence interval; DM: Dermatomyositis.
a Adjusted for sex, birth year and residential area of index individuals, and birth year of first-degree relatives. Model assuming no direct effects between IIM and cancer.
b The false discovery rate (FDR) threshold was 0.0009. The only association (observed p value=0.0006) remained statistically significant after FDR correction.

Familial co-aggregation of IIM and specific cancer types

In the exploratory analyses by specific cancer types, we only observed statistically significant familial associations for myeloid malignancies (aOR=2.3 95%CI 1.4-3.6) and liver cancer (aOR=2.0 95%CI 1.2-3.3) in male first-degree relatives of patients with IIM after adjusting for multiple testing (**Table 15**).

Table 15. The familial co-aggregation of IIM and specific cancer types among first-degree relatives by sex concordance of relative pairs

	Female concordant		Male concordant		Discordant	
	Patients with IIM/individuals without IIM, n (%)	aOR (95% CI)[a]	Patients with IIM/individuals without IIM, n (%)	aOR (95% CI)[a]	Patients with IIM/individuals without IIM, n (%)	aOR (95% CI)[a]
Haematological malignancies	25 (1.0%)/118 (1.0%)	0.99 (0.72-1.35)	35 (1.9%)/ 119 (1.4%)	1.30 (0.99-1.71)	50 (1.2%)/241 (1.2%)	1.02 (0.81-1.30)
Lymphoid	22 (0.9%)/80 (0.7%)	1.19 (0.87-1.63)	25 (1.4%)/102 (1.2%)	1.12 (0.80-1.57)	38 (0.9%)/189 (0.9%)	0.99 (0.76-1.30)
B cell lymphoma	17 (0.7%)/55 (0.4%)	1.33 (0.94-1.87)	17 (0.9%)/60 (0.7%)	1.25 (0.84-1.85)	26 (0.6%)/105 (0.5%)	1.13 (0.82-1.56)
Hodgkin lymphoma	< 5/8 (0.1%)	-	< 5/11 (0.1%)	-	< 5/20 (0.1%)	-
Myeloid malignancies	< 5/34 (0.3%)	-	9 (0.5%)/13 (0.2%)	2.27 (1.43-3.60)[b]	10 (0.2%)/47 (0.2%)	1.07 (0.63-1.82)
Chronic myeloid leukaemia	< 5/11 (0.1%)	-	< 5/5 (0.1%)	-	6 (0.1%)/14 (0.1%)	2.38 (1.18-4.81)
Acute myeloid leukaemia	< 5/16 (0.1%)	-	6 (0.3%)/7 (0.1%)	2.50 (1.32-4.72)	< 5/22 (0.1%)	-
Solid cancers	479 (19.6%)/2,274 (18.6%)	1.06 (0.97-1.16)	293 (16.2%)/1,226 (14.8%)	1.14 (1.02-1.27)	716 (17.0%)/3,391 (16.5%)	0.98 (0.91-1.06)
Low-grade brain	5 (0.2%)/37 (0.3%)	0.80 (0.37-1.70)	9 (0.5%)/23 (0.3%)	1.86 (1.10-3.14)	11 (0.3%)/67 (0.3%)	0.77 (0.46-1.28)
Head and neck	13 (0.5%)/72 (0.6%)	0.88 (0.55-1.42)	6 (0.3%)/65 (0.8%)	0.49 (0.23-1.02)	23 (0.5%)/147 (0.7%)	0.72 (0.51-1.03)
Non-oropharyngeal	13 (0.5%)/68 (0.6%)	0.93 (0.58-1.49)	5 (0.3%)/53 (0.6%)	0.49 (0.21-1.13)	20 (0.5%)/124 (0.6%)	0.74 (0.51-1.08)
Breast	109 (4.5%)/523 (4.3%)	1.04 (0.87-1.23)	< 5/< 5	-	61 (1.4%)/340 (1.7%)	0.86 (0.69-1.07)
Lung	21 (0.9%)/95 (0.8%)	1.15 (0.81-1.63)	35 (1.9%)/90 (1.1%)	1.67 (1.28-2.19)	44 (1.0%)/229 (1.1%)	0.89 (0.69-1.14)
Liver	< 5/15 (0.1%)	-	7 (0.4%)/12 (0.1%)	2.01 (1.21-3.33)[b]	7 (0.2%)/38 (0.2%)	0.88 (0.45-1.71)
Colorectal	47 (1.9%)/188 (1.5%)	1.20 (0.92-1.55)	35 (1.9%)/162 (1.9%)	1.00 (0.75-1.34)	75 (1.8%)/375 (1.8%)	0.95 (0.78-1.16)
Pancreatic	18 (0.7%)/89 (0.7%)	0.86 (0.57-1.28)	9 (0.5%)/56 (0.7%)	0.86 (0.47-1.55)	33 (0.8%)/150 (0.7%)	0.94 (0.70-1.26)
Kidney	12 (0.5%)/34 (0.3%)	1.83 (1.11-3.02)	8 (0.4%)/37 (0.4%)	1.09 (0.57-2.08)	20 (0.5%)/87 (0.4%)	1.03 (0.70-1.52)
Bladder	6 (0.2%)/44 (0.4%)	0.72 (0.37-1.41)	19 (1.1%)/76 (0.9%)	1.19 (0.79-1.81)	29 (0.7%)/122 (0.6%)	1.07 (0.79-1.44)
Uterine	24 (1.0%)/103 (0.8%)	1.07 (0.76-1.51)	- -	-	- -	-
Cervical	116 (4.7%)/514 (4.2%)	1.13 (0.96-1.33)	-	-	-	-
Ovarian	17 (0.7%)/77 (0.6%)	1.11 (0.75-1.64)	-	-	-	-
Prostate	-	-	78 (4.3%)/394 (4.7%)	0.92 (0.75-1.12)	-	-
Melanoma	12 (0.5%)/127 (1.0%)	0.53 (0.32-0.89)	21 (1.2%)/52 (0.6%)	1.60 (1.17-2.20)	42 (1.0%)/168 (0.8%)	1.16 (0.91-1.47)
Non-melanoma skin	33 (1.4%)/145 (1.2%)	1.04 (0.78-1.38)	20 (1.1%)/95 (1.1%)	0.99 (0.67-1.46)	49 (1.2%)/214 (1.0%)	1.03 (0.80-1.32)

IIM: Idiopathic inflammatory myopathy; aOR: Adjusted Odds ratio; CI: Confidence interval.
a Adjusted for sex, birth year and residential area of index individuals, and sex and birth year of first-degree relatives. For subgroup analyses of sex-concordant relative pairs and for sex-specific cancers, sex of first-degree relatives was not adjusted. Model assuming no direct effects between IIM and cancer.
b The false discovery rate (FDR) threshold was 0.0008. Only these associations (observed *p* values: 0.0002 for liver cancer and 0.0005 for myeloid malignancies) remained statistically significant after FDR correction.

Due to small number of cases, only some of the cancer types were analysed in all three types of kinship and these associations did not differ from the those in any first-degree relatives. Furthermore, additionally adjusting for cancer or specific cancer types in index individuals or IIM in first-degree relatives in the analyses showed results in line with the main analyses.

4.4 Study IV: Genetic overlap between IIM and common B-cell lymphoma subtypes

The number of common SNPs/loci from each data set included the LAVA and the PleioFDR analyses for each disease pair are presented in **Table 16**.

Local genetic correlations

The GWAS data

Among the loci showing significant heritability with both diseases of the GWAS data, there were in total seven loci presenting significant genetic correlations, with strengths varying from -0.5 to 0.72; the detected loci were [chr17:62,112,374-63,548,724] for DM-FL, [chr6:32,629,240-32,682,213] and [chr9:114443476-115390112] for DM-CLL, [chr1:238,094,456-238,704,857] for PM-DLBCL, [chr3:158460059-159478751],

[chr7:55161395-56303513] and [chr12:126871453-127545377] for PM-MZL (**Table 17**). The squared local genetic correlation (r_g^2) represented the proportion of the heritability of an IIM subtype that could be explained by the heritability of a B-cell lymphoma subtype in a given locus. We found the highest r_g^2 in the locus [chr6:32,629,240-32,682,213] that showed a strong r_g between DM and CLL, that is, about half of the local heritability of DM could be explained by that of CLL in this locus. For the other detected local r_g, the r_g^2 ranged from 23% to 39%. Of note, the local heritability presented in each disease pair was low.

Table 16. The number of SNPs and loci included in the LAVA analyses and the pleioFDR analyses for the GWAS and the ImmunoChip disease pairs of IIM and common B-cell lymphoma subtypes

		LAVA		PleioFDR
	Number of common SNPs[a]	**Number of total loci in univariate analysis (both:IIM:B-cell lymphomas)[b]**	**Number of loci in bivariate analysis**	**Number of common SNPs[a]**
GWAS data				
DM-DLBCL	6,415,387	2,471 (1,836:1,876:2,431)	680	6,260,170
DM-FL	6,416,691	2,469 (1,829:1,889:2,409)	611	6,259,969
DM-CLL	6,417,297	2,474 (1,828:1,881:2,421)	741	6,260,163
DM-MZL	6,384,382	2,451 (1,786:2,337:1,900)	813	6,239,939
PM-DLBCL	6,415,351	2,466 (1,073:1,111:2,428)	326	6,260,153
PM-FL	6,416,654	2,465 (1,052:1,111:2,406)	284	6,259,952
PM-CLL	6,417,260	2,472 (1,068:1,118:2,422)	369	6,260,145
PM-MZL	6,384,346	2,460 (1,025:1,147:2,338)	264	6,239,921
ImmunoChip data				
DM-DLBCL	882,537	1,697 (614:1,020:1,291)	37	850,547
DM-FL	882,935	1,694 (607:1,017:1,284)	34	850,673
DM-CLL	882,730	1,698 (616:1,019:1,295)	36	850,596
DM-MZL	880,432	1,671 (567:1,023:1,215)	24	849,056
PM-DLBCL	882,537	1,740 (627:1,078:1,289)	35	850,547
PM-FL	882,935	1,746 (614:1,078:1,282)	36	850,673
PM-CLL	882,730	1,712 (661:1,078:1,295)	36	850,596
PM-MZL	880,432	1,691 (602:1,080:1,213)	23	849,056

SNP: Single nucleotide polymorphism; LAVA: Local analysis of (co) variant association; pleioFDR: Pleiotropy-informed false discovery rate; GWAS: Genome-wide association study; IIM: Idiopathic inflammatory myopathy; DM: Dermatomyositis; PM: Polymyositis; common B-cell lymphoma subtypes included diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL).
a Nonoverlapping SNPs between diseases and SNPs that were not aligned to the reference panel in the LAVA or in the pleioFDR were excluded.
b Some loci were dropped in the univariate analyses due to number of SNPs < 2 in the locus or negative variance estimate for both diseases.

The ImmunoChip data

We identified more significant local genetic correlations for the ImmunoChip disease pairs and most of these correlations were found in the HLA region (**Table 17**). There were three HLA loci detected for DM-DLBCL, four HLA loci and one non-HLA locus on chromosome 1 detected for DM-FL, three HLA loci detected for DM-CLL, four HLA loci and one non-HLA loci on chromosome 9 detected for PM-DLBCL, and one HLA locus for PM-CLL. We observed a very strong genetic correlation in the locus [chr6:32,208,902-32,454,577] for DM-CLL, corresponding to a r_g^2 of 70%. The loci [chr6:30,715,007-31,106,493] and [chr6: 31,320,269-31,427,209] were associated with both disease pairs DM-DLBCL and PM-DLBCL. Moreover, the locus [chr6:31,427,210-32,208,901] was associated with DM-DLBCL and DM-FL. The loci [chr6:32,208,902-32,454,577] and [chr6:32,586,785-32,629,239] were associated with DM-FL and DM-CLL. The locus [chr6:32,629,240-32,682,213] was associated with DM-CLL and PM-CLL.

Table 17. The loci with significant local genetic correlations (r_g) for the GWAS and the ImmunoChip disease pairs of IIM and common B-cell lymphoma subtypes

Disease pair	Chr	Start	Stop	n.snps	n.pcs	h^2_{IIM}	$h^2_{B\text{-cell lymphoma}}$	r_g (95% CI)	r_g^2 (95% CI)	P value
GWAS data										
DM-FL	17	62112374	63548724	2483	283	0,02	0,01	-0,50 (-0,72 - -0,32)	0,25 (0,10 - 0,52)	$5,29 \times 10^{-7}$
DM-CLL	6	32629240	32682213	45	20	0,04	0,01	0,72 (0,53 - 0,89)	0,51 (0,28 - 0,79)	$1,38 \times 10^{-8}$
DM-CLL	9	114443476	115390112	2459	204	0,02	0,02	0,62 (0,54 - 0,70)	0,39 (0,30 - 0,49)	$3,39 \times 10^{-38}$
PM-DLBCL	1	238094456	238704857	2081	171	0,00	0,02	0,48 (0,30 - 0,66)	0,23 (0,09 - 0,44)	$7,62 \times 10^{-7}$
PM-MZL	3	158460059	159478751	1961	218	0,00	0,00	0,54 (0,33 - 0,78)	0,29 (0,11 - 0,61)	$3,31 \times 10^{-6}$
PM-MZL	7	55161395	56303513	3322	226	0,02	0,01	0,53 (0,32 - 0,77)	0,28 (0,10 - 0,59)	$3,03 \times 10^{-6}$
PM-MZL	12	126871453	127545377	2021	243	0,01	0,01	0,57 (0,38 - 0,78)	0,33 (0,15 - 0,61)	$6,83 \times 10^{-8}$
ImmunoChip data										
DM-DLBCL	6	30715007	31106493	544	60	0,03	0,01	0,77 (0,48 - 1,00)	0,59 (0,23 - 1,00)	$1,05 \times 10^{-5}$
DM-DLBCL	6	31320269	31427209	299	42	0,04	0,01	0,70 (0,45 - 0,98)	0,49 (0,20 - 0,96)	$1,08 \times 10^{-5}$
DM-DLBCL	6	31427210	32208901	400	79	0,04	0,01	0,58 (0,33 - 0,88)	0,34 (0,11 - 0,77)	$5,86 \times 10^{-5}$
DM-FL	1	205917549	208162951	2918	288	0,01	0,01	0,51 (0,27 - 0,78)	0,26 (0,07 - 0,60)	$6,59 \times 10^{-5}$
DM-FL	6	31427210	32208901	400	79	0,04	0,03	0,35 (0,20 - 0,51)	0,13 (0,04 - 0,26)	$2,47 \times 10^{-5}$
DM-FL	6	32208902	32454577	152	23	0,03	0,04	0,42 (0,25 - 0,58)	0,18 (0,06 - 0,34)	$3,43 \times 10^{-6}$
DM-FL	6	32539568	32586784	23	13	0,01	0,04	0,43 (0,22 - 0,63)	0,19 (0,05 - 0,40)	$1,65 \times 10^{-4}$
DM-FL	6	32586785	32629239	75	37	0,04	0,04	0,50 (0,35 - 0,63)	0,25 (0,12 - 0,40)	$3,01 \times 10^{-9}$
DM-CLL	6	32208902	32454577	151	22	0,03	0,01	0,84 (0,62 - 1,00)	0,70 (0,39 - 1,00)	$2,18 \times 10^{-7}$
DM-CLL	6	32586785	32629239	75	37	0,04	0,01	0,77 (0,53 - 1,00)	0,59 (0,28 - 1,00)	$1,52 \times 10^{-6}$
DM-CLL	6	32629240	32682213	45	20	0,07	0,01	0,54 (0,37 - 0,70)	0,29 (0,14 - 0,49)	$1,77 \times 10^{-8}$
PM-DLBCL	6	30070718	30715006	617	77	0,03	0,01	0,63 (0,36 - 0,96)	0,40 (0,13 - 0,92)	$5,4 \times 10^{-5}$
PM-DLBCL	6	30715007	31106493	544	60	0,05	0,01	0,62 (0,35 - 0,94)	0,39 (0,12 - 0,88)	$6,54 \times 10^{-5}$
PM-DLBCL	6	31320269	31427209	299	43	0,06	0,01	0,68 (0,44 - 0,94)	0,46 (0,19 - 0,88)	$3,3 \times 10^{-6}$
PM-DLBCL	9	138995792	140097759	1042	192	0,02	0,01	-0,46 (-0,70 - -0,23)	0,21 (0,05 - 0,48)	$1,69 \times 10^{-4}$
PM-CLL	6	32629240	32682213	45	20	0,10	0,01	0,51 (0,33 - 0,67)	0,26 (0,11 - 0,45)	$4,48 \times 10^{-6}$

GWAS: Genome-wide association study; IIM: Idiopathic inflammatory myopathy; DM: Dermatomyositis; PM: Polymyositis; common B-cell lymphoma subtypes included diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL); n.snps: Number of single nucleotide polymorphisms in each locus; chr: Chromosome number; n.pcs: Number of principal components in each locus; h^2_{IIM}: Local heritability of IIM at observed scale; $h^2_{B\text{-cell lymphoma}}$: Local heritability of B-cell lymphoma at observed scale; CI: Confidence interval.

Heatmaps of the r_g in the HLA region for the GWAS and the ImmunoChip disease pairs

We also compared the local genetic correlations in all tested HLA loci between the GWAS and the ImmunoChip data (**Figure 8**). The SNP overlapping rates between the two data sets ranged from 69% to 100%. Overall, we observed consistent findings in terms of strength and direction in loci that were tested in both the GWAS and the ImmunoChip disease pairs. The genetic correlation detected in the locus [chr6:32,629,240-32,682,213] in the GWAS disease pair DM-CLL was replicated in the ImmunoChip disease pair DM-CLL.

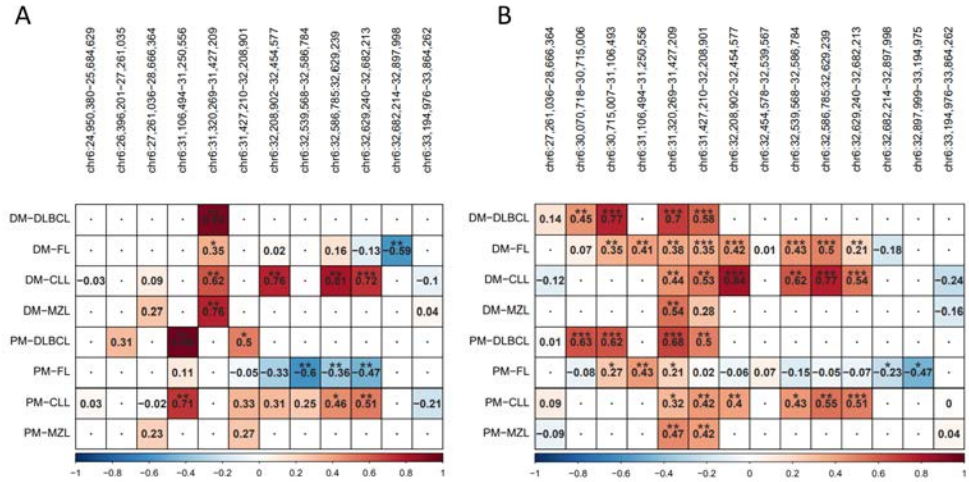


Figure 8. The heatmaps presenting the local genetic correlations (rg) of all tested human leukocyte antigen loci for the GWAS (A) and the ImmunoChip (B) disease pairs of IIM and common B-cell lymphoma subtypes. *GWAS: Genome-wide association study; IIM: Idiopathic inflammatory myopathy; DM: Dermatomyositis; PM: Polymyositis; common B-cell lymphoma subtypes included diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL); *** P value < 0.00 (significant threshold); ** P value < 0.01; * P value < 0.05.*

Pleiotropic enrichment

We observed typical pleiotropic enrichment only in the ImmunoChip disease pairs DM-CLL and PM-CLL. For SNPs with p-values < 0.001 in CLL, there were up to 400- and 225-fold enrichment of SNPs with -$\log_{10}$(p)=7 for DM and PM compared to all SNPs, respectively. Similarly, there were 45- and 60-fold enrichment of SNPs with -$\log_{10}$(p)=7 for CLL when conditioning on SNPs with p-values < 0.001 in DM and PM, respectively.

Jointly associated SNPs

The GWAS data

Among the GWAS disease pairs, we identified in total 12 lead jointly associated SNPs and the majority of them resided in the HLA region: one associated with DM-DLBCL, two with DM-FL, one with DM-CLL, one with DM-MZL, three with PM-DLBCL, two with PM-FL, one with PM-CLL and one with PM-MZL (**Table 18**). The lead SNP rs3130923 was associated with both PM-DLBCL and PM-MZL.

The ImmunoChip data

We detected more jointly associated SNPs among the ImmunoChip disease pairs and again most of them were located in the HLA region (**Table 18**). The two lead jointly associated SNPs rs9270493 and rs2596500 detected respectively for the disease pairs PM-FL and PM-CLL of the GWAS data were replicated in the ImmunoChip data.

Table 18. The lead jointly associated SNPs detected for the GWAS and the ImmunoChip disease pairs of IIM and common B-cell lymphoma subtypes

Disease pair	SNP ID	Reference allele	Effect allele	Z_{IIM}	$Z_{B\text{-cell lymphoma}}$	Chr	Position	P value$_{IIM}$	P value$_{B\text{-cell lymphoma}}$	conjFDR
GWAS data										
DM-DLBCL	rs3094005	G	T	7,00	4,32	6	31465047	$2,64 \times 10^{-12}$	$1,56 \times 10^{-5}$	$8,21 \times 10^{-3}$
DM-FL	rs2596462	T	C	4,23	4,80	6	31414580	$2,39 \times 10^{-5}$	$1,56 \times 10^{-6}$	$9,00 \times 10^{-3}$
DM-FL	rs7450278	C	T	6,41	-4,97	6	32439048	$1,46 \times 10^{-10}$	$6,68 \times 10^{-7}$	$3,36 \times 10^{-4}$
DM-CLL	rs9273325	G	A	8,61	5,86	6	32623193	$7,50 \times 10^{-18}$	$4,54 \times 10^{-9}$	$2,60 \times 10^{-6}$
DM-MZL	rs3130490	G	T	7,25	4,74	6	31739120	$4,02 \times 10^{-13}$	$2,09 \times 10^{-6}$	$1,04 \times 10^{-3}$
PM-DLBCL	rs1611929	G	A	-3,81	-4,00	3	21457961	$1,41 \times 10^{-4}$	$6,31 \times 10^{-5}$	$4,29 \times 10^{-2}$
PM-DLBCL	rs3130923	G	A	8,51	4,28	6	31462135	$1,69 \times 10^{-17}$	$1,87 \times 10^{-5}$	$8,86 \times 10^{-3}$
PM-DLBCL	rs117408955	G	A	3,90	-4,28	13	112538767	$9,45 \times 10^{-5}$	$1,85 \times 10^{-5}$	$3,07 \times 10^{-2}$
PM-FL	rs130071	G	A	4,24	5,25	6	31116210	$2,26 \times 10^{-5}$	$1,5 \times 10^{-7}$	$9,43 \times 10^{-3}$
PM-FL	rs9270493	T	C	5,20	-5,66	6	32559110	$2,01 \times 10^{-7}$	$1,51 \times 10^{-8}$	$6,95 \times 10^{-5}$
PM-CLL	rs2596500	A	C	8,47	4,04	6	31321267	$2,49 \times 10^{-17}$	$5,42 \times 10^{-5}$	$4,17 \times 10^{-2}$
PM-MZL	rs3130923	G	A	8,52	5,09	6	31462135	$1,65 \times 10^{-17}$	$3,62 \times 10^{-7}$	$1,89 \times 10^{-4}$
ImmunoChip data										
DM-DLBCL	rs3093958	A	G	12,30	4,84	6	31410521	$9,56 \times 10^{-35}$	$1,28 \times 10^{-6}$	$7,68 \times 10^{-5}$
DM-FL	rs10173316	C	T	3,54	3,31	2	152170442	$4,00 \times 10^{-4}$	$9,30 \times 10^{-4}$	$3,78 \times 10^{-2}$
DM-FL	rs9273504	T	C	11,18	3,26	6	32628407	$5,22 \times 10^{-29}$	$1,12 \times 10^{-3}$	$4,40 \times 10^{-2}$
DM-CLL	rs12203592	C	T	6,83	7,32	6	396321	$8,78 \times 10^{-12}$	$2,51 \times 10^{-13}$	$5,74 \times 10^{-10}$
DM-CLL	rs9348747	A	G	-3,41	3,22	6	27002406	$6,60 \times 10^{-4}$	$1,27 \times 10^{-3}$	$3,21 \times 10^{-2}$
DM-CLL	rs3093958	A	G	12,30	3,76	6	31410521	$8,73 \times 10^{-35}$	$1,72 \times 10^{-4}$	$4,14 \times 10^{-3}$
DM-CLL	rs511515	A	G	-3,49	5,03	6	33541507	$4,78 \times 10^{-4}$	$4,99 \times 10^{-7}$	$1,54 \times 10^{-2}$
DM-CLL	rs2872812	G	A	-3,42	-3,36	17	38758650	$6,22 \times 10^{-4}$	$7,89 \times 10^{-4}$	$1,99 \times 10^{-2}$
DM-MZL	rs3093958	A	G	12,31	5,13	6	31410521	$8,24 \times 10^{-35}$	$2,86 \times 10^{-7}$	$1,46 \times 10^{-5}$
PM-DLBCL	rs2596500	A	C	15,37	5,06	6	31321267	$2,54 \times 10^{-53}$	$4,18 \times 10^{-7}$	$2,09 \times 10^{-5}$
PM-FL	rs2023472	A	G	-4,39	3,40	6	30075864	$1,13 \times 10^{-5}$	$6,81 \times 10^{-4}$	$4,55 \times 10^{-2}$
PM-FL	rs9270493	T	C	12,22	-5,63	6	32559110	$2,36 \times 10^{-34}$	$1,82 \times 10^{-8}$	$8,89 \times 10^{-7}$
PM-CLL	rs12203592	C	T	5,38	7,32	6	396321	$7,36 \times 10^{-8}$	$2,51 \times 10^{-13}$	$4,9 \times 10^{-6}$
PM-CLL	rs10946859	A	G	-3,60	-3,38	6	26851785	$3,22 \times 10^{-4}$	$7,18 \times 10^{-4}$	$2,92 \times 10^{-2}$
PM-CLL	rs2523990	A	G	3,84	3,27	6	30077229	$1,25 \times 10^{-4}$	$1,06 \times 10^{-3}$	$4,33 \times 10^{-2}$
PM-CLL	rs2596500	A	C	15,37	3,97	6	31321267	$2,80 \times 10^{-53}$	$7,20 \times 10^{-5}$	$2,62 \times 10^{-3}$
PM-MZL	rs2596500	A	C	15,37	5,31	6	31321267	$2,44 \times 10^{-53}$	$1,07 \times 10^{-7}$	$4,81 \times 10^{-6}$

SNP: Single nucleotide polymorphism; GWAS: Genome-wide association study; IIM: Idiopathic inflammatory myopathy; DM: Dermatomyositis; PM: Polymyositis; common B-cell lymphoma subtypes included diffuse large B-cell lymphoma (DLBCL), follicular lymphoma (FL), chronic lymphocytic leukaemia (CLL) and marginal zone lymphoma (MZL); chr: Chromosome number; Z_{IIM}: Z score of SNP of IIM; $Z_{B\text{-cell lymphoma}}$: Z score of SNP of B-cell lymphoma; conjFDR: Conjunctional false discovery rate; P values were corrected for genomic inflation.

4.5 Study V: The impact of cancer on the prognosis of patients with IIM

Study population

This study included 1,826 patients with adult-onset IIM diagnosed between 1998 and 2020, of 310 patients had *cancer before IIM* and 306 patients had *first cancer after IIM* (**Figure 9**). **Table 19** presents the basic demographic characteristics of all patients overall and by timing of cancer. Briefly, patients with cancer (before and after IIM) were born earlier than all patients with IIM. There were more DM cases in patients with *cancer before IIM* than the other two patient groups.

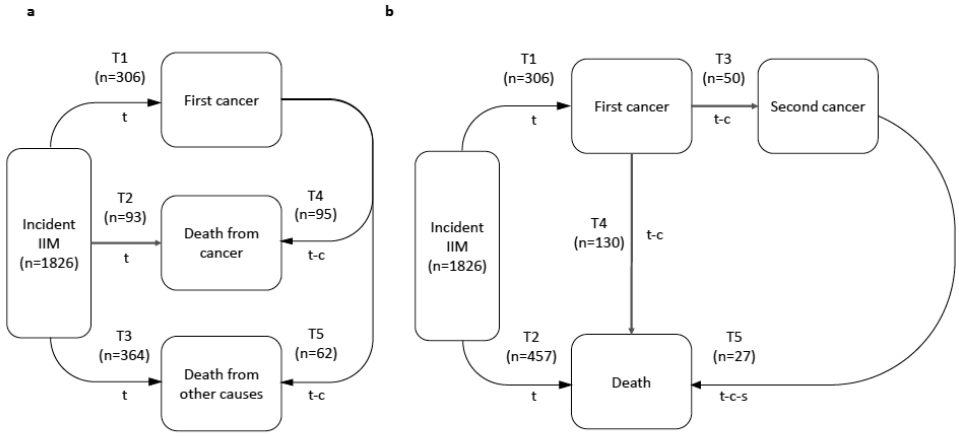


Figure 9. The multistate models consist of states and transitions to different states since IIM diagnosis, with number of patients and time scale in each transition specified. *a. The first cancer multistate model included transitions to first cancer, death from cancer and death from other causes since IIM diagnosis; b. The second cancer multistate model included transitions to first cancer, second cancer and death since IIM diagnosis. The time scales for transitions were specified as t, time since incident IIM; c, time at first cancer; s, time at second cancer; t-c, time since first cancer; t-c-s, time since second cancer. IIM: Idiopathic inflammatory myopathy.*

Table 19. The baseline characteristics of all patients with IIM, overall and stratified by timing of cancer

	All patients (n=1,826)	Patients with cancer before IIM (n=310)	Patients with first cancer after IIM (n=306)
Birth year, median (IQR)	1949 (1939-1963)	1942 (1934-1950)	1942 (1934-1949)
Women, n (%)	1,149 (59.53%)	199 (64.19%)	163 (53.27%)
Age at IIM diagnosis, median (IQR)	62 (48-72)	-	-
Age at cancer diagnosis, median (IQR)	-	64 (52-73)	71 (63-78)
DM, n (%)	638 (34.94%)	143 (46.13%)	116 (37.91%)
Other IIM, n (%)	1,188 (65.06%)	167 (53.87%)	190 (62.09%)

IIM: Idiopathic inflammatory myopathy; IQR: Interquartile range; DM: Dermatomyositis; Among patients with *cancer before IIM*, the characteristics were defined based on the cancer diagnosis that occurred closest in time to the IIM diagnosis.

Cancer

Of 310 patients with *cancer before IIM*, 49 (16%) and 11 (4%) had *first cancer after IIM* and *second cancer after IIM*, respectively (**Table 20**). Among 306 patients with *first cancer after IIM*, 50 (16%) had a second primary cancer. The cumulative incidence of *first cancer after*

IIM was 28% (95% CI 25%-33%) after approximately 23 years of follow-up. With the same duration of follow-up since the diagnosis of *first cancer after IIM*, the cumulative incidence of *second cancer after IIM* was 24% (95% CI 18%-31%). In both patients with *cancer before IIM* and patients with *first cancer after IIM*, the majority of cancer cases were cancer in situ or logically or regionally aggressive at diagnosis, and half of them were diagnosed within three years of the IIM diagnosis. In patients with *cancer before IIM*, the most frequent subtypes were breast cancer (n=54, 17%), cervical cancer (n=39, 13%), skin cancer (n=39, 13%), colorectal cancer (n=32, 10%) and prostate cancer (n=30, 9.68%) while in patients with *first cancer after IIM*, the top five prevalent subtypes were skin cancer (n= 87, 28%), prostate cancer (n=29, 9%), breast cancer (n=28, 9%), colorectal cancer (n=26, 9%) and lung cancer (n=24, 7.84%).

Table 20. The cancer characteristics among patients with cancer before IIM and patients with first cancer after IIM

	Patients with cancer before IIM (n=310)	**Patients with first cancer after IIM (n=306)**
Cancer before IIM, n (%)	-	49 (16.01%)
First cancer after IIM, n (%)	49 (15.81%)	-
Second cancer and IIM, n (%)	11 (3.55%)	50 (16.34%)
Disease extent of cancer, n (%)		
Cancer in-situ	27 (8.71%)	52 (16.99%)
Locally aggressive	86 (27.74%)	110 (35.95%)
Locally or regionally aggressive	29 (9.35%)	37 (12.09%)
Distant metastasis	12 (3.87%)	26 (8.50%)
Missing	37 (11.94%)	58 (18.95%)
Missing due to diagnosed before 2004	119 (38.39%)	23 (7.52%)
Time from cancer/IIM to IIM/first cancer, n (%)		
≤ 1 year	89 (28.71%)	84 (27.45%)
(1-3] years	62 (20.00%)	60 (19.61%)
> 3 years	159 (51.29%)	162 (52.94%)
Time from cancer/IIM to IIM/first cancer, median (IQR)	3.20 (0.71-9.94)	3.49 (0.88-8.11)
Number of cancers diagnosed before/after IIM, n (%)		
1	252 (81.29%)	256 (83.66%)
2-3	56 (18.06%)	42 (13.73%)
> 3	< 5	8 (2.61%)

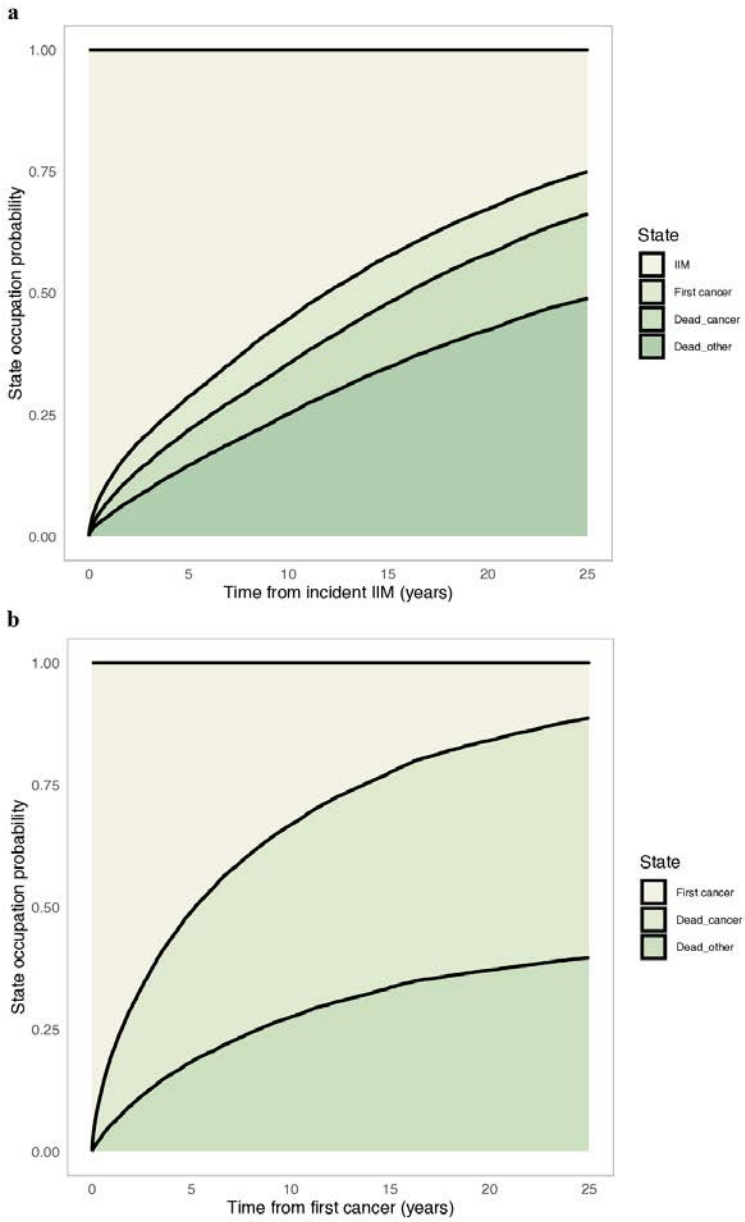
IIM: Idiopathic inflammatory myopathy; IQR: Interquartile range; Among patients with *cancer after IIM*, the characteristics were defined based on the cancer diagnosis that occurred closest in time to the IIM diagnosis.

Causes of death

Among patients with no cancer that occurred at or after IIM diagnosis, 93 died from cancer, and 364 died from other causes, specifically of 128 (28%) deaths due to diseases of the circulatory system, 52 deaths due to IIM and 46 deaths due to diseases of the respiratory system (**Figure 9a**). Notably, there were 20 patients who died from cancer but had no cancer diagnosis registered in the SCR. Among patients with *first cancer after IIM*, we observed 95 deaths from cancer and 62 deaths from other causes, mostly diseases of the circulatory system (n=23, 15%) and the respiratory system (n=13, 8%). Of 50 patients with *second cancer after* IIM, 27 died during follow-up (**Figure 9b**).

First cancer multistate model

The state occupation probability of each state at several follow-up time points predicted from the crude first cancer multistate model is presented in **Figure 10**. Compared to patients with no cancer diagnosis after IIM, patients with *first cancer after IIM* had a greater risk of death; at 25th year of follow-up, the risk of overall death was 66% for patients with no cancer diagnosis after IIM and 89% for patients with *first cancer after IIM*. Furthermore, patients with no cancer diagnosis after IIM were more likely to have died from other causes than from cancer and the risk of death from other causes increased gradually from 4% at 1st year to 35% at 10th year to 50% at 25th year of follow-up. Whereas patients with *first cancer after IIM* were at a greater risk of death from cancer than other causes; the risk of death from cancer was 14% at 1st year and it rapidly rose to 31% at 5th year and further increased to 49% at 25th year of follow-up. However, the risk of death from other causes among patients with *first cancer after IIM* also increased during follow-up and was comparable to that in patients with no cancer diagnosis after IIM. The state occupation probabilities estimated from the parametric models were comparable to the ones estimated by the AJ estimator.



Follow-up year	IIM state	From incident IIM state			From first cancer state		
		First cancer state	Death from cancer state	Death from other causes state	First cancer state	Death from cancer state	Death from other causes state
1	88.55%	4.10%	3.09%	4.26%	80.08%	14.29%	5.63%
5	71.28%	6.84%	7.28%	14.60%	50.93%	30.73%	18.34%
10	55.33%	9.16%	10.31%	25.20%	33.13%	39.37%	27.50%
15	42.39%	9.61%	13.28%	34.72%	22.36%	44.03%	33.61%
20	32.75%	9.25%	15.62%	42.38%	15.90%	47.01%	37.09%
25	25.03%	8.70%	17.37%	48.90%	11.32%	49.03%	39.65%

Figure 10. The state occupation probabilities estimated from the crude parametric first cancer multistate model, starting from "IIM" state (a) and "First cancer" state (b). *IIM: Idiopathic inflammatory myopathy.*

Table 21 presents the predicted length of stay in a non-fatal state at a given time point of follow-up. At one year after IIM diagnosis, patients spent almost the entire year in the "Incident IIM" state. The proportion of time spent in the "Incident IIM" state dropped gradually during follow-up as the risks of first cancer and death increased. Similarly, patients with *first cancer after IIM* were alive nearly the whole first year after cancer diagnosis but the proportion of duration spent in the "First cancer" state dropped sharply after five years of follow-up.

Table 21. The length of stay in year with 95% CI at 1st, 5th, 10th, 15th, 20th and 25th year of follow-up, estimated from the crude parametric first cancer multistate model, starting from the "Incident IIM" state and the "First cancer" state

	From incident IIM state		From first cancer state
Follow-up year	**IIM state**	**First cancer state**	**First cancer state**
1	0.93 (0.92-0.94)	0.03 (0.02-0.04)	0.88 (0.85-0.91)
5	4.09 (4.00-4.18)	0.26 (0.22-0.31)	3.42 (3.21-3.65)
10	7.25 (7.05-7.46)	0.67 (0.58-0.78)	5.48 (5.03-5.97)
15	9.68 (9.36-10.01)	1.15 (0.99-1.33)	6.85 (6.17-7.60)
20	11.55 (11.09-12.03)	1.62 (1.38-1.90)	7.79 (6.89-8.81)
25	12.97 (12.36-13.62)	2.06 (1.74-2.45)	8.46 (7.35-9.73)

IIM: Idiopathic inflammatory myopathy; CI: Confidence interval.

The effect estimates of the studied prognostic factors on each transition are presented in **Table 22**. Having IIM diagnosis at age > 60 years, being men and having DM were associated with *first cancer after IIM*. In patients without *first cancer after IIM*, having IIM at age > 60 years and DM conferred a higher risk of death from cancer; having IIM at age > 60 years and *cancer before IIM* increased the risk of death from other causes while having IIM diagnosis between 2009 and 2020 was a protective factor. In patients with *first cancer after IIM*, having *cancer before IIM*, DM and a cancer diagnosis within a year after IIM diagnosis were associated with a greater mortality from cancer while having *first cancer after IIM* between 2009 and 2020 was a protective factor; having *first cancer after IIM* at age > 70 years, being men and having a cancer diagnosis within a year after IIM diagnosis were risk factors for death from other causes.

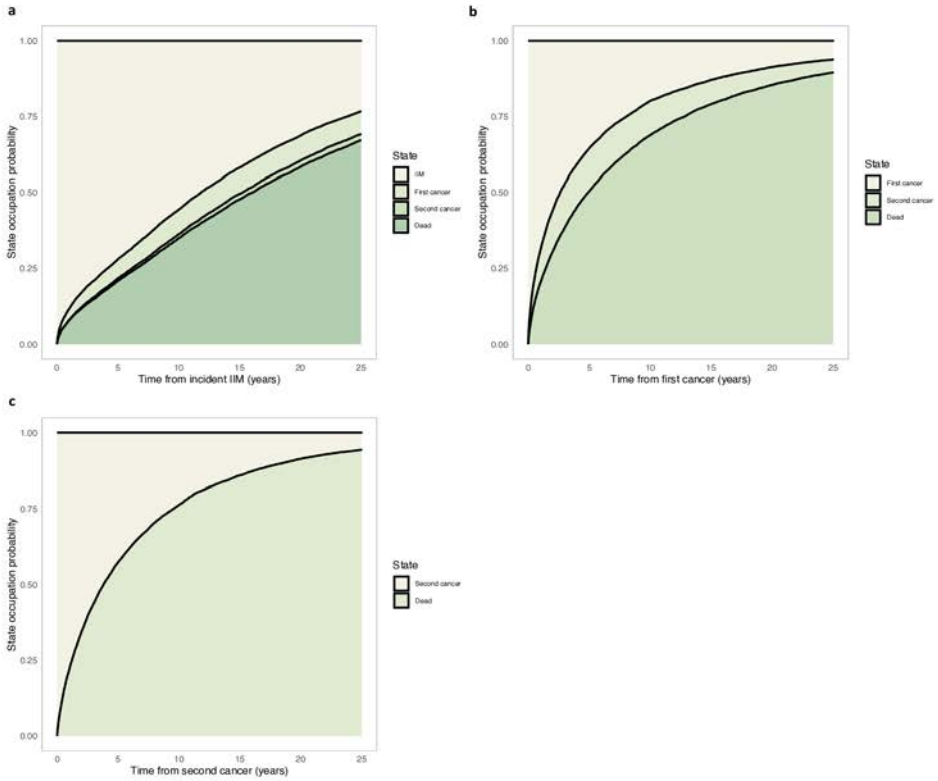
Table 22. The hazard ratios with 95% CI of the covariates in each transition of the first cancer multistate model

Transition	Covariate	Hazard ratio (95% CI)
T1: IIM to first cancer	Age at IIM diagnosis > 60 years	2.89 (2.25-3.72)
	Male sex	1.34 (1.07-1.68)
	Cancer before IIM	1.16 (0.85-1.60)
	DM	1.54 (1.22-1.94)
	Calendar year at IIM diagnosis 2009-2020	0.92 (0.72-1.17)
T2: IIM to death from cancer	Age at IIM diagnosis > 60 years	3.94 (2.43-6.39)
	Male sex	1.15 (0.76-1.74)
	DM	4.24 (2.77-6.48)
	Calendar year at IIM diagnosis 2009-2020	0.71 (0.47-1.08)
T3: IIM to death from other causes	Age at IIM diagnosis > 60 years	6.46 (4.89-8.54)
	Male sex	1.07 (0.87-1.32)
	Cancer before IIM	1.35 (1.03-1.77)
	DM	0.92 (0.73-1.17)
	Calendar year at IIM diagnosis 2009-2020	0.68 (0.54-0.85)
T4: First cancer after IIM to death from cancer	Age at first cancer diagnosis after IIM > 70 years	1.44 (0.93-2.21)
	Male sex	0.99 (0.66-1.50)
	Cancer before IIM	1.74 (1.07-2.85)
	DM	2.16 (1.42-3.28)
	Calendar year at first cancer diagnosis after IIM 2009-2020	0.50 (0.33-0.78)
	Time from IIM to first cancer < 1 year	2.29 (1.42-3.70)
	Time from IIM to first cancer (1-3] years	1.08 (0.60-1.97)
T5: First cancer after IIM to death from other causes	Age at first cancer diagnosis after IIM > 70 years	2.56 (1.47-4.46)
	Male sex	1.77 (1.03-3.04)
	Male sex_tvc_df1	2.52 (1.41-4.50)
	Cancer before IIM	0.89 (0.43-1.82)
	DM	0.87 (0.49-1.52)
	Calendar year at first cancer diagnosis after IIM 2009-2020	1.13 (0.62-2.08)
	Time from IIM to first cancer < 1 year	1.87 (1.01-3.46)
	Time from IIM to first cancer (1-3] years	1.42 (0.73-2.75)

IIM: Idiopathic inflammatory myopathy; CI: Confidence interval; T: Transition; tvc_df1: The corresponding estimate represents the change in hazard ratio of the time-dependent covariate for every unit increase in logarithm of time between time 0 and the first centile of time.

Second cancer multistate model

In patients with *first cancer after IIM*, the one-year risk of having *second cancer after IIM* was 11% and it increased to 14% at 5th year and gradually decreased to 4% at 25th year of follow-up (**Figure 11b**). The risk of death was much higher in patients with *first cancer after IIM* and patients with *second cancer after IIM* compared to patients without any cancer diagnosed at or after IIM diagnosis. Among patients with *second cancer after IIM*, the risk of death rose from 22% at 1st year, to 86% at 15th year to 94% at 25th year of follow-up. Regarding the length of stay, similar to patients with *first cancer after IIM*, patients with *second cancer after IIM* stayed alive for almost the entire first year after cancer diagnosis and the duration of survival reduced during the follow-up (**Table 23**).



	From incident IIM state				From first cancer state			From second cancer state	
Follow-up year	Incident IIM state	First cancer state	Second cancer state	Death state	First cancer state	Second cancer state	Death state	Second cancer state	Death state
1	88.45%	3.70%	0.15%	7.70%	69.06%	10.75%	20.19%	77.90%	22.10%
5	72.09%	6.35%	0.82%	20.74%	35.51%	14.10%	50.39%	42.66%	57.34%
10	55.67%	8.12%	1.32%	34.89%	19.87%	11.31%	68.82%	23.76%	76.24%
15	41.81%	8.73%	1.82%	47.64%	12.92%	7.97%	79.11%	14.05%	85.95%
20	31.33%	8.24%	2.08%	58.35%	8.69%	5.89%	85.42%	8.63%	91.37%
25	23.31%	7.56%	1.99%	67.14%	6.20%	4.29%	89.51%	5.64%	94.36%

Figure 11. The state occupation probabilities estimated from the crude parametric second cancer multistate model, starting from "Incident IIM" state (a), "First cancer" state (b) and "Second cancer" state (c). *IIM: Idiopathic inflammatory myopathy.*

Table 23. The length of stay in year with 95% CI at 1st, 5th, 10th, 15th, 20th and 25th year of follow-up, estimated from the crude parametric second cancer multistate model, starting from "Incident IIM" state, "First cancer" state and "Second cancer" state

	From incident IIM state			From first cancer state		From second cancer state
Follow-up year	IIM state	First cancer state	Second cancer state	First cancer state	Second cancer state	Second cancer state
1	0.93 (0.91-0.94)	0.03 (0.02-0.03)	0.00 (0.00-0.00)	0.85 (0.82-0.88)	0.03 (0.02-0.05)	0.87 (0.78-0.96)
5	4.11 (4.03-4.19)	0.24 (0.20-0.29)	0.02 (0.02-0.03)	3.06 (2.88-3.26)	0.37 (0.27-0.51)	3.17 (2.63-3.82)
10	7.30 (7.11-7.50)	0.61 (0.52-0.72)	0.08 (0.05-0.12)	4.70 (4.31-5.13)	0.82 (0.59-1.13)	4.77 (3.76-6.07)
15	9.72 (9.42-10.03)	1.04 (0.88-1.21)	0.16 (0.10-0.25)	5.72 (5.12-6.40)	1.17 (0.82-1.68)	5.69 (4.22-7.66)
20	11.54 (11.12-11.98)	1.46 (1.23-1.73)	0.26 (0.16-0.43)	6.42 (5.59-7.37)	1.44 (0.97-2.13)	6.24 (4.38-8.90)
25	12.89 (12.33-13.48)	1.86 (1.54-2.24)	0.36 (0.21-0.63)	6.90 (5.87-8.11)	1.65 (1.07-2.54)	6.59 (4.39-9.88)

IIM: Idiopathic inflammatory myopathy; CI: Confidence interval.

Table 24 shows the prognostic factors associated with *first cancer after IIM*, *second cancer after IIM* and death. The covariates associated with an increased risk of *first cancer after IIM* were the same as in the first cancer multistate model. Furthermore, the factors associated with an increased overall mortality before the diagnosis of *second cancer after IIM* were in line with those observed in the first cancer multistate model. Being men was the only risk factor significantly associated with *second cancer after IIM*. In patients with *second cancer after IIM*, DM was associated with a four-fold higher risk of death compared to other IIM.

Table 24. The hazard ratios with 95% CI of the covariates in each transition of the second cancer multistate model

Transition	Covariate	Hazard ratio (95% CI)
T1: IIM to first cancer	Age at IIM diagnosis > 60 years	2.89 (2.25-3.72)
	Male sex	1.34 (1.07-1.68)
	Cancer before IIM	1.16 (0.85-1.60)
	DM	1.54 (1.22-1.94)
	Calendar year at IIM diagnosis 2009-2020	0.92 (0.72-1.17)
T2: IIM to death	Age at IIM diagnosis > 60 years	4.83 (3.67-6.38)
	Age at IIM diagnosis > 60 years_tvc_df1	0.99 (0.78-1.26)
	Age at IIM diagnosis > 60 years_tvc_df2	0.83 (0.69-1.00)
	Male sex	1.14 (0.94-1.37)
	Cancer before IIM	2.46 (1.98-3.07)
	Cancer before IIM_tvc_df1	1.15 (0.95-1.38)
	Cancer before IIM_tvc_df2	1.43 (1.21-1.68)
	DM	1.57 (1.27-1.95)
	DM_tvc_df1	0.58 (0.49-0.69)
	DM_tvc_df2	1.20 (1.06-1.36)
	Calendar year at IIM diagnosis 2009-2020	0.67 (0.55-0.83)
	Calendar year at IIM diagnosis 2009-2020_tvc_df1	1.07 (0.91-1.27)
	Calendar year at IIM diagnosis 2009-2020_tvc_df2	0.84 (0.73-0.96)
T3: First cancer to second cancer after IIM	Age at first cancer diagnosis after IIM > 70 years	1.48 (0.81-2.73)
	Male sex	1.88 (1.06-3.34)
	Cancer before IIM	1.98 (0.97-4.02)
	DM	0.53 (0.27-1.07)
	Calendar year at first cancer diagnosis after IIM 2009-2020	0.97 (0.48-1.95)
	Time from IIM to first cancer < 1 year	0.78 (0.36-1.66)
	Time from IIM to first cancer (1-3] years	1.04 (0.48-2.28)
T4: First cancer after IIM to death	Age at first cancer diagnosis after IIM > 70 years	1.61 (1.11-2.32)
	Male sex	1.11 (0.78-1.58)
	Cancer before IIM	1.30 (0.82-2.05)
	DM	1.54 (1.08-2.21)
	Calendar year at first cancer diagnosis after IIM 2009-2020	0.61 (0.42-0.89)
	Time from IIM to first cancer < 1 year	2.34 (1.55-3.52)
	Time from IIM to first cancer (1-3] years	1.34 (0.81-2.21)
T5: Second cancer after IIM to death	Age at first cancer diagnosis after IIM > 70 years	2.67 (0.84-8.47)
	Male sex	1.67 (0.70-3.97)
	Cancer before IIM	1.55 (0.62-3.87)
	DM	3.77 (1.46-9.75)
	Calendar year at first cancer diagnosis after IIM 2009-2020	0.97 (0.30-3.14)
	Time from first cancer to second cancer after IIM (1-3] years	0.97 (0.34-2.79)
	Time from first cancer to second cancer after IIM < 1 year	1.05 (0.34-3.22)

IIM: Idiopathic inflammatory myopathy; CI: Confidence interval; T: Transition; tvc_df1: The corresponding estimate represents the change in hazard ratio of the time-dependent covariate for every unit increase in logarithm of time between time 0 and the first centile of time; tvc_df2: The corresponding estimate represents the change in hazard ratio of the time-dependent covariate for every unit increase in logarithm of time between the first centile and the second centile of time.

5 Discussion

In this section, I aim to comprehensively discuss the major findings of this thesis by comparing with previous literature, highlighting strengths, limitations, scientific and clinical implications, as well as methodological considerations.

5.1 Understanding the familial and genetic susceptibility of IIM with population data

The nationwide registration system integrating population and healthcare data are powerful data sources that can overcome many challenges presented in traditional family study including small and highly selective sample, laborious data collection, recall bias, short duration of follow-up and high cost (360). In this thesis, **Study I-III** with a population-based case-control family design included a large representative sample of patients with IIM. Furthermore, random sampling from the general population, identification of their blood relatives, accurate determination of the presence of a wide range of diseases in patients, comparators and their relatives spanning a long time period were easily done in a cost-effective way. With appropriately performed statistical analyses, findings from **Study I-III** provided important insights into the genetic susceptibility of IIM that we as researchers have made huge amount of efforts to answer in the past decade. The significance of this was exemplified in **Study I**, confirming the familial risk of IIM and pioneering the report on family-based heritability of IIM. This sheds light on the potential upper limit of the proportion of the phenotypic variance of IIM that could be explained by additive genetic variance.

The familial risk of IIM was estimated to be four times higher in patients with IIM compared to the general population or three times higher in first-degree relative pairs of patients with IIM compared to those of the general population. Our findings are line with two previous population-based family studies; one used Danish nationwide register data and reported an familial OR of having parents or siblings affected by IIM in 949 patients with IIM to be 3.9 (95%CI 0.6-27.7); another Swedish study used similar approach and presented a significant familial risk of IIM in parents or full-siblings of 2,668 patients with IIM (SIR=4.0 95%CI 1.3-8.4) (61, 62). Both of these two studies defined IIM based on one outpatient or inpatient visit and analysed familial aggregation of IIM based on very small number of cases, which led to the wide confidence intervals. **Study I**, where IIM was defined using a robust algorithm, presented a more precise estimate of familial aggregation of IIM through analysing all first-degree relative pairs.

The family-based heritability of IIM was estimated to be 22% (95%CI 12%-31%), which is higher than the previously reported SNP-based heritability (estimated using SNPs presented on the ImmunoChip, 5.5% for DM and 8.3% for PM) (65). This discrepancy could be explained by an overestimation in our family-based heritability of IIM due to the presence of shared environmental factors or an underestimation in the SNP-based heritability due to selective genomic coverage (missing heritability) . The former reason is possible since it is unlikely that similarity between first-degree relatives solely stems from genetics, particularly in the case of full siblings who often share the same upbringing environment. However, in

Study I, we found that the heritability of IIM among full siblings was just two percent higher than the heritability of IIM among any first-degree relatives, suggesting that our family-based heritability of IIM was unlikely biased upward to a large extent due to the inclusion of shared environmental factors in the estimation. Missing heritability was a more convincing explanation as the used ImmunoChip covered only 186 immune-related loci. Moreover, the study was likely underpowered to detect SNPs with small to moderate effects and the analytical method, genome-wide complex trait analysis, could underestimate the heritability of a rare disease (66, 67). Considering the family-based heritability as the upper limit and the SNP-based heritability as the lower limit for the heritability of IIM, the missing heritability was roughly 12%. This suggests that there might be numerous more IIM-associated variants yet to be identified. Advancement in genotyping, imputation techniques and analytical methods, as well as international collaboration have led to discovery of many non-HLA variants linked to IIM (with GWAS or suggestive level of significance) (**Table 2**) (3, 5, 6, 72, 75, 76). Nearly all of the identified genetic variants are common SNPs. We know little about if there are other types of genetic variants associated with IIM. Our team is currently conducting an ongoing study involving the genotyping of patients with clearly defined IIM using whole-genome sequencing, which may shed light on this question.

The findings from **Study I** have important clinical implications. Firstly, family history of IIM (among first-degree relatives) could be considered an indication pointing towards a diagnosis of IIM although familial IIM is exceptionally rare. Once family history of IIM has been observed, at least in the Swedish setting where a diagnostic code of IIM is strictly given to only confirmed cases, it could serve as supporting evidence for a genuine IIM case.

When advising patients about family planning, it is crucial to interpret the findings from **Study I** cautiously. Patients need to understand that although the familial risk of IIM appeared significant in the study, this was based on a comparison of small proportions. The observed frequency of familial IIM in **Study I** was extremely low. Additionally, when we refer to the heritability of IIM, we are discussing the proportion of the variance in IIM that can be explained by additive genetic variance. It is important to note that this is about variation at the population level and doesn't determine the likelihood of a child inheriting IIM from an affected parent. Moreover, heritability is specific to the study population, the time of observation and the environment. Therefore, the generalizability of the family-based heritability of IIM reported in **Study I** to other populations with distinct settings is limited (361).

In **study II**, we confirmed the familial associations with other RIDs and AITD in patients with IIM, as previously suggested in studies (62, 303-305), and also reported statistically significant familial associations with CeD and IBD for the first time (309). However, some previous studies examining familial associations with IIM in patients with SSc, SS and IBD reported no significant findings, probably due to lack of statistical power (306-310). It is because a sufficiently large sample of patients with a specific autoimmune disease is required to detect familial association with a rare disease like IIM. For example, a Swedish population-based study examining the familial association with IIM in 25,846 patients with ulcerative colitis and 18,885 patients with Crohn's disease observed a SIR of 1.0 (95%CI 0.1-4.0) based on three cases in siblings for ulcerative colitis and a SIR of 2.3 (95%CI 0.6-7.2) based on six cases in siblings for Crohn's disease (309). **Study II** investigated the familial autoimmunity

in a large representative sample of IIM might avoid overlooking potential associations. Furthermore, our analyses were performed under a causal framework and found the observed familial associations were robust to the presence of causal relationships between IIM and the autoimmune diseases, supporting that the observed familial associations could be attributed to shared familial factors.

When two diseases are found to run in families, it suggests that they might share genetic susceptibility. This insight can guide genetic studies, helping us better understand the genetic contribution to a disease. Several genetic studies of IIM were guided by the knowledge of common genetic architecture among autoimmune diseases, particularly RIDs, and successfully identified numerous non-HLA loci associated with IIM via using the ImmunoChip or meta-analysing GWAS summary statistics of seropositive RIDs (3, 5, 75, 76). In **Study II**, the strength of familial associations of IIM with CeD and IBD were comparable to that of familial association between IIM and other RIDs. Among the reported IIM-associated genetic variants, five were associated with CeD and 15 were associated with IBD, respectively, compared to 27 for SSc and 26 for SLE (4-6, 69-71, 75-77, 80, 83, 85, 87, 95-97, 100, 101, 104, 105). Considering the previously successful examples and the access to publicly available GWAS summary statistics of CeD and IBD, secondary analysis of their and IIM's GWAS summary statistics may lead to identification of novel disease-associated genetic variants and enhance understanding of the shared pathogenesis between IIM and CeD and IBD (343).

In diagnostic assessment, family history of autoimmune diseases is one of the indicators towards a diagnosis of IIM. The findings of **Study II** add to the knowledge of which autoimmune diseases cluster in families of patients with IIM (53). There should be additional attention to suspected cases of IIM whose first-degree relatives are affected by other RIDs, IBD, AITD and CeD.

5.2 IIM and cancer: from genetics to consequences

Investigating the genetic contribution to the co-occurrence of IIM and cancer is one of the primary aims of this thesis, given the relevant knowledge is lacking. The findings from **Study III** (familial co-aggregation of IIM and cancer) and **IV** (genetic correlation between IIM and B-cell lymphomas) help to fill the knowledge gap and suggest that IIM to a low extent shares genetic susceptibility with cancer.

In **Study III**, when studying cancer overall, we observed statistically significant familial co-aggregation of cancer only in male relative pairs of patients with DM, suggesting that shared familial factors might be of more importance to explain the co-occurrence of DM and cancer in men than other subgroup of patients affected by IIM. Interestingly, both male sex and DM are well-known risk factors of cancer in IIM (206). To date, no sex difference in SNP associations with IIM has been noted (3, 4). However, an association between IIM and X chromosome abnormality (47, XXY) in men has been repored (107). Besides genetics, environmental exposure or lifestyle behaviours shared between male relatives might also explain this observation. For example, male relatives may exhibit similar smoking habits, a known factor associated with both IIM and cancer (113).

We also examined the familial co-aggregation of IIM and a variety of specific cancer types in **Study III** and observed significant findings only for myeloid malignancies and liver cancer in male relative pairs. Only familial associations between IIM and haematological malignancies including HL, MM, DLBCL, FL and CLL have been examined in previous studies but none of these association reach statistical significance (246-251). In **Study III**, the aORs for familial associations with HL and B-cell lymphomas were 1.1 (95%CI 0.6-1.8) and 1.2 (95%CI 1.0-2.5), respectively. These point estimates align with findings from previous studies (249, 251).

In **Study III**, we explored if first-degree relatives of patients with IIM had an increased risk of cancer before 50 years of age since early-onset cancer might be additional evidence supporting the presence of shared genetic susceptibility between IIM and cancer (362). We found that only offspring of patients with IIM were more likely to have early-onset cancer compared to offspring of the general population without IIM. Besides shared genetics, we could not preclude that this observation could be due to cancer surveillance in offspring of patients with IIM. However, this explanation is not supported by the clinical practice. To date, there is no advice on cancer screening among offspring of patients with IIM (363). Nevertheless, considering many autoimmune diseases are associated with an increased risk of cancer, IIM in an affected parent might indirectly lead to an increased cancer surveillance in their offspring via familial association with other autoimmune diseases, especially when the offspring was diagnosed with an autoimmune disease younger than 50 years of age.

Importantly, **Study III** is the first population-based family study exploring the familial co-aggregation of IIM and cancer, and the findings were found to be robust. However, replication studies in other populations are warranted to make conclusive interpretation on the role of familial factors in the pathological link between IIM and cancer.

In **Study IV**, our analyses focused on examining the genetic correlation, specifically exploring loci genetic correlation and jointly associated SNPs, between DM and PM with four common B-cell lymphoma subtypes including DLBCL, FL, CLL and MZL. Despite not finding a significant familial association between IIM and B-cell lymphomas in **Study III**, we, in **Study IV,** uncovered genetic correlations in a limited number of loci and all disease pairs showed genetic correlations and/or had jointly associated SNPs in the HLA region. These findings highlight the significance of the HLA region as a central factor in the limited shared genetic susceptibility observed between major IIM and B-cell lymphoma subtypes. This coincides with the recognized importance of HLA alleles in the genetic susceptibility of both IIM and common B-cell lymphoma subtypes. Specifically, for DM, PM and the four B-cell lymphoma subtypes, associations with various HLA class I-III alleles have been reported (4-7, 69-73, 327). Moreover, associations with HLA class II alleles have been identified in CLL and FL and associations with HLA class I alleles have been reported in DLBCL and MZL (86, 88-91, 93). In our study, we observed corresponding shared genetic susceptibility in HLA class II loci for the disease pairs of DM and PM with FL and CLL, as well as in HLA class I loci for the disease pairs of DM and PM with DLBCL and MZL. Though both DM and PM show associations with HLA class III alleles, we only found significant genetic correlations in HLA class III loci for the DM disease pairs.

Both DM and PM exhibit associations with haematological malignancies in comparable manners concerning the strength of association and the timing of cancer diagnosis (8, 9, 364). However, it is unclear if the underlying biological mechanisms between DM and PM overlap. In **Study IV**, we observed both similarities and differences in the shared genetic susceptibility between the DM and PM disease pairs. For example, there were two HLA class I loci associated with both the disease pairs DM-DLBCL and PM-DLBCL. For the disease pair PM-DLBCL, we also found shared genetic susceptibility in non-HLA loci. These findings suggest the presence of both common and distinct biological pathways in the development of B-cell lymphomas between DM and PM.

The findings from **Study III** and **IV** suggest that familial factors and genetics are not a major factor contributing to the co-occurrence of IIM and cancer overall, especially regarding B-cell lymphomas specifically. This is expected given the complex relationship between IIM and cancer. It is unlikely that there is a factor that could explain the pathological link between IIM and cancer substantially. The biological mechanism behind the co-occurrence of IIM and cancer may involve numerous genetic and environment factors, as well as their interactions, with small to moderate effects. Moreover, the biological mechanism may differ among disease subsets with different characteristics. For instance, findings from patients with DM and anti-TIF1γ autoantibodies suggest that DM may develop due to the interplay between anti-tumour immunity and tumorigenesis (227, 232). Conversely, for other IIM subtypes with a long-term risk of cancer, IIM-related factors including chronic inflammation and immunosuppressive treatments may contribute to cancer development, particularly skin cancer (9, 237, 239-241).

This thesis does not only comprehend our knowledge of the pathological link between IIM and cancer, but also offers useful information to guide future research. In **Study IV**, we detected shared genetic susceptibility in non-HLA loci for the disease pairs DM-FL, DM-CLL, PM-DLBCL, PM-CLL and PM-MZL. No genetic variant association in these loci has been reported to be associated with IIM and B-cell lymphomas. Concerning missing heritability has been suggested for IIM and B-cell lymphomas (65, 365, 366), further investigation in these genomic regions may lead to identification of novel disease-associated genetic variants.

The focus of our discussion is now moved to the findings of **Study V** where we investigated cancer-related disease burden by following patients' trajectories to cancer and death since IIM diagnosis.

In **Study V**, we revealed an increased risk in mortality and a change in disease burden, wherein the primary cause of death shifted from other causes to cancer in patients with *first cancer after IIM*, compared to patients without any subsequent cancer diagnosis after IIM. The predicted one- and five-year overall survival rates were 80% and 51% in patients with *first cancer after IIM*, respectively, versus 93% and 78% in patients without any cancer diagnosis after IIM. Moreover, compared to patients with no cancer diagnosis after IIM, patients with *first cancer after IIM* had a much higher risk of death from cancer and a comparable risk of death from other causes (the five-year risk of death: 31% versus 7% for cancer and 18% versus 15% for other causes). It is also important to note that the jump in risk of death from cancer in patients with *first cancer after IIM* was more profound within

the first five years of the diagnosis of *first cancer after IIM*. These findings are line with previous studies where patients with CAM had lower five-year survival rate than patients with primary IIM only and cancer as the most common cause of death (193, 196, 198, 289-293).

Furthermore, we made the first attempt to estimate the risk of having a second primary cancer after IIM diagnosis in patients with IIM and to examine its impact on mortality. We found that the risk of having a second primary cancer in patients with *first cancer after IIM* was around 10% within the first 10 years of follow-up and it gradually decreased to below 5% at 25th year of follow-up. Moreover, a diagnosis of *second cancer after IIM* slightly increased the risk of death in patients with *first cancer after IIM*.

Knowing what patient characteristics are associated with an increased cancer risk and an elevated mortality overall or from a specific cause of death is important to assess disease burden and direct intervention strategy. First, we found that old age at IIM diagnosis (> 60 years), being men and having DM were associated with *first cancer after IIM*, in line with previous findings (206). We also found that being men was significantly associated with an increased risk of a second cancer in patients with *first cancer after IIM*. In the latest cancer screening guideline for IIM, risk of a second cancer is not accounted for (363). Although having more than one cancer diagnosis in patients with IIM is uncommon, it may become a more frequently observed complication given that the survival of patients with IIM has been improved in recent years (149-151). Attention to second cancer may be needed in male patients with a cancer diagnosis after IIM.

Among patients with no cancer diagnosis after IIM, having IIM at age older than 60 years was associated with increased risks of death from cancer and other causes but it was a stronger risk factor for death from other causes. We found that having *cancer before IIM* was a risk factor of death from other causes while having IIM diagnosis between 2009 and 2020 was a protective factor, probably due to improved disease management in recent years (149-151). Among patients with *first cancer after IIM*, having *cancer before IIM*, DM, first cancer diagnosis within one year after IIM diagnosis conferred a higher risk of death from cancer while having first cancer diagnosis between 2009 and 2020 was a protective factor. For increased mortality from other causes, the risk factors were having first cancer diagnosis at age older than 70 years, being men and having first cancer diagnosis within one year after IIM diagnosis. In patients with *second cancer after IIM*, having DM was associated with an increased mortality. These findings corroborate with previous studies where a close temporal relationship between IIM and cancer and male sex were suggested to be associated with an increased mortality in patients affected by IIM and cancer (177, 188, 196, 210, 289).

5.3 Other considerations

Limitations associated with the Swedish nationwide register data

The Swedish nationwide register data is not perfect and come with limitations. First, its use on population scale often comprises detailed information. Using the NPR data, only ICD-10 codes could be used for IIM classification. As a result, patients with IIM in **Study I-III** and **V** were broadly classified into DM and other IIM. The other IIM group was heterogeneous and could include ASyS, IMNM, IBM, OM and PM as these subtypes share ICD-10 codes

M33.2 and M33.9. This prevented us from performing analyses in more homogeneous group of IIM subtype. Specifically, in **Study III**, the estimates of familial co-aggregation of other IIM and cancer might be diluted due to inclusion of patients with ASyS, IBM and OM in the analysis, subtypes that has little evidence supporting an association with cancer (44, 45, 176, 188). Another problem associated with the NPR is that it does not include data from primary care visit. This will become problematic when studying a disease that is usually managed at primary care. For instance, CeD is mostly detected and treated at primary care and only patients with severe symptoms require inpatient or hospital-based outpatient care (367). Therefore, in **Study II**, the detected familial association between IIM and CeD might be overestimated. It is also essential to take into account the timeframe encompassed by a nationwide register, especially when studying early-onset disease in a family study including parents of patients with a late-onset disease. For example, in **Study II**, T1DM in parents and full siblings might be missed as there was no nationwide coverage for inpatient visit till 1997 and outpatient visit till 2001. A sensitivity analysis including only first-degree relatives alive in 2001 was therefore performed to examine if this might affect the estimation.

Although the coverage of the SCR is nearly complete, underreporting may be more of a problem for patients with aggressive cancer, old age and comorbid condition (319). When ascertaining prevalent cancer cases, data from the CDR can be used to identify missed cases, as such in **Study III**. However, for incident cancer case, data from the CDR are less useful since it holds no information on date of cancer diagnosis. In **Study V**, there were 20 potential unreported cases of cancer. A validation study found that only 50% of cancer cases registered only in the NPR should be reported to the SCR (319). Therefore, in the main analysis, we might misclassify 10 patients with cancer as patients with primary IIM only, which was unlikely to have a significant impact on the estimation. A sensitivity analysis excluding the 20 patients found no significant difference compared to the main findings. Another concern is that not all variables included in the SCR are available since 1958. The SNOMED codes, with higher accuracy and specificity for cancer classification than ICD codes, were not introduced till 1993. In **Study III**, where we analyzed prevalent cancer cases diagnosed between 1958 and 2016, we classified cancer cases into broader categories to reconcile the varying levels of detail between the SNOMED and ICD codes. Consequently, further investigations into familial associations with specific B-cell lymphoma subtypes such as DLBCL and FL in patients with IIM were unfeasible. Similarly, **Study V** encountered a comparable challenge. As TNM and FIGO codes for cancer staging have become available only since 2004, and these systems often have missing values, only 219 (72%) patients with a first cancer diagnosis following IIM had information on TNM and/or FIGO codes. Consequently, with this reduced sample size, it was not feasible to explore the impact of cancer stage on mortality in the multivariable analyses.

As a researcher actively engaged in register-based research within Sweden, it is essential to acknowledge both the strengths and limitations inherent in the register data. Moreover, equipping oneself with knowledge and skills to mitigate potential biases is fundamental in conducting high-quality studies using population-based register data.

Misclassifying inherited IIM mimics as IIM, a big issue within the Swedish context?

Misdiagnosing inherited muscular diseases as PM is not uncommon given the shared clinical and histopathological presentations. However, as mentioned in section 1.3.1, the diagnostic work-up of IIM within the Swedish healthcare system is comprehensive and with a high awareness to potential IIM mimics. IIM-related ICD codes are strictly given based on evidence from all examination tests. Even if a misdiagnosis happens, it should be corrected within a short timeframe given that patients' clinical courses are closely monitored at specific hospital-based clinics.

In **Study I**, we performed additional investigations into the risk of misdiagnosing inherited IIM mimics as IIM. By checking the inpatient and outpatient visits of patients with IIM and their first-degree relatives with IIM between 1987 and 2017, we found that none of them had a visit indicating muscular dystrophies or metabolic myopathies. Given that DM present distinct skin manifestations, which set it apart from inherited IIM mimics, we performed a sensitivity analysis that only included patients with DM or JDM and their matched comparators. This yielded an aOR of 1.4 (95% 0.4-5.2) and a heritability of 5% (95% -8%-19%). Although the estimates became statistically insignificant, they still suggested the familial aggregation and heritability of IIM in the same direction as the main analysis. These findings bolster the confidence in the robustness of our **Study I** findings and suggest a low risk of misdiagnosing inherited IIM mimics as IIM within the Swedish context. However, since inherited muscular diseases stem from entirely different biological mechanisms than IIM, the risk of misdiagnosing them as IIM is a critical concern in the realm of IIM research, particularly genetic studies.

6 Conclusions

This thesis addresses a wide scope of research questions spanning from genetic inquiries of IIM to an investigation of cancer-related disease burden on patients with IIM. The key findings, scientific and clinical implications of **Study I** to **V** are summarized as below.

Key findings

- Compared to the Swedish general population, patients with IIM had a four-fold higher risk of having at least one first-degree relative affected by IIM. In Sweden, about 22% of the phenotypic variance of IIM could be explained by additive genetic variance.
- The autoimmune diseases aggregated within families of patients with IIM included other RIDs, IBD, AITD and CeD.
- We observed a modest familial co-aggregation of IIM and cancer among male first-degree relatives of patients with DM. The odds of having a cancer diagnosis before 50 years of age in offspring of patients with IIM was 14% higher than in those of comparators without IIM. Among a variety of specific cancer types, family history of myeloid malignancies and liver cancer in male first-degree relatives were significantly associated with IIM.
- Major IIM (DM and PM) and B-cell lymphoma subtypes (DLBCL, FL, CLL and MZL) shared genetic suscepitiblity in a limited number of loci, primarily within the HLA region. Some HLA loci were associated with more than one disease pair. The non-HLA loci showing significant genetic correlations were disease pair specific.
- Compared to patients with no cancer diangosis after IIM, patients with cancer diagnosis after IIM had an increased mortality, notably exhibiting a significantly greater risk of death from cancer and a comparable risk of death from other causes. The risk of having a second primary cancer was approximately 10% within the first decade after a cancer diagnosis following IIM. We also reported prognistic factors associated with increased risks of cancer and death events that happened after IIM.

Scientific implications

- Our family-based heritability of IIM surpassed the SNP-based heritability by roughly 12%, suggesting the presence of missing heritability and yet-to-be-discovered IIM-associated genetic variants.
- IIM to a low extent shared familial and genetic susceptibility with cancer. Moreover, our findings suggest that the HLA region was a key player in the limited shared genetic susceptibility between IIM and common B-cell lymphoma subtypes.
- The identified loci presenting significant genetic correlations with the disease pairs of DM and PM with B-cell lymphoma subtypes could serve as potential candidate genomic regions for uncovering of novel disease-associated genetic variants.

Clinical implications

- The actual prevalence of familial IIM was extremely low (0.8%). Moreover, heritability estimate tells nothing about the chance of having a child inherited a

disease from an affected parent. Rather it is an estimate to quantify how much additive genetics may contribute to the phenotypic variance of a disease in a specific population.

- Family history of IIM, other RIDs, IBD, AITD and CeD could serve as indicators pointing towards a IIM diagnosis.
- **Study V** sheds light on the cancer-related disease burden on patients with IIM and identified key prognotic factors. Patients who are diagnosed with cancer after IIM, particularly those exhibiting risk factors associated with increased risks of second cancer and mortality, require additional care.

7 Acknowledgements

I am grateful to have met many kind and talented individuals during my four-year PhD journey at Karolinska Institutet. I would have been able to accomplish this thesis without help and support from the following people and stakeholders.

Marie Holmqvist, my guiding main supervisor, I vividly recall our initial meeting at KEP. Your warm reception introduced me to the intricate realm of IIM research. Your steadfast belief in my abilities and the knowledge you have imparted me in the past six years together have been invaluable. Thank you for everything.

Ingrid Lundberg, my inspiring co-supervisor, you serve as a role model to many researchers in the field of rheumatology, and I am fortunate to count myself among them! Your dedication has created an incredibly positive atmosphere within the myositis team, and your invaluable insights and guidance have significantly enhanced the quality of this thesis. Thank you!

Helga Westerlind, my supportive co-supervisor, your wealth of experience in family and genetic studies has been a guiding light. Moreover, your encouragement and the teaching opportunities you provided have significantly contributed to my academic growth. Thank you!

Karin Hellgren, my caring co-supervisor, your sincere commitment to scientific research has been truly inspiring. I am also deeply grateful for your generous support and guidance in preparing for my thesis defence.

I want to extend my immensely gratitude to all co-authors involved in this thesis: **Fredrik Baecklund**, your expertise in oncology has been instrumental in shaping this thesis; **Ralf Kuja-Halkola**, your guidance in family study and statistical analysis has significantly contributed to the accomplishments of this thesis; **Karin Ekström Smedby**, your generous assistance in applying the GWAS summary statistics and your expertise in haematology have significantly contributed to the success of Study IV; **Anton Öberg Sysojev**, your insightful advice on the methodologies for estimating genetic correlation and your patient guidance in programming were instrumental in helping me navigate the challenges encountered during the execution of Study IV; **Karina Patasova**, your generous guidance in data wrangling and estimation of genetic correlation have been immensely valuable; **Hannah Brower**, your guidance was a beacon of light when I was navigating the complexities of multistate modelling and your advice has been instrumental in advancing Study V; **Catherine Zhu**, collaborating with you has been a joy and I am grateful for your insights and shared experiences in genetic studies.

Yu-Tao Xiang, my compassionate mentor, I am deeply grateful for your invaluable advice on on career planning. Your qualities as an exceptional teacher are truly inspiring.

I extend my heartfelt appreciation to **Martin Neovius**, the chair of my thesis defence. Your participation and contributions have been invaluable.

My colleagues in the Holmqvist team including **Karin Gunnarsson**, my study buddy, as well as **Karin Wadström** and **Majd Bairkdar**, have my deeply gratitude for engaging in insightful discussions and contributing to our collective learning and progress.

My valued colleagues within the Askling team and at KEP, including **Johan Askling**, **Anna-Maria Wintler**, **Arda Yilal**, **Bénédicte Delcoigne**, **Björn Pasternak, Daniela Di Guseppe**, **Diego Hernan Giunta**, **Elizabeth Arkema**, **Gustaf Bruze**, **Renata Zelic**, **Kelsi Smith**, **Kristin Waldenlind**, **Laura Pazzagli**, **Marina Dehara**, **Matilda Morin**, **Mehdi Osooli**, **Peter Apling**, **Simon Englund**, **Simon Steiger**, **Sandra Eloranta**, **Thomas Frisell**, **Joshua Entrop, Viktor Molander**, **Viktor Wintzell** and **Viet Ngoc Nguyen**, have my deepest appreciation for fostering an enriching learning environment and providing invaluable feedback that greatly contributed to my growth and development.

My colleagues in the Myositis team including **Angeles Shunashy Galindo-Feria**, **Alexandra Argyriou**, **Antonella Notarnicola**, **Anneli Dihkan**, **Begum Horuluoglu**, **Deniz Demirdal**, **Eva Lindroos**, **Eveline Van Gompel**, **Fabricio Espinosa**, **Gloria Rostvall**, **Helene Alexanderson**, **Helene Sandlund**, **Hideaki Tsuji**, **Irene Peralta**, **Ingrid Gerhardson**, **Karin Lodin**, **Karina Gheorghe**, **Kristofer Andreasson**, **Lara Dani**, **Louise Ekholm**, **Maryam Dastmalchi**, **Maho Nakazawa**, **Valérie Leclair** and **Yuebei Luo**: Being part of this inspiring and supportive team has been an honour and a joy. Your collective contributions have made my PhD journey immensely fulfilling. A heartfelt thank you to **Anneli** for taking time to review the Swedish popular science summary.

My participation in the National Clinical Research School in Chronic Inflammatory Diseases and the Swedish Interdisciplinary Graduate School in register-based research have significantly contributed to my academic and personal growth. I extend my heartfelt appreciation to **Helena Haris**, **Cecilla Aulin**, **Anita Berglund**, **Karin Modig** and **Johanna Bergman**, along with all the lecturers and fellow talented PhD students, for their collective efforts in making these research schools immensely successful.

I extend my sincere gratitude to the IMACS and InterLymph for fostering international networks that facilitate collaborations among researchers worldwide. I wish to express my heartfelt appreciation to **Janine A. Lamb**, **Christopher Amos**, **Lisa Rider**, **Sonja Berndt**, **Arial Harrison**, **Dakai Zhu**, **Dennis Robinson**, **Rebecca Montalvan**, and **Karen Pettit** for their generous assistance in data application, data transfer, and peer review.

Michael Fored, **Helena Nord**, **Pernilla Appelquist**, **Monica Ringheim** and the exceptional **KEP IT** team, thank you for unwavering support throughout my PhD journey.

The thesis received support and funding from the Swedish Research Council, the Wallenberg Foundation (Marianne and Marcus Wallenberg Clinical Fellow), the Swedish Rheumatism Association, the King Gustaf V:s 80 years Foundation, and the Professor Nanna Svartz Foundation and Region Stockholm (ALF-project). Your invaluable support has enabled us to contribute significantly to the scientific development in the field of IIM research.

我由細玩到大的好朋友，阿珊，翠珊和恩祺，感謝你們的陪伴和幫助。希望我回澳門之後可以多聚多見面。

感謝家人一直以來對我的支持和鼓勵。
特別感謝我親愛的媽咪，永遠關心我是否健康快樂，你無條件的愛是我最大的依靠。

8 References

1. Wagner E. Fall einer seltenen Muskelkrankheit. Arch Heilkd. 1863;4:282.
2. Lundberg IE, Fujimoto M, Vencovsky J, Aggarwal R, Holmqvist M, Christopher-Stine L, et al. Idiopathic inflammatory myopathies. Nat Rev Dis Primers. 2021;7(1):86.
3. Miller FW, Cooper RG, Vencovsky J, Rider LG, Danko K, Wedderburn LR, et al. Genome-wide association study of dermatomyositis reveals genetic overlap with other autoimmune disorders. Arthritis Rheum. 2013;65(12):3239-47.
4. Miller FW, Chen W, O'Hanlon TP, Cooper RG, Vencovsky J, Rider LG, et al. Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. Genes Immun. 2015;16(7):470-80.
5. Rothwell S, Cooper RG, Lundberg IE, Miller FW, Gregersen PK, Bowes J, et al. Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups. Ann Rheum Dis. 2016;75(8):1558-66.
6. Rothwell S, Chinoy H, Lamb JA, Miller FW, Rider LG, Wedderburn LR, et al. Focused HLA analysis in Caucasians with myositis identifies significant associations with autoantibody subgroups. Ann Rheum Dis. 2019;78(7):996-1002.
7. Leclair V, Galindo-Feria AS, Rothwell S, Krystufkova O, Zargar SS, Mann H, et al. Distinct HLA associations with autoantibody-defined subgroups in idiopathic inflammatory myopathies. EBioMedicine. 2023;96:104804.
8. Hill CL, Zhang Y, Sigurgeirsson B, Pukkala E, Mellemkjaer L, Airio A, et al. Frequency of specific cancer types in dermatomyositis and polymyositis: a population-based study. Lancet. 2001;357(9250):96-100.
9. Dani L, Ian Che W, Lundberg IE, Hellgren K, Holmqvist M. Overall and site-specific cancer before and after diagnosis of idiopathic inflammatory myopathies: A nationwide study 2002-2016. Semin Arthritis Rheum. 2021;51(1):331-7.
10. Dobloug GC, Svensson J, Lundberg IE, Holmqvist M. Mortality in idiopathic inflammatory myopathy: results from a Swedish nationwide population-based cohort study. Ann Rheum Dis. 2018;77(1):40-7.
11. Oddis CV, Conte CG, Steen VD, Medsger TA, Jr. Incidence of polymyositis-dermatomyositis: a 20-year study of hospital diagnosed cases in Allegheny County, PA 1963-1982. J Rheumatol. 1990;17(10):1329-34.
12. Ahlstrom G, Gunnarsson LG, Leissner P, Sjoden PO. Epidemiology of neuromuscular diseases, including the postpolio sequelae, in a Swedish county. Neuroepidemiology. 1993;12(5):262-9.
13. Koh ET, Seow A, Ong B, Ratnagopal P, Tjia H, Chng HH. Adult onset polymyositis/dermatomyositis: clinical and laboratory features and treatment response in 75 patients. Ann Rheum Dis. 1993;52(12):857-61.
14. Kusumi M, Nakashima K, Nakayama H, Takahashi K. Epidemiology of inflammatory neurological and inflammatory neuromuscular diseases in Tottori Prefecture, Japan. Psychiatry Clin Neurosci. 1995;49(3):169-74.
15. Kaipiainen-Seppanen O, Aho K. Incidence of rare systemic rheumatic and connective tissue diseases in Finland. J Intern Med. 1996;240(2):81-4.
16. Weitoft T. Occurrence of polymyositis in the county of Gavleborg, Sweden. Scand J Rheumatol. 1997;26(2):104-6.
17. Patrick M, Buchbinder R, Jolley D, Dennett X, Buchanan R. Incidence of inflammatory myopathies in Victoria, Australia, and evidence of spatial clustering. J Rheumatol. 1999;26(5):1094-100.

18. Lynn SJ, Sawyers SM, Moller PW, O'Donnell JL, Chapman PT. Adult-onset inflammatory myopathy: North Canterbury experience 1989-2001. Intern Med J. 2005;35(3):170-3.
19. Bernatsky S, Joseph L, Pineau CA, Belisle P, Boivin JF, Banerjee D, et al. Estimating the prevalence of polymyositis and dermatomyositis from administrative data: age, sex and regional differences. Ann Rheum Dis. 2009;68(7):1192-6.
20. Kuo CF, See LC, Yu KH, Chou IJ, Chang HC, Chiou MJ, et al. Incidence, cancer risk and mortality of dermatomyositis and polymyositis in Taiwan: a nationwide population study. Br J Dermatol. 2011;165(6):1273-9.
21. Barnabe C, Joseph L, Belisle P, Labrecque J, Barr SG, Fritzler MJ, et al. Prevalence of autoimmune inflammatory myopathy in the first nations population of Alberta, Canada. Arthritis Care Res (Hoboken). 2012;64(11):1715-9.
22. Furst DE, Amato AA, Iorga SR, Gajria K, Fernandes AW. Epidemiology of adult idiopathic inflammatory myopathies in a U.S. managed care plan. Muscle Nerve. 2012;45(5):676-83.
23. Smoyer-Tomic KE, Amato AA, Fernandes AW. Incidence and prevalence of idiopathic inflammatory myopathies among commercially insured, Medicare supplemental insured, and Medicaid enrolled populations: an administrative claims analysis. BMC Musculoskelet Disord. 2012;13:103.
24. Gupta R, Gow PJ. Inflammatory myopathies--a review of newly diagnosed patients (2004-2008) in the Counties Manukau region. N Z Med J. 2013;126(1370):89-95.
25. Rosa J, Garrot LF, Navarta DA, Saucedo C, Scolnik M, Bedran Z, et al. Incidence and prevalence of polymyositis and dermatomyositis in a health management organization in Buenos Aires. J Clin Rheumatol. 2013;19(6):303-7.
26. Tan JA, Roberts-Thomson PJ, Blumbergs P, Hakendorf P, Cox SR, Limaye V. Incidence and prevalence of idiopathic inflammatory myopathies in South Australia: a 30-year epidemiologic study of histology-proven cases. Int J Rheum Dis. 2013;16(3):331-8.
27. Yu KH, See LC, Kuo CF, Chou IJ, Chou MJ. Prevalence and incidence in patients with autoimmune rheumatic diseases: a nationwide population-based study in Taiwan. Arthritis Care Res (Hoboken). 2013;65(2):244-50.
28. Nagy-Vincze M, Bodoki L, Griger Z, Danko K. [Epidemiology of idiopathic inflammatory myopathy in Hungary]. Orv Hetil. 2014;155(41):1643-6.
29. Ohta A, Nagai M, Nishina M, Tomimitsu H, Kohsaka H. Prevalence and incidence of polymyositis and dermatomyositis in Japan. Mod Rheumatol. 2014;24(3):477-80.
30. Dobloug C, Garen T, Bitter H, Stjarne J, Stenseth G, Grovle L, et al. Prevalence and clinical characteristics of adult polymyositis and dermatomyositis; data from a large and unselected Norwegian cohort. Ann Rheum Dis. 2015;74(8):1551-6.
31. Parker MJS, Oldroyd A, Roberts ME, Ollier WE, New RP, Cooper RG, et al. Increasing incidence of adult idiopathic inflammatory myopathies in the City of Salford, UK: a 10-year epidemiological study. Rheumatol Adv Pract. 2018;2(2):rky035.
32. Cho SK, Kim H, Myung J, Nam E, Jung SY, Jang EJ, et al. Incidence and Prevalence of Idiopathic Inflammatory Myopathies in Korea: a Nationwide Population-based Study. J Korean Med Sci. 2019;34(8):e55.
33. Ostrovrsnik J, Hocevar A, Rotar Z, Krosel M, Cucnik S, Jurcic V, et al. The incidence of idiopathic inflammatory myopathies in the adult Slovenian population. Clin Rheumatol. 2019;38(2):279-83.
34. Svensson J, Arkema EV, Lundberg IE, Holmqvist M. Incidence and prevalence of idiopathic inflammatory myopathies in Sweden: a nationwide population-based study. Rheumatology (Oxford). 2017;56(5):802-10.

35. Nossent J, Keen H, Preen DB, Inderjeeth CA. The spectrum of idiopathic inflammatory myopathies in Western Australia: epidemiological characteristics and mortality over time. Rheumatol Int. 2023.
36. Meyer A, Meyer N, Schaeffer M, Gottenberg JE, Geny B, Sibilia J. Incidence and prevalence of inflammatory myopathies: a systematic review. Rheumatology (Oxford). 2015;54(1):50-63.
37. Badrising UA, Maat-Schieman M, van Duinen SG, Breedveld F, van Doorn P, van Engelen B, et al. Epidemiology of inclusion body myositis in the Netherlands: a nationwide study. Neurology. 2000;55(9):1385-7.
38. Mariampillai K, Granger B, Amelin D, Guiguet M, Hachulla E, Maurier F, et al. Development of a New Classification System for Idiopathic Inflammatory Myopathies Based on Clinical Manifestations and Myositis-Specific Autoantibodies. JAMA Neurology. 2018;75(12):1528.
39. Khoo T, Lilleker JB, Thong BY, Leclair V, Lamb JA, Chinoy H. Epidemiology of the idiopathic inflammatory myopathies. Nat Rev Rheumatol. 2023;19(11):695-712.
40. Bernstein RM, Morgan SH, Chapman J, Bunn CC, Mathews MB, Turner-Warwick M, et al. Anti-Jo-1 antibody: a marker for myositis with interstitial lung disease. Br Med J (Clin Res Ed). 1984;289(6438):151-2.
41. Cavagna L, Nuno L, Scire CA, Govoni M, Longo FJL, Franceschini F, et al. Clinical Spectrum Time Course in Anti Jo-1 Positive Antisynthetase Syndrome: Results From an International Retrospective Multicenter Study. Medicine (Baltimore). 2015;94(32):e1144.
42. Iaccarino L, Gatto M, Bettio S, Caso F, Rampudda M, Zen M, et al. Overlap connective tissue disease syndromes. Autoimmun Rev. 2013;12(3):363-73.
43. Horimoto AM, Silveira AF, Costa IP. Familial autoimmunity and polyautoimmunity in 60 Brazilian Midwest patients with systemic sclerosis. Rev Bras Reumatol. 2016.
44. Nuno-Nuno L, Joven BE, Carreira PE, Maldonado-Romero V, Larena-Grijalba C, Llorente Cubas I, et al. Overlap myositis, a distinct entity beyond primary inflammatory myositis: A retrospective analysis of a large cohort from the REMICAM registry. Int J Rheum Dis. 2019;22(8):1393-401.
45. Lilleker JB, Vencovsky J, Wang G, Wedderburn LR, Diederichsen LP, Schmidt J, et al. The EuroMyositis registry: an international collaborative tool to facilitate myositis research. Ann Rheum Dis. 2018;77(1):30-9.
46. Shah M, Mamyrova G, Targoff IN, Huber AM, Malley JD, Rice MM, et al. The clinical phenotypes of the juvenile idiopathic inflammatory myopathies. Medicine (Baltimore). 2013;92(1):25-41.
47. Rider LG, Nistala K. The juvenile idiopathic inflammatory myopathies: pathogenesis, clinical and autoantibody phenotypes, and outcomes. J Intern Med. 2016;280(1):24-38.
48. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). N Engl J Med. 1975;292(7):344-7.
49. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). N Engl J Med. 1975;292(8):403-7.
50. Griggs RC, Askanas V, DiMauro S, Engel A, Karpati G, Mendell JR, et al. Inclusion body myositis and myopathies. Ann Neurol. 1995;38(5):705-13.
51. Rose MR. 188th ENMC International Workshop: Inclusion Body Myositis, 2–4 December 2011, Naarden, The Netherlands. Neuromuscul Disord. 2013;23(12):1044-55.
52. Lundberg IE, Tjärnlund A, Bottai M, Werth VP, Pilkington C, Visser M, et al. 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Ann Rheum Dis. 2017;76(12):1955-64.

53. Aggarwal R, Oddis CV. Managing Myositis: A Practical Guide. Switzerland: Springer International Publishing; 2019.
54. Wedgwood RJ, Cook CD, Cohen J. Dermatomyositis; report of 26 cases in children with a discussion of endocrine therapy in 13. Pediatrics. 1953;12(4):447-66.
55. Mastaglia F, Price P, Walters S, Fabian V, Miller J, Zilko P. Familial inclusion body myositis in a mother and son with different ancestral MHC haplotypes. Neuromuscul Disord. 2006;16(11):754-8.
56. Limaye V, Walker J, Ahern M, Bardy P, Cox S, Roberts-Thomson P, et al. Monozygotic twins with distinct forms of idiopathic inflammatory myositis. Rheumatology (Oxford). 2009;48(7):855-7.
57. Ozaki T, Yamashita H, Hosono Y, Nakashima R, Mimori T, Takahashi Y, et al. Two patients in the same family with anti-ARS antibody-associated myositis. Mod Rheumatol. 2014;24(4):699-700.
58. Khosa S, Vazirian S, Mishra S, Khanlou N, Heinze E, Singh S, et al. Familial Inclusion Body Myositis with Negative Anti-cytosolic 5′-nucleotidase 1A (cN1A) Antibody Test. (P3.4-016). Neurology. 2019;92(15 Supplement):P3.4-016.
59. Ginn LR, Lin JP, Plotz PH, Bale SJ, Wilder RL, Mbauya A, et al. Familial autoimmunity in pedigrees of idiopathic inflammatory myopathy patients suggests common genetic risk factors for many autoimmune diseases. Arthritis Rheum. 1998;41(3):400-5.
60. Niewold TB, Wu SC, Smith M, Morgan GA, Pachman LM. Familial aggregation of autoimmune disease in juvenile dermatomyositis. Pediatrics. 2011;127(5):e1239-46.
61. Eaton WW, Rose NR, Kalaydjian A, Pedersen MG, Mortensen PB. Epidemiology of autoimmune diseases in Denmark. J Autoimmun. 2007;29(1):1-9.
62. Thomsen H, Li X, Sundquist K, Sundquist J, Forsti A, Hemminki K. Familial associations for rheumatoid autoimmune diseases. Rheumatol Adv Pract. 2020;4(2):rkaa048.
63. Rider LG, Gurley RC, Pandey JP, Garcia de la Torre I, Kalovidouris AE, O'Hanlon TP, et al. Clinical, serologic, and immunogenetic features of familial idiopathic inflammatory myopathy. Arthritis Rheum. 1998;41(4):710-9.
64. Mayhew AJ, Meyre D. Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities. Curr Genomics. 2017;18(4):332-40.
65. Rothwell S, Lamb JA, Chinoy H. New developments in genetics of myositis. Curr Opin Rheumatol. 2016;28(6):651-6.
66. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. Arthritis Res Ther. 2011;13(1):101.
67. Golan D, Lander ES, Rosset S. Measuring missing heritability: inferring the contribution of common variants. Proc Natl Acad Sci U S A. 2014;111(49):E5272-81.
68. Gambino CM, Aiello A, Accardi G, Caruso C, Candore G. Autoimmune diseases and 8.1 ancestral haplotype: An update. HLA. 2018;92(3):137-43.
69. O'Hanlon TP, Carrick DM, Arnett FC, Reveille JD, Carrington M, Gao X, et al. Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1 and -DQA1 allelic profiles and motifs define clinicopathologic groups in caucasians. Medicine (Baltimore). 2005;84(6):338-49.
70. Deakin CT, Bowes J, Rider LG, Miller FW, Pachman LM, Sanner H, et al. Association with HLA-DRbeta1 position 37 distinguishes juvenile dermatomyositis from adult-onset myositis. Hum Mol Genet. 2022;31(14):2471-81.
71. O'Hanlon TP, Carrick DM, Targoff IN, Arnett FC, Reveille JD, Carrington M, et al. Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1, and -DQA1 allelic profiles distinguish European American patients with different myositis autoantibodies. Medicine (Baltimore). 2006;85(2):111-27.

72. Bianchi M, Kozyrev SV, Notarnicola A, Hultin Rosenberg L, Karlsson A, Pucholt P, et al. Contribution of Rare Genetic Variation to Disease Susceptibility in a Large Scandinavian Myositis Cohort. Arthritis Rheumatol. 2022;74(2):342-52.
73. Zhou D, King EH, Rothwell S, Krystufkova O, Notarnicola A, Coss S, et al. Low copy numbers of complement C4 and C4A deficiency are risk factors for myositis, its subgroups and autoantibodies. Ann Rheum Dis. 2023;82(2):235-45.
74. Lintner KE, Patwardhan A, Rider LG, Abdul-Aziz R, Wu YL, Lundstrom E, et al. Gene copy-number variations (CNVs) of complement C4 and C4A deficiency in genetic risk and pathogenesis of juvenile dermatomyositis. Ann Rheum Dis. 2016;75(9):1599-606.
75. Rothwell S, Amos CI, Miller FW, Rider LG, Lundberg IE, Gregersen PK, et al. Identification of Novel Associations and Localization of Signals in Idiopathic Inflammatory Myopathies Using Genome-Wide Imputation. Arthritis Rheumatol. 2023;75(6):1021-7.
76. Acosta-Herrera M, Kerick M, Gonzalez-Serna D, Myositis Genetics C, Scleroderma Genetics C, Wijmenga C, et al. Genome-wide meta-analysis reveals shared new loci in systemic seropositive rheumatic diseases. Ann Rheum Dis. 2019;78(3):311-9.
77. Armstrong DL, Zidovetzki R, Alarcon-Riquelme ME, Tsao BP, Criswell LA, Kimberly RP, et al. GWAS identifies novel SLE susceptibility genes and explains the association of the HLA region. Genes Immun. 2014;15(6):347-54.
78. Cruz-Tapias P, Rojas-Villarraga A, Maier-Moore S, Anaya JM. HLA and Sjogren's syndrome susceptibility. A meta-analysis of worldwide studies. Autoimmun Rev. 2012;11(4):281-7.
79. Patsopoulos NA, Barcellos LF, Hintzen RQ, Schaefer C, van Duijn CM, Noble JA, et al. Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS Genet. 2013;9(11):e1003926.
80. Goyette P, Boucher G, Mallon D, Ellinghaus E, Jostins L, Huang H, et al. High-density mapping of the MHC identifies a shared role for HLA-DRB1*01:03 in inflammatory bowel diseases and heterozygous advantage in ulcerative colitis. Nat Genet. 2015;47(2):172-9.
81. Cucca F, Lampis R, Congia M, Angius E, Nutland S, Bain SC, et al. A correlation between the relative predisposition of MHC class II alleles to type 1 diabetes and the structure of their proteins. Hum Mol Genet. 2001;10(19):2025-37.
82. Maciel LM, Rodrigues SS, Dibbern RS, Navarro PA, Donadi EA. Association of the HLA-DRB1*0301 and HLA-DQA1*0501 alleles with Graves' disease in a population representing the gene contribution from several ethnic backgrounds. Thyroid. 2001;11(1):31-5.
83. Medrano LM, Dema B, Lopez-Larios A, Maluenda C, Bodas A, Lopez-Palacios N, et al. HLA and celiac disease susceptibility: new genetic factors bring open questions about the HLA influence and gene-dosage effects. PLoS One. 2012;7(10):e48403.
84. Seldin MF, Alkhairy OK, Lee AT, Lamb JA, Sussman J, Pirskanen-Matell R, et al. Genome-Wide Association Study of Late-Onset Myasthenia Gravis: Confirmation of TNFRSF11A and Identification of ZBTB10 and Three Distinct HLA Associations. Mol Med. 2016;21(1):769-81.
85. Arnett FC, Gourh P, Shete S, Ahn CW, Honey RE, Agarwal SK, et al. Major histocompatibility complex (MHC) class II alleles, haplotypes and epitopes which confer susceptibility or protection in systemic sclerosis: analyses in 1300 Caucasian, African-American and Hispanic cases and 1000 controls. Ann Rheum Dis. 2010;69(5):822-7.
86. Skibola CF, Berndt SI, Vijai J, Conde L, Wang Z, Yeager M, et al. Genome-wide association study identifies five susceptibility loci for follicular lymphoma outside the HLA region. Am J Hum Genet. 2014;95(4):462-71.
87. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet. 2010;42(6):508-14.

88. Gragert L, Fingerson S, Albrecht M, Maiers M, Kalaycio M, Hill BT. Fine-mapping of HLA associations with chronic lymphocytic leukemia in US populations. Blood. 2014;124(17):2657-65.
89. Conde L, Halperin E, Akers NK, Brown KM, Smedby KE, Rothman N, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nat Genet. 2010;42(8):661-4.
90. Cerhan JR, Berndt SI, Vijai J, Ghesquieres H, McKay J, Wang SS, et al. Genome-wide association study identifies multiple susceptibility loci for diffuse large B cell lymphoma. Nat Genet. 2014;46(11):1233-8.
91. Vijai J, Wang Z, Berndt SI, Skibola CF, Slager SL, de Sanjose S, et al. A genome-wide association study of marginal zone lymphoma shows association to the HLA region. Nat Commun. 2015;6:5751.
92. Nejentsev S, Howson JM, Walker NM, Szeszko J, Field SF, Stevens HE, et al. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. Nature. 2007;450(7171):887-92.
93. Berndt SI, Camp NJ, Skibola CF, Vijai J, Wang Z, Gu J, et al. Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia. Nat Commun. 2016;7:10933.
94. Cooper JD, Simmonds MJ, Walker NM, Burren O, Brand OJ, Guo H, et al. Seven newly identified loci for autoimmune thyroid disease. Hum Mol Genet. 2012;21(23):5202-8.
95. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012;491(7422):119-24.
96. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. 2014;506(7488):376-81.
97. Bentham J, Morris DL, Graham DSC, Pinder CL, Tombleson P, Behrens TW, 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):1457-64.
98. Mayes MD, Bossini-Castillo L, Gorlova O, Martin JE, Zhou X, Chen WV, et al. Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. Am J Hum Genet. 2014;94(1):47-61.
99. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet. 2009;41(6):703-7.
100. Trynka G, Hunt KA, Bockett NA, Romanos J, Mistry V, Szperl A, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat Genet. 2011;43(12):1193-201.
101. Lopez-Isac E, Acosta-Herrera M, Kerick M, Assassi S, Satpathy AT, Granja J, et al. GWAS for systemic sclerosis identifies multiple risk loci and highlights fibrotic and vasculopathy pathways. Nat Commun. 2019;10(1):4955.
102. Kleinstern G, Yan H, Hildebrandt MAT, Vijai J, Berndt SI, Ghesquieres H, et al. Inherited variants at 3q13.33 and 3p24.1 are associated with risk of diffuse large B-cell lymphoma and implicate immune pathways. Hum Mol Genet. 2020;29(1):70-9.
103. International Multiple Sclerosis Genetics C, Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet. 2013;45(11):1353-60.
104. Ramos PS, Criswell LA, Moser KL, Comeau ME, Williams AH, Pajewski NM, et al. A comprehensive analysis of shared loci between systemic lupus erythematosus (SLE)

and sixteen autoimmune diseases reveals limited genetic overlap. PLoS Genet. 2011;7(12):e1002406.
105. Wang YF, Zhang Y, Lin Z, Zhang H, Wang TY, Cao Y, et al. Identification of 38 novel loci for systemic lupus erythematosus and genetic heterogeneity between ancestral groups. Nat Commun. 2021;12(1):772.
106. Scofield RH, Bruner GR, Namjou B, Kimberly RP, Ramsey-Goldman R, Petri M, et al. Klinefelter's syndrome (47,XXY) in male systemic lupus erythematosus patients: support for the notion of a gene-dose effect from the X chromosome. Arthritis Rheum. 2008;58(8):2511-7.
107. Scofield RH, Lewis VM, Cavitt J, Kurien BT, Assassi S, Martin J, et al. 47XXY and 47XXX in Scleroderma and Myositis. ACR Open Rheumatol. 2022;4(6):528-33.
108. Hengstman GJ, van Venrooij WJ, Vencovsky J, Moutsopoulos HM, van Engelen BG. The relative prevalence of dermatomyositis and polymyositis in Europe exhibits a latitudinal gradient. Ann Rheum Dis. 2000;59(2):141-2.
109. Okada S, Weatherhead E, Targoff IN, Wesley R, Miller FW, International Myositis Collaborative Study G. Global surface ultraviolet radiation intensity may modulate the clinical and immunologic expression of autoimmune muscle disease. Arthritis Rheum. 2003;48(8):2285-93.
110. Parkes JE, Rothwell S, Oldroyd A, Chinoy H, Lamb JA, Myositis Genetics Consortium M. Genetic background may contribute to the latitude-dependent prevalence of dermatomyositis and anti-TIF1-gamma autoantibodies in adult patients with myositis. Arthritis Res Ther. 2018;20.
111. Shah M, Targoff IN, Rice MM, Miller FW, Rider LG, Childhood Myositis H. Ultraviolet Radiation Exposure Is Associated With Clinical and Autoantibody Phenotypes in Juvenile Myositis. Arthritis Rheum. 2013;65(7):1934-41.
112. Love LA, Weinberg CR, McConnaughey DR, Oddis CV, Medsger TA, Reveille JD, et al. Ultraviolet Radiation Intensity Predicts the Relative Distribution of Dermatomyositis and Anti-Mi-2 Autoantibodies in Women. Arthritis Rheum. 2009;60(8):2499-504.
113. Chinoy H, Adimulam S, Marriage F, New P, Vincze M, Zilahi E, et al. Interaction of HLA-DRB1*03 and smoking for the development of anti-Jo-1 antibodies in adult idiopathic inflammatory myopathies: a European-wide case study. Ann Rheum Dis. 2012;71(6):961-5.
114. Blanc PD, Järvholm B, Torén K. Prospective risk of rheumatologic disease associated with occupational exposure in a cohort of male construction workers. Am J Med. 2015;128(10):1094-101.
115. Goldman JA. Connective tissue disease in people exposed to organic chemical solvents: systemic sclerosis (scleroderma) in dry cleaning plant and aircraft industry workers. J Clin Rheumatol. 1996;2(4):185-90.
116. Koeger AC, Lang T, Alcaix D, Milleron B, Rozenberg S, Chaibi P, et al. Silica-associated connective tissue disease. A study of 24 cases. Medicine (Baltimore). 1995;74(5):221-37.
117. Ying DV, Schmajuk G, Trupin L, Blanc PD. Inorganic Dust Exposure During Military Service as a Predictor of Rheumatoid Arthritis and Other Autoimmune Conditions. Acr Open Rheumatology. 2021;3(7):466-74.
118. Labirua-Iturburu A, Selva-O'Callaghan A, Zock JP, Orriols R, Martinez-Gomez X, Vilardell-Tarres M. Occupational exposure in patients with the antisynthetase syndrome. Clin Rheumatol. 2014;33(2):221-5.
119. Campos G, Eisenreich MA, Lopes LM, D'Avila R, do Prado AD, Bredemeier M. Antisynthetase syndrome after acute massive inhalation of wood and paint dust. Scand J Rheumatol. 2016;45(5):425-6.

120. Webber MP, Moir W, Zeig-Owens R, Glaser MS, Jaber N, Hall C, et al. Nested Case-Control Study of Selected Systemic Autoimmune Diseases in World Trade Center Rescue/Recovery Workers. Arthritis & Rheumatology. 2015;67(5):1369-76.
121. Selva-O'Callaghan A, Labirua-Iturburu A, Pinal-Fernandez I. Antisynthetase antibodies in World Trade Center rescue and recovery workers with inflammatory myositis: comment on the article by Webber et al. Arthritis Rheumatol. 2015;67(10):2791.
122. Kawassaki AM, Kairalla RA, Carvalho CR, Costa AN. Hypersensitive pneumonitis: A trigger of autoimmunity in inflammatory myositis - A case series. Am J Respir Crit Care Med. 2010;181(1).
123. Leff RL, Burgess SH, Miller FW, Love LA, Targoff IN, Dalakas MC, et al. Distinct seasonal patterns in the onset of adult idiopathic inflammatory myopathy in patients with anti-Jo-1 and anti-signal recognition particle autoantibodies. Arthritis Rheum. 1991;34(11):1391-6.
124. Sarkar K, Weinberg CR, Oddis CV, Medsger TA, Jr., Plotz PH, Reveille JD, et al. Seasonal influence on the onset of idiopathic inflammatory myopathies in serologically defined groups. Arthritis Rheum. 2005;52(8):2433-8.
125. Vegosen LJ, Weinberg CR, O'Hanlon TP, Targoff IN, Miller FW, Rider LG. Seasonal birth patterns in myositis subgroups suggest an etiologic role of early environmental exposures. Arthritis Rheum. 2007;56(8):2719-28.
126. Nishina N, Sato S, Masui K, Gono T, Kuwana M. Seasonal and residential clustering at disease onset of anti-MDA5-associated interstitial lung disease. RMD Open. 2020;6(2).
127. So H, So J, Lam TT, Wong VT, Ho R, Li WL, et al. Seasonal Effect on Disease Onset and Presentation in Anti-MDA5 Positive Dermatomyositis. Front Med (Lausanne). 2022;9:837024.
128. Svensson J, Holmqvist M, Lundberg IE, Arkema EV. Infections and respiratory tract disease as risk factors for idiopathic inflammatory myopathies: a population-based case-control study. Ann Rheum Dis. 2017;76(11):1803-8.
129. De Santis M, Isailovic N, Motta F, Ricordi C, Ceribelli A, Lanza E, et al. Environmental triggers for connective tissue disease: the case of COVID-19 associated with dermatomyositis-specific autoantibodies. Curr Opin Rheumatol. 2021;33(6):514-21.
130. Giannini M, Ohana M, Nespola B, Zanframundo G, Geny B, Meyer A. Similarities between COVID-19 and anti-MDA5 syndrome: what can we learn for better care? Eur Respir J. 2020;56(3).
131. Gokhale Y, Patankar A, Holla U, Shilke M, Kalekar L, Karnik ND, et al. Dermatomyositis during COVID-19 Pandemic (A Case Series): Is there a Cause Effect Relationship? J Assoc Physicians India. 2020;68(11):20-4.
132. Christopher-Stine L, Casciola-Rosen LA, Hong G, Chung T, Corse AM, Mammen AL. A novel autoantibody recognizing 200-kd and 100-kd proteins is associated with an immune-mediated necrotizing myopathy. Arthritis Rheum. 2010;62(9):2757-66.
133. Wu YF, Lach B, Provias JP, Tarnopolsky MA, Baker SK. Statin-associated Autoimmune Myopathies: A Pathophysiologic Spectrum. Can J Neurol Sci. 2014;41(5):638-47.
134. Tiniakou E, Pinal-Fernandez I, Lloyd TE, Albayda J, Paik J, Werner JL, et al. More severe disease and slower recovery in younger patients with anti-3-hydroxy-3-methylglutaryl-coenzyme A reductase-associated autoimmune myopathy. Rheumatology (Oxford). 2017;56(5):787-94.
135. Sailler L, Pereira C, Bagheri A, Lapeyre-Mestre M, Montastruc JL, Arlet P, et al. Increased exposure to statins in patients developing chronic muscle diseases: a 2-year retrospective study. Ann Rheum Dis. 2008;67(5):614-9.

136. Caughey GE, Gabb GM, Ronson S, Ward M, Beukelman T, Hill CL, et al. Association of Statin Exposure With Histologically Confirmed Idiopathic Inflammatory Myositis in an Australian Population. JAMA Intern Med. 2018;178(9):1224-9.
137. Rider LG, Farhadi PN, Bayat N, Wilkerson J, Faiq A, McGrath J, et al. Infections and medications associated with onset of myositis in myovision, a national myositis patient registry. Arthritis and Rheumatology. 2017;69.
138. Aldrich J, Pundole X, Tummala S, Palaskas N, Andersen CR, Shoukier M, et al. Inflammatory Myositis in Cancer Patients Receiving Immune Checkpoint Inhibitors. Arthritis Rheumatol. 2021;73(5):866-74.
139. Xu M, Nie Y, Yang Y, Lu YT, Su Q. Risk of Neurological Toxicities Following the Use of Different Immune Checkpoint Inhibitor Regimens in Solid Tumors: A Systematic Review and Meta-analysis. Neurologist. 2019;24(3):75-83.
140. Moreira A, Loquai C, Pföhler C, Kähler KC, Knauss S, Heppt MV, et al. Myositis and neuromuscular side-effects induced by immune checkpoint inhibitors. Eur J Cancer. 2019;106:12-23.
141. Anquetil C, Salem JE, Lebrun-Vignes B, Johnson DB, Mammen AL, Stenzel W, et al. Immune Checkpoint Inhibitor-Associated Myositis: Expanding the Spectrum of Cardiac Complications of the Immunotherapy Revolution. Circulation. 2018;138(7):743-5.
142. Zong M, Bruton JD, Grundtman C, Yang H, Li JH, Alexanderson H, et al. TLR4 as receptor for HMGB1 induced muscle dysfunction in myositis. Ann Rheum Dis. 2013;72(8):1390-9.
143. Okiyama N, Ichimura Y, Shobo M, Tanaka R, Kubota N, Saito A, et al. Immune response to dermatomyositis-specific autoantigen, transcriptional intermediary factor 1gamma can result in experimental myositis. Ann Rheum Dis. 2021;80(9):1201-8.
144. Pinal-Fernandez I, Casal-Dominguez M, Derfoul A, Pak K, Plotz P, Miller FW, et al. Identification of Distinctive Interferon Gene Signatures in Different Types of Inflammatory Myopathy. Arthritis & Rheumatology. 2019;71.
145. Song J, Kim D, Hong J, Kim GW, Jung J, Park S, et al. Meta-Analysis of Polymyositis and Dermatomyositis Microarray Data Reveals Novel Genetic Biomarkers. Genes. 2019;10(11).
146. Askanas V, Engel WK, Nogalska A. Sporadic inclusion-body myositis: A degenerative muscle disease associated with aging, impaired muscle protein homeostasis and abnormal mitophagy. Biochim Biophys Acta. 2015;1852(4):633-43.
147. Kridin K, Kridin M, Amital H, Watad A, Khamaisi M. Mortality in Patients with Polymyositis and Dermatomyositis in an Israeli Population. Isr Med Assoc J. 2020;22(10):623-7.
148. Bai Z, Shen G, Dong L. Analysis of risk factors of interstitial lung disease and mortality rates in Chinese patients with idiopathic inflammatory myopathy. Int J Rheum Dis. 2021;24(6):815-27.
149. Li L, D'Silva KM, Lu N, Huang K, Esdaile JM, Choi HK, et al. Mortality trends in polymyositis and dermatomyositis: A general population-based study. Semin Arthritis Rheum. 2020;50(5):834-9.
150. D'Silva KM, Li L, Lu N, Ogdie A, Avina-Zubieta JA, Choi HK. Persistent premature mortality gap in dermatomyositis and polymyositis: a United Kingdom general population-based cohort study. Rheumatology (Oxford). 2021;60(6):2653-60.
151. Qiao P, Guo Q, Gao J, Ma D, Liu S, Gao X, et al. Long-term secular trends in dermatomyositis and polymyositis mortality in the USA from 1981 to 2020 according to underlying and multiple cause of death mortality data. Arthritis Res Ther. 2023;25(1):16.
152. Yang XL, Hao YJ, Zhang XH, Geng Y, Ji LL, Li GT, et al. Mortality of Chinese patients with polymyositis and dermatomyositis. Clin Rheumatol. 2020;39(5):1569-79.

153. Chuah TY, Kwan YH, Lui NL, Fong W. Risk Factors Associated with Mortality in Inflammatory Myositis: An Asian Perspective. Arthritis & Rheumatology. 2018;70.
154. Go DJ, Lee EY, Lee EB, Song YW, Konig MF, Park JK. Elevated Erythrocyte Sedimentation Rate Is Predictive of Interstitial Lung Disease and Mortality in Dermatomyositis: a Korean Retrospective Cohort Study. J Korean Med Sci. 2016;31(3):389-96.
155. Johnson C, Pinal-Fernandez I, Parikh R, Paik J, Albayda J, Mammen AL, et al. Assessment of Mortality in Autoimmune Myositis With and Without Associated Interstitial Lung Disease. Lung. 2016;194(5):733-7.
156. Li YH, Gao XJ, Li YM, Jia XH, Zhang XW, Xu Y, et al. Predictors and Mortality of Rapidly Progressive Interstitial Lung Disease in Patients With Idiopathic Inflammatory Myopathy: A Series of 474 Patients. Frontiers in Medicine. 2020;7.
157. Cobo-Ibanez T, Lopez-Longo FJ, Joven B, Carreira PE, Munoz-Fernandez S, Maldonado-Romero V, et al. Long-term pulmonary outcomes and mortality in idiopathic inflammatory myopathies associated with interstitial lung disease. Clin Rheumatol. 2019;38(3):803-15.
158. Wu CY, Wang Q, He LR, Yang EH, Zeng XF. Hospitalization mortality and associated risk factors in patients with polymyositis and dermatomyositis: A retrospective case-control study. PLoS One. 2018;13(2).
159. Mukae H, Ishimoto H, Sakamoto N, Hara S, Kakugawa T, Nakayama S, et al. Clinical differences between interstitial lung disease associated with clinically amyopathic dermatomyositis and classic dermatomyositis. Chest. 2009;136(5):1341-7.
160. Soni D, Maniscalco D, Veeraraghavan S, Cheeley J, Khosroshahi A. Anti-MDA5 Antibody Positive Dermatomyositis Is Not Always Associated with Recalcitrant Lung Disease or Mortality. Arthritis & Rheumatology. 2019;71.
161. Joven-Ibanez B, Nuno L, Lopez-Longo FJ, Martinez-Barrio J, Larena C, Maldonado V, et al. IS DISEASE SEVERITY A RISK FACTOR FOR CARDIOVASCULAR MORTALITY IN INFLAMMATORY MYOSITIS? ANALYSIS OF THE REGISTRY OF INFLAMMATORY MYOSITIS FROM THE MADRID COMMUNITY (REMICAM). Ann Rheum Dis. 2019;78:1560-+.
162. Huang LL, Fu Q, Ye Y, Lin YW, Yan QR, Chen S. High incidence and mortality of Pneumocystis jirovecii infection in anti-MDA5-antibody-positive dermatomyositis: experience from a single center. Arthritis Res Ther. 2021;23(1).
163. Muhammed H, Gupta L, Zanwar AA, Misra DP, Lawrence A, Agarwal V, et al. Infections Are Leading Cause of In-Hospital Mortality in Indian Patients With Inflammatory Myopathy. Jcr-Journal of Clinical Rheumatology. 2021;27(3):114-9.
164. Bai ZQ, Shen GF, Dong LL. Analysis of risk factors of interstitial lung disease and mortality rates in Chinese patients with idiopathic inflammatory myopathy. Int J Rheum Dis. 2021;24(6):815-27.
165. Mehta P, Agarwal V, Gupta L. High early mortality in idiopathic inflammatory myopathies: results from the inception cohort at a tertiary care centre in northern India. Rheumatology (Oxford). 2021;60(9):4281-90.
166. Stertz G. Polymyositis. Berl klin Wschr. 1916;53:489.
167. Sigurgeirsson B, Lindelof B, Edhag O, Allander E. Risk of cancer in patients with dermatomyositis or polymyositis. A population-based study. N Engl J Med. 1992;326(6):363-7.
168. Stockton D, Doherty VR, Brewster DH. Risk of cancer in patients with dermatomyositis or polymyositis, and follow-up implications: a Scottish population-based cohort study. Br J Cancer. 2001;85(1):41-5.
169. Troyanov Y, Targoff IN, Tremblay JL, Goulet JR, Raymond Y, Senecal JL. Novel classification of idiopathic inflammatory myopathies based on overlap syndrome

features and autoantibodies: analysis of 100 French Canadian patients. Medicine (Baltimore). 2005;84(4):231-49.
170. Leatham H, Schadt C, Chisolm S, Fretwell D, Chung L, Callen JP, et al. Evidence supports blind screening for internal malignancy in dermatomyositis: Data from 2 large US dermatology cohorts. Medicine (Baltimore). 2018;97(2):e9639.
171. Fardet L, Dupuy A, Gain M, Kettaneh A, Cherin P, Bachelez H, et al. Factors associated with underlying malignancy in a retrospective cohort of 121 patients with dermatomyositis. Medicine (Baltimore). 2009;88(2):91-7.
172. Dani L, Holmqvist M, Martinez MA, Trallero-Araguas E, Dastmalchi M, Svensson J, et al. Anti-transcriptional intermediary factor 1 gamma antibodies in cancer-associated myositis: a longitudinal study. Clin Exp Rheumatol. 2020;38(1):67-73.
173. Aussy A, Freret M, Gallay L, Bessis D, Vincent T, Jullien D, et al. The IgG2 Isotype of Anti-Transcription Intermediary Factor 1gamma Autoantibodies Is a Biomarker of Cancer and Mortality in Adult Dermatomyositis. Arthritis Rheumatol. 2019;71(8):1360-70.
174. Aljohani G, Bin Awad EA, Alshahrani K, Alsaqar MM, Albogami B, Almotywee SH, et al. The prevalence, clinical features, predictive factors and investigations to screen for cancer in patients with inflammatory myositis. A case series from two tertiary care centers in Riyadh, Saudi Arabia. Saudi Med J. 2021;42(1):100-4.
175. Dobloug GC, Garen T, Brunborg C, Gran JT, Molberg O. Survival and cancer risk in an unselected and complete Norwegian idiopathic inflammatory myopathy cohort. Semin Arthritis Rheum. 2015;45(3):301-8.
176. Huang YL, Chen YJ, Lin MW, Wu CY, Liu PC, Chen TJ, et al. Malignancies associated with dermatomyositis and polymyositis in Taiwan: a nationwide population-based study. Br J Dermatol. 2009;161(4):854-60.
177. Sung YK, Jung SY, Kim H, Choi S, Im SG, Cha EJ, et al. Temporal relationship between idiopathic inflammatory myopathies and malignancies and its mortality: a nationwide population-based study. Clin Rheumatol. 2020;39(11):3409-16.
178. Buchbinder R, Forbes A, Hall S, Dennett X, Giles G. Incidence of malignant disease in biopsy-proven inflammatory myopathy. A population-based cohort study. Ann Intern Med. 2001;134(12):1087-95.
179. Azuma K, Yamada H, Ohkubo M, Yamasaki Y, Yamasaki M, Mizushima M, et al. Incidence and predictive factors for malignancies in 136 Japanese patients with dermatomyositis, polymyositis and clinically amyopathic dermatomyositis. Mod Rheumatol. 2011;21(2):178-83.
180. Qiang JK, Kim WB, Baibergenova A, Alhusayen R. Risk of Malignancy in Dermatomyositis and Polymyositis. J Cutan Med Surg. 2017;21(2):131-6.
181. Lim CH, Tseng CW, Lin CT, Huang WN, Chen YH, Chen YM, et al. The clinical application of tumor markers in the screening of malignancies and interstitial lung disease of dermatomyositis/polymyositis patients: A retrospective study. SAGE Open Med. 2018;6:2050312118781895.
182. Fujimoto M, Hamaguchi Y, Kaji K, Matsushita T, Ichimura Y, Kodera M, et al. Myositis-specific anti-155/140 autoantibodies target transcription intermediary factor 1 family proteins. Arthritis Rheum. 2012;64(2):513-22.
183. Kadoya M, Hida A, Maeda MH, Taira K, Ikenaga C, Uchio N, et al. Cancer association as a risk factor for anti-HMGCR antibody-positive myopathy. Neurology-Neuroimmunology & Neuroinflammation. 2016;3(6).
184. Allenbach Y, Keraen J, Bouvier AM, Jooste V, Champtiaux N, Hervier B, et al. High risk of cancer in autoimmune necrotizing myopathies: usefulness of myositis specific antibody. Brain. 2016;139:2131-5.

185. Limaye V, Luke C, Tucker G, Hill C, Lester S, Blumbergs P, et al. The incidence and associations of malignancy in a large cohort of patients with biopsy-determined idiopathic inflammatory myositis. Rheumatol Int. 2013;33(4):965-71.
186. Naddaf E, Shelly S, Mandrekar J, Chamberlain AM, Hoffman EM, Ernste FC, et al. Survival and associated comorbidities in inclusion body myositis. Rheumatology (Oxford). 2022;61(5):2016-24.
187. Shelly S, Beecher G, Milone M, Liewluck T, Ernste F, Triplett J, et al. Cancer and immune-mediated necrotizing myopathy: a longitudinal referral case-controlled outcomes evaluation. Rheumatology (Oxford). 2022;62(1):281-9.
188. Yang H, Peng Q, Yin L, Li S, Shi J, Zhang Y, et al. Identification of multiple cancer-associated myositis-specific autoantibodies in idiopathic inflammatory myopathies: a large longitudinal cohort study. Arthritis Res Ther. 2017;19(1):259.
189. Chen YJ, Wu CY, Huang YL, Wang CB, Shen JL, Chang YT. Cancer risks of dermatomyositis and polymyositis: a nationwide cohort study in Taiwan. Arthritis Res Ther. 2010;12(2).
190. Li LB, Liu CX, Wang Q, Wu CY, Zhang YF, Cheng LL, et al. Analysis of myositis autoantibodies in Chinese patients with cancer-associated myositis. J Clin Lab Anal. 2020;34(8).
191. Chen YJ, Wu CY, Shen JL. Predicting factors of malignancy in dermatomyositis and polymyositis: a case-control study. Br J Dermatol. 2001;144(4):825-31.
192. Peng JC, Sheen TS, Hsu MM. Nasopharyngeal carcinoma with dermatomyositis. Analysis of 12 cases. Arch Otolaryngol Head Neck Surg. 1995;121(11):1298-301.
193. Li Y, Jia X, Sun X, Shi L, Lin F, Gan Y, et al. Risk factors for cancer-associated myositis: A large-scale multicenter cohort study. Int J Rheum Dis. 2021;24(2):268-73.
194. Ang P, Sugeng MW, Chua SH. Classical and amyopathic dermatomyositis seen at the National Skin Centre of Singapore: a 3-year retrospective review of their clinical characteristics and association with malignancy. Annals of the Academy of Medicine, Singapore. 2000;29(2):219-23.
195. So MW, Koo BS, Kim YG, Lee CK, Yoo B. Idiopathic inflammatory myopathy associated with malignancy: a retrospective cohort of 151 Korean patients with dermatomyositis and polymyositis. J Rheumatol. 2011;38(11):2432-5.
196. Kang EH, Lee SJ, Ascherman DP, Lee YJ, Lee EY, Lee EB, et al. Temporal relationship between cancer and myositis identifies two distinctive subgroups of cancers: impact on cancer risk and survival in patients with myositis. Rheumatology (Oxford). 2016;55(9):1631-41.
197. Chow WH, Gridley G, Mellemkjaer L, McLaughlin JK, Olsen JH, Fraumeni JF, Jr. Cancer risk following polymyositis and dermatomyositis: a nationwide cohort study in Denmark. Cancer Causes Control. 1995;6(1):9-13.
198. Andras C, Ponyi A, Constantin T, Csiki Z, Szekanecz E, Szodoray P, et al. Dermatomyositis and polymyositis associated with malignancy: a 21-year retrospective study. J Rheumatol. 2008;35(3):438-44.
199. Antiochos BB, Brown LA, Li Z, Tosteson TD, Wortmann RL, Rigby WF. Malignancy is associated with dermatomyositis but not polymyositis in Northern New England, USA. J Rheumatol. 2009;36(12):2704-10.
200. Chow WH, Gridley G, Mellemkjaer L, McLaughlin JK, Olsen JH, Fraumeni JF. CANCER RISK FOLLOWING POLYMYOSITIS AND DERMATOMYOSITIS - A NATIONWIDE COHORT STUDY IN DENMARK. Cancer Causes Control. 1995;6(1):9-13.
201. Oldroyd A, Sergeant JC, New P, McHugh NJ, Betteridge Z, Lamb JA, et al. The temporal relationship between cancer and adult onset anti-transcriptional intermediary factor 1 antibody-positive dermatomyositis. Rheumatology (Oxford). 2019;58(4):650-5.

202. Zhao Y, Su H, Yin X, Hou H, Wang Y, Xu Y, et al. Cancer associated autoantibodies in idiopathic inflammatory myopathies: A retrospective cohort from a single center in China. Med Clin (Barc). 2023;160(1):10-6.
203. Zhao L, Xie S, Zhou B, Shen C, Li L, Pi W, et al. Machine Learning Algorithms Identify Clinical Subtypes and Cancer in Anti-TIF1gamma+ Myositis: A Longitudinal Study of 87 Patients. Front Immunol. 2022;13:802499.
204. Lu X, Yang H, Shu X, Chen F, Zhang Y, Zhang S, et al. Factors predicting malignancy in patients with polymyositis and dermatomyostis: a systematic review and meta-analysis. PLoS One. 2014;9(4):e94128.
205. Wang J, Guo G, Chen G, Wu B, Lu L, Bao L. Meta-analysis of the association of dermatomyositis and polymyositis with cancer. Br J Dermatol. 2013;169(4):838-47.
206. Oldroyd AG, Allard AB, Callen JP, Chinoy H, Chung L, Fiorentino D, et al. A systematic review and meta-analysis to inform cancer screening guidelines in idiopathic inflammatory myopathies. Rheumatology. 2021;60(6):2615-28.
207. Nicoletis I, Pasco J, Maillot F, Goupille P, Corcia P, Grammatico-Guillon L, et al. High pre-treatment neutrophil-to-lymphocyte ratio in patients with dermatomyositis/polymyositis predicts an increased risk of cancer. Eur J Dermatol. 2020;30(2):133-9.
208. Shao C, Li S, Sun Y, Zhang Y, Xu K, Zhang X, et al. Clinical characteristics and prognostic analysis of Chinese dermatomyositis patients with malignancies. Medicine (Baltimore). 2020;99(34):e21899.
209. Wong RC, Brown S, Clarke BE, Klingberg S, Zimmerman PV. Transient elevation of the tumor markers CA 15-3 and CASA as markers of interstitial lung disease rather than underlying malignancy in dermatomyositis sine myositis. J Clin Rheumatol. 2002;8(4):204-7.
210. Kadota H, Gono T, Yamaguchi Y, Watanabe E, Okazaki Y, Nakajima Y, et al. INITIAL PREDICTORS FOR MORTALITY IN PATIENTS WITH CANCER-ASSOCIATED MYOSITIS: A MULTICENTER RETROSPECTIVE COHORT IN JAPAN. Ann Rheum Dis. 2018;77:752-.
211. Albayda J, Mecoli C, Casciola-Rosen L, Danoff SK, Lin CT, Hines D, et al. A North American Cohort of Anti-SAE Dermatomyositis: Clinical Phenotype, Testing, and Review of Cases. ACR Open Rheumatol. 2021;3(5):287-94.
212. Fiorentino DF, Chung LS, Christopher-Stine L, Zaba L, Li SF, Mammen AL, et al. Most Patients With Cancer-Associated Dermatomyositis Have Antibodies to Nuclear Matrix Protein NXP-2 or Transcription Intermediary Factor 1 gamma. Arthritis and Rheumatism. 2013;65(11):2954-62.
213. Allenzara A, Sadeghi N, Alvarez C, Maczuga S, Helm M, Olsen N, et al. Obesity is an Independent Risk Factor for Cancer Development Following Diagnosis of Dermatomyositis. Seminars in Arthritis and Rheumatism. 2023:152283.
214. Chen H, Yang H, Cheng QX, Ge YP, Peng QL, Zhang YM, et al. A novel autoantibody targeting calreticulin is associated with cancer in patients with idiopathic inflammatory myopathies. Clin Transl Immunology. 2020;9(10):e1195.
215. Trallero-Araguas E, Rodrigo-Pendas JA, Selva-O'Callaghan A, Martinez-Gomez X, Bosch X, Labrador-Horrillo M, et al. Usefulness of anti-p155 autoantibody for diagnosing cancer-associated dermatomyositis: a systematic review and meta-analysis. Arthritis Rheum. 2012;64(2):523-32.
216. Chinoy H, Fertig N, Oddis CV, Ollier WER, Cooper RG. The diagnostic utility of myositis autoantibody testing for predicting the risk of cancer-associated myositis. Ann Rheum Dis. 2007;66(10):1345-9.
217. Fiorentino D, Mecoli CA, Igusa T, Albayda J, Paik JJ, Tiniakou E, et al. Association of Anti-CCAR1 Autoantibodies With Decreased Cancer Risk Relative to the

General Population in Patients With Anti-Transcriptional Intermediary Factor 1gamma-Positive Dermatomyositis. Arthritis Rheumatol. 2023;75(7):1238-45.
218. Hosono Y, Sie B, Pinal-Fernandez I, Pak K, Mecoli CA, Casal-Dominguez M, et al. Coexisting autoantibodies against transcription factor Sp4 are associated with decreased cancer risk in patients with dermatomyositis with anti-TIF1gamma autoantibodies. Ann Rheum Dis. 2023;82(2):246-52.
219. Shih Y, Chen S, Huang J, Chen Y, Zhu Z, Zhao Q, et al. Serum level of galectin-9 as a potential biomarker for high risk of malignancy in dermatomyositis. Rheumatology (Oxford). 2023.
220. Inaguma G, Shimada A, Tsunoda J, Matsuzaki T, Nishi T, Seki H, et al. Inflammatory breast cancer associated with amyopathic dermatomyositis: a case report. Surgical Case Reports. 2020;6(1).
221. Kubecek O, Soukup T, Paulik A, Kopecky J. Dermatomyositis with anti-TIF-1 gamma antibodies as a presenting symptom of underlying triple-negative breast cancer: a case report. Bmc Cancer. 2016;16.
222. Daido W, Yamasaki M, Morio Y, Funaishi K, Ishiyama S, Deguchi N, et al. A Rare Combination of Dermatomyositis, Interstitial Pneumonia, and Lung Cancer in a Patient Treated with Immunosuppressive Therapy and Chemotherapy. Internal Medicine. 2018;57(6):849-53.
223. Patel SR, Saliba AN, Steel S, Liewluck T, Mahipal A. Metastatic Colon Cancer Presenting With Immune-mediated Necrotizing Myopathy. Clinical Colorectal Cancer. 2021;20(1):E71-E3.
224. Ponyi A, Constantin T, Garami M, Andras C, Tallai B, Vancsa A, et al. Cancer-associated myositis: clinical features and prognostic signs. Ann N Y Acad Sci. 2005;1051:64-71.
225. Ogawa-Momohara M, Muro Y, Mitsuma T, Katayama M, Yanaba K, Nara M, et al. Strong correlation between cancer progression and anti-transcription intermediary factor 1 gamma antibodies in dermatomyositis patients. Clinical and Experimental Rheumatology. 2018;36(6):990-5.
226. Ikeda N, Yamaguchi Y, Kanaoka M, Ototake Y, Akita A, Watanabe T, et al. Clinical significance of serum levels of anti-transcriptional intermediary factor 1-gamma antibody in patients with dermatomyositis. J Dermatol. 2020;47(5):490-6.
227. Pinal-Fernandez I, Ferrer-Fabregas B, Trallero-Araguas E, Balada E, Martinez MA, Milisenda JC, et al. Tumour TIF1 mutations and loss of heterozygosity related to cancer-associated myositis. Rheumatology. 2018;57(2):388-96.
228. Mohassel P, Rosen P, Casciola-Rosen L, Pak K, Mammen AL. Expression of the dermatomyositis autoantigen transcription intermediary factor 1gamma in regenerating muscle. Arthritis Rheumatol. 2015;67(1):266-72.
229. Zampieri S, Doria A, Adami N, Biral D, Vecchiato M, Savastano S, et al. Subclinical myopathy in patients affected with newly diagnosed colorectal cancer at clinical onset of disease: evidence from skeletal muscle biopsies. Neurol Res. 2010;32(1):20-5.
230. Pinal-Fernandez I, Ferrer-Fabregas B, Trallero-Araguas E, Balada E, Martinez MA, Milisenda JC, et al. Tumour TIF1 mutations and loss of heterozygosity related to cancer-associated myositis. Rheumatology (Oxford). 2018;57(2):388-96.
231. Cordel N, Derambure C, Coutant S, Mariette X, Jullien D, Debarbieux S, et al. TRIM33 gene somatic mutations identified by next generation sequencing in neoplasms of patients with anti-TIF1gamma positive cancer-associated dermatomyositis. Rheumatology (Oxford). 2021;60(12):5863-7.
232. Selva-O'Callaghan A, Ros J, Gil-Vila A, Vila-Pijoan G, Trallero-Araguas E, Pinal-Fernandez I. Malignancy and myositis, from molecular mimicry to tumor infiltrating lymphocytes. Neuromuscul Disord. 2019;29(11):819-25.

233. Mascella F, Gianni L, Affatato A, Fantini M. Aromatase inhibitors and anti-synthetase syndrome. Int J Immunopathol Pharmacol. 2016;29(3):494-7.
234. Pokhai G, Buzzola R, Abrudescu A. Letrozole-induced very early systemic sclerosis in a patient with breast cancer: a case report. Archives of Rheumatology. 2014;29(2):126-9.
235. Fisher J, Patel M, Miller M, Burris K. Anastrozole-induced subacute cutaneous lupus erythematosus. Cutis. 2016;98(2):E22-6.
236. Lundberg IE, Fujimoto M, Vencovsky J, Aggarwal R, Holmqvist M, Christopher-Stine L, et al. Idiopathic inflammatory myopathies. Nature Reviews Disease Primers. 2021;7(1):86.
237. Zhao H, Wu L, Yan G, Chen Y, Zhou M, Wu Y, et al. Inflammation and tumor progression: signaling pathways and targeted intervention. Signal Transduct Target Ther. 2021;6(1):263.
238. Haller MC, Royuela A, Nagler EV, Pascual J, Webster AC. Steroid avoidance or withdrawal for kidney transplant recipients. Cochrane Database Syst Rev. 2016;2016(8):CD005632.
239. Ognenovski VM, Marder W, Somers EC, Johnston CM, Farrehi JG, Selvaggi SM, et al. Increased incidence of cervical intraepithelial neoplasia in women with systemic lupus erythematosus treated with intravenous cyclophosphamide. J Rheumatol. 2004;31(9):1763-7.
240. Jiyad Z, Olsen CM, Burke MT, Isbel NM, Green AC. Azathioprine and Risk of Skin Cancer in Organ Transplant Recipients: Systematic Review and Meta-Analysis. Am J Transplant. 2016;16(12):3490-503.
241. Cho HG, Kuo KY, Xiao K, Batra P, Li S, Tang JY, et al. Azathioprine and risk of multiple keratinocyte cancers. J Am Acad Dermatol. 2018;78(1):27-8 e1.
242. Papaioannou A, Chevet E. Driving Cancer Tumorigenesis and Metastasis Through UPR Signaling. Curr Top Microbiol Immunol. 2018;414:159-92.
243. Wang H, Jiang S, Zhang Y, Pan K, Xia J, Chen M. High expression of thymosin beta 10 predicts poor prognosis for hepatocellular carcinoma after hepatectomy. World J Surg Oncol. 2014;12:226.
244. Zhang X, Ren D, Guo L, Wang L, Wu S, Lin C, et al. Thymosin beta 10 is a key regulator of tumorigenesis and metastasis and a novel serum marker in breast cancer. Breast Cancer Res. 2017;19(1):15.
245. Pachman LM, Hayford JR, Hochberg MC, Pallansch MA, Chung A, Daugherty CD, et al. New-onset juvenile dermatomyositis: comparisons with a healthy cohort and children with juvenile rheumatoid arthritis. Arthritis Rheum. 1997;40(8):1526-33.
246. Mellemkjaer L, Alexander F, Olsen JH. Cancer among children of parents with autoimmune diseases. Br J Cancer. 2000;82(7):1353-7.
247. Engels EA, Cerhan JR, Linet MS, Cozen W, Colt JS, Davis S, et al. Immune-related conditions and immune-modulating medications as risk factors for non-Hodgkin's lymphoma: a case-control study. Am J Epidemiol. 2005;162(12):1153-61.
248. Landgren O, Engels EA, Caporaso NE, Gridley G, Mellemkjaer L, Hemminki K, et al. Patterns of autoimmunity and subsequent chronic lymphocytic leukemia in Nordic countries. Blood. 2006;108(1):292-6.
249. Landgren O, Engels EA, Pfeiffer RM, Gridley G, Mellemkjaer L, Olsen JH, et al. Autoimmunity and susceptibility to Hodgkin lymphoma: a population-based case-control study in Scandinavia. J Natl Cancer Inst. 2006;98(18):1321-30.
250. Landgren O, Linet MS, McMaster ML, Gridley G, Hemminki K, Goldin LR. Familial characteristics of autoimmune and hematologic disorders in 8,406 multiple myeloma patients: a population-based case-control study. Int J Cancer. 2006;118(12):3095-8.

251. Hemminki K, Forsti A, Sundquist K, Sundquist J, Li X. Familial associations of lymphoma and myeloma with autoimmune diseases. Blood Cancer J. 2017;7(1):e515.
252. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nat Genet. 2017;49(7):1126-32.
253. Urayama KY, Jarrett RF, Hjalgrim H, Diepstra A, Kamatani Y, Chabrier A, et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. J Natl Cancer Inst. 2012;104(3):240-53.
254. Leo PJ, Madeleine MM, Wang S, Schwartz SM, Newell F, Pettersson-Kymmer U, et al. Defining the genetic susceptibility to cervical neoplasia-A genome-wide association study. PLoS Genet. 2017;13(8):e1006866.
255. Rashkin SR, Graff RE, Kachuri L, Thai KK, Alexeeff SE, Blatchins MA, et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. Nat Commun. 2020;11(1):4423.
256. Shi Y, Li L, Hu Z, Li S, Wang S, Liu J, et al. A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. Nat Genet. 2013;45(8):918-22.
257. Jiang DK, Sun J, Cao G, Liu Y, Lin D, Gao YZ, et al. Genetic variants in STAT4 and HLA-DQ genes confer risk of hepatitis B virus-related hepatocellular carcinoma. Nat Genet. 2013;45(1):72-5.
258. Osman Y, Elsharkawy T, Hashim TM, Alratroot JA, Alsuwat HS, Otaibi WMA, et al. Functional multigenic variations associated with hodgkin lymphoma. Int J Lab Hematol. 2021;43(6):1472-82.
259. Adolphe C, Xue A, Fard AT, Genovesi LA, Yang J, Wainwright BJ. Genetic and functional interaction network analysis reveals global enrichment of regulatory T cell genes influencing basal cell carcinoma susceptibility. Genome Med. 2021;13(1):19.
260. Liyanage UE, Law MH, Han X, An J, Ong JS, Gharahkhani P, et al. Combined analysis of keratinocyte cancers identifies novel genome-wide loci. Hum Mol Genet. 2019;28(18):3148-60.
261. Brandes N, Linial N, Linial M. Genetic association studies of alterations in protein function expose recessive effects on cancer predisposition. Sci Rep. 2021;11(1):14901.
262. Yamaguchi K, Iwamoto H, Sakamoto S, Horimasu Y, Masuda T, Miyamoto S, et al. AGER rs2070600 polymorphism elevates neutrophil-lymphocyte ratio and mortality in metastatic lung adenocarcinoma. Oncotarget. 2017;8(55):94382-92.
263. Danaei G, Vander Hoorn S, Lopez AD, Murray CJ, Ezzati M, Comparative Risk Assessment collaborating g. Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. Lancet. 2005;366(9499):1784-93.
264. Carr S, Smith C, Wernberg J. Epidemiology and Risk Factors of Melanoma. Surg Clin North Am. 2020;100(1):1-12.
265. Oldroyd AGS, Callen JP, Chinoy H, Chung L, Fiorentino D, Gordon P, et al. International Guideline for Idiopathic Inflammatory Myopathy-Associated Cancer Screening: an International Myositis Assessment and Clinical Studies Group (IMACS) initiative. Nat Rev Rheumatol. 2023;19(12):805-17.
266. Kardes S, Gupta L, Aggarwal R. Cancer and myositis: Who, when, and how to screen. Best Practice & Research Clinical Rheumatology. 2022;36(2):101771.
267. Selva-O'Callaghan A, Trallero-Araguas E, Ros J, Gil-Vila A, Lostes J, Agusti A, et al. Management of Cancer-Associated Myositis. Curr Treatm Opt Rheumatol. 2022;8(4):91-104.
268. Ferronato M, Lalanne C, Quarneti C, Cevolani M, Ricci C, Granito A, et al. Paraneoplastic Anti-Tif1-gamma Autoantibody-positive Dermatomyositis as Clinical

Presentation of Hepatocellular Carcinoma Recurrence. Journal of Clinical and Translational Hepatology. 2023;11(1):253-9.
269. Suga T, Abe T, Oiwa H. Anti-nuclear Matrix Protein 2 Antibody-positive Dermatomyositis Associated with Cervical Cancer Recurrence after 6 Years of Stable Disease. Internal Medicine. 2022;61(22):3445-8.
270. Gu JN, Yan W, Gao QL, Chen L. Anti-OJ antibody-positive anti-synthetase syndrome with repeated arthritis, fever, and recurrent liver cancer: a case report. Journal of Gastrointestinal Oncology. 2022;13(5):2672-8.
271. Baleiras MM, Maduro L, Vasques C, Ferreira F, Pinto MM, Martins A. Paraneoplastic dermatomyositis and prostate cancer: Myopathy regression under cancer-directed therapy. Dermatology Reports. 2021;13(3).
272. Mahajan C, Ray C, Singhal A, Thakur B, Hadi R, Singh P. Dermatomyositis as a Paraneoplastic Manifestation of Ovarian Cancer: A Case Report and Review of Literature. Indian Journal of Gynecologic Oncology. 2020;18(1).
273. Lim D, Landon-Cardinal O, Belisle A, Davar S. A case of dermatomyositis with anti-TIF1γ antibodies revealing isolated para-aortic lymphadenopathy metastatic recurrence of endometrial cancer: A case report. Sage Open Medical Case Reports. 2020;8.
274. Bientinesi R, Ragonese M, Pinto F, Bassi PF, Sacco E. Paraneoplastic Dermatomyositis Associated With Panurothelial Transitional Cell Carcinoma: A Case Report and Literature Review. Clinical Genitourinary Cancer. 2016;14(2):E199-E201.
275. Landriscina M, Gerardi AM, Fersini A, Modoni S, Stoppino LP, Macarini L, et al. Multiple skeletal muscle metastases from colon carcinoma preceded by paraneoplastic dermatomyositis. Case Rep Med. 2013;2013:392609.
276. Fitzpatrick J, Wallace WA, Lang S, Aboumarzouk OM, Windsor P, Nabi G. Recurrent dermatomyositis manifesting as a sign of recurrent transitional cell carcinoma of urinary bladder: Long-term survival. Urol Ann. 2014;6(3):264-6.
277. Nagano Y, Inoue Y, Shimura T, Fujikawa H, Okugawa Y, Hiro J, et al. Exacerbation of Dermatomyositis with Recurrence of Rectal Cancer: A Case Report. Case Rep Oncol. 2015;8(3):482-6.
278. Ono K, Shimomura M, Toyota K, Kagimoto A, Tsukiyama N, Shishida M, et al. Successful resection of liver metastasis detected by exacerbation of skin symptom in a patient with dermatomyositis accompanied by rectal cancer: a case report and literature review. Surg Case Rep. 2017;3(1):3.
279. Field C, Goff BA. Dermatomyositis - key to diagnosing ovarian cancer, monitoring treatment and detecting recurrent disease: Case report. Gynecol Oncol Rep. 2018;23:1-3.
280. Murgic J, Prpic M, Kirac I, Camino-Varela AM, Bolanca A, Kusic Z. Dermatomyositis as Paraneoplastic Syndrome of Peritoneal and Ovarian Relapse after Long-Term Complete Remission in Patient with Metastatic Bilateral Breast Cancer. Collegium Antropologicum. 2012;36(1):325-9.
281. Lai-Tiong F. Dermatomyositis revealing a 18-year breast cancer recurrence. European Journal of Gynaecological Oncology. 2018;39(2):292-3.
282. Psomiadou V, Gkegkes ID, Iavazzo C. Dermatomyositis and/or polymyositis as a paraneoplastic manifestation of ovarian cancer: a systematic review. Wspolczesna Onkologia-Contemporary Oncology. 2020;24(4):252-7.
283. Gkegkes ID, Minis EE, Iavazzo C. Dermatomyositis and colorectal cancer: a systematic review. Ir J Med Sci. 2018;187(3):615-20.
284. Hemminki K, Liu X, Ji J, Sundquist J, Sundquist K. Effect of autoimmune diseases on risk and survival in histology-specific lung cancer. Eur Respir J. 2012;40(6):1489-95.

285. Park JK, Yang JA, Ahn EY, Chang SH, Song YW, Curtis JR, et al. Survival rates of cancer patients with and without rheumatic disease: a retrospective cohort analysis. BMC Cancer. 2016;16:381.
286. Li N, Gao L, Bai C, Zhao L, Shao Y. Clinical features and prognosis of lung cancer in patients with connective tissue diseases: a retrospective cohort study. Front Oncol. 2023;13:1167213.
287. Wakata N, Kurihara T, Saito E, Kinoshita M. Polymyositis and dermatomyositis associated with malignancy: a 30-year retrospective study. Int J Dermatol. 2002;41(11):729-34.
288. Andras C, Bodoki L, Nagy-Vincze M, Griger Z, Csiki E, Danko K. Retrospective Analysis of Cancer-Associated Myositis Patients over the Past 3 Decades in a Hungarian Myositis Cohort. Pathol Oncol Res. 2020;26(3):1749-55.
289. Basset-Seguin N, Roujeau JC, Gherardi R, Guillaume JC, Revuz J, Touraine R. Prognostic factors and predictive signs of malignancy in adult dermatomyositis. A study of 32 cases. Arch Dermatol. 1990;126(5):633-7.
290. Mebazâa A, Boussen H, Nouira R, Rokbani L, Ben Osman-Dhahri A, Bouaouina N, et al. Dermatomyositis and malignancy in Tunisia: A multicenter national retrospective study of 20 cases. Journal of the American Academy of Dermatology. 2003;48(4):530-4.
291. Allenbach Y, Keraen J, Bouvier AM, Jooste V, Champtiaux N, Hervier B, et al. High risk of cancer in autoimmune necrotizing myopathies: usefulness of myositis specific antibody. Brain. 2016;139(Pt 8):2131-5.
292. Neri R, Simone B, Iacopetti V, Iacopetti G, Pepe P, d'Ascanio A, et al. Cancer-associated myositis: a 35-year retrospective study of a monocentric cohort. Rheumatology International. 2014;34(4):565-9.
293. Park JK, Yang JA, Ahn EY, Chang SH, Song YW, Curtis JR, et al. Survival rates of cancer patients with and without rheumatic disease: a retrospective cohort analysis. BMC Cancer. 2016;16.
294. Kaneko Y, Nunokawa T, Taniguchi Y, Yamaguchi Y, Gono T, Masui K, et al. Clinical characteristics of cancer-associated myositis complicated by interstitial lung disease: a large-scale multicentre cohort study. Rheumatology (Oxford). 2020;59(1):112-9.
295. Ng KP, Smith CR, Isenberg DA. Muscular dystrophy mimicking refractory idiopathic inflammatory myositis: a trio of cases. Rheumatology (Oxford). 2007;46(10):1618-9.
296. Nguyen K, Bassez G, Krahn M, Bernard R, Laforet P, Labelle V, et al. Phenotypic study in 40 patients with dysferlin gene mutations: high frequency of atypical phenotypes. Arch Neurol. 2007;64(8):1176-82.
297. Griger Z, Nagy-Vincze M, Bodoki L, Gherardi RK, Danko K, Hortobagyi T. Late onset dysferlinopathy mimicking treatment resistant polymyositis. Joint Bone Spine. 2016;83(3):355-6.
298. Camelo-Filho AE, Martins MRA, Souza JLB, Mauricio RB, Nobrega PR. Pompe disease misdiagnosed as polymyositis. Pract Neurol. 2023.
299. Rojas-Villarraga A, Amaya-Amaya J, Rodriguez-Rodriguez A, Mantilla RD, Anaya JM. Introducing polyautoimmunity: secondary autoimmune diseases no longer exist. Autoimmune Dis. 2012;2012:254319.
300. Danielsson O, Lindvall B, Hallert C, Vrethem M, Dahle C. Increased prevalence of celiac disease in idiopathic inflammatory myopathies. Brain Behav. 2017;7(10):e00803.
301. Tseng CC, Chang SJ, Liao WT, Chan YT, Tsai WC, Ou TT, et al. Increased Cumulative Incidence of Dermatomyositis in Ulcerative Colitis: a Nationwide Cohort Study. Sci Rep. 2016;6:28175.

302. Fang F, Sveinsson O, Thormar G, Granqvist M, Askling J, Lundberg IE, et al. The autoimmune spectrum of myasthenia gravis: a Swedish population-based study. J Intern Med. 2015;277(5):594-604.
303. Hemminki K, Li X, Sundquist J, Sundquist K. The epidemiology of Graves' disease: evidence of a genetic and an environmental contribution. J Autoimmun. 2010;34(3):J307-13.
304. Kuo CF, Grainge MJ, Valdes AM, See LC, Luo SF, Yu KH, et al. Familial Aggregation of Systemic Lupus Erythematosus and Coaggregation of Autoimmune Diseases in Affected Families. JAMA Intern Med. 2015;175(9):1518-26.
305. Kuo CF, Luo SF, Yu KH, See LC, Zhang W, Doherty M. Familial risk of systemic sclerosis and co-aggregation of autoimmune diseases in affected families. Arthritis Res Ther. 2016;18(1):231.
306. Kuo CF, Grainge MJ, Valdes AM, See LC, Yu KH, Shaw SWS, et al. Familial aggregation of rheumatoid arthritis and co-aggregation of autoimmune diseases in affected families: a nationwide population-based study. Rheumatology (Oxford). 2017;56(6):928-33.
307. Hemminki K, Li X, Sundquist J, Sundquist K. Familial associations of rheumatoid arthritis with autoimmune diseases and related conditions. Arthritis Rheum. 2009;60(3):661-8.
308. Arora-Singh RK, Assassi S, del Junco DJ, Arnett FC, Perry M, Irfan U, et al. Autoimmune diseases and autoantibodies in the first degree relatives of patients with systemic sclerosis. J Autoimmun. 2010;35(1):52-7.
309. Hemminki K, Li X, Sundquist K, Sundquist J. Familial association of inflammatory bowel diseases with other autoimmune and related diseases. Am J Gastroenterol. 2010;105(1):139-47.
310. Kuo CF, Grainge MJ, Valdes AM, See LC, Luo SF, Yu KH, et al. Familial Risk of Sjogren's Syndrome and Co-aggregation of Autoimmune Diseases in Affected Families: A Nationwide Population Study. Arthritis Rheumatol. 2015;67(7):1904-12.
311. Nielsen NM, Frisch M, Rostgaard K, Wohlfahrt J, Hjalgrim H, Koch-Henriksen N, et al. Autoimmune diseases in patients with multiple sclerosis and their first-degree relatives: a nationwide cohort study in Denmark. Mult Scler. 2008;14(6):823-9.
312. Hemminki K, Huang WQ, Sundquist J, Sundquist K, Ji JG. Autoimmune diseases and hematological malignancies: Exploring the underlying mechanisms from epidemiological evidence. Semin Cancer Biol. 2020;64:114-21.
313. Anderson LA, Gadalla S, Morton LM, Landgren O, Pfeiffer R, Warren JL, et al. Population-based study of autoimmune conditions and the risk of specific lymphoid malignancies. Int J Cancer. 2009;125(2):398-405.
314. Smedby KE, Ponzoni M. The aetiology of B-cell lymphoid malignancies with a focus on chronic inflammation and infections. J Intern Med. 2017;282(5):360-70.
315. Rosenquist R, Bea S, Du MQ, Nadel B, Pan-Hammarstrom Q. Genetic landscape and deregulated pathways in B-cell lymphoid malignancies. J Intern Med. 2017;282(5):371-94.
316. Berndt SI, Vijai J, Benavente Y, Camp NJ, Nieters A, Wang Z, et al. Distinct germline genetic susceptibility profiles identified for common non-Hodgkin lymphoma subtypes. Leukemia. 2022;36(12):2835-44.
317. Laugesen K, Ludvigsson JF, Schmidt M, Gissler M, Valdimarsdottir UA, Lunde A, et al. Nordic Health Registry-Based Research: A Review of Health Care Systems and Key Registries. Clin Epidemiol. 2021;13:533-54.
318. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.
319. Barlow L, Westergren K, Holmberg L, Talbäck M. The completeness of the Swedish Cancer Register – a sample survey for year 1998. Acta Oncol. 2009;48(1):27-33.

320. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. Eur J Epidemiol. 2017;32(9):765-73.
321. Johansson LA, Björkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. J Clin Epidemiol. 2009;62(11):1202-9.
322. Ludvigsson JF, Almqvist C, Bonamy AK, Ljung R, Michaelsson K, Neovius M, et al. Registers of the Swedish total population and their use in medical research. Eur J Epidemiol. 2016;31(2):125-36.
323. Ekbom A. The Swedish Multi-generation Register. Methods Mol Biol. 2011;675:215-20.
324. Dahlen T, Zhao J, Magnusson PKE, Pawitan Y, Lavrod J, Edgren G. The frequency of misattributed paternity in Sweden is low and decreasing: A nationwide cohort study. J Intern Med. 2022;291(1):95-100.
325. Wettermark B, Hammar N, Fored CM, Leimanis A, Otterblad Olausson P, Bergman U, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. Pharmacoepidemiol Drug Saf. 2007;16(7):726-35.
326. Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature. 2021;590(7845):290-9.
327. Din L, Sheikh M, Kosaraju N, Smedby KE, Bernatsky S, Berndt SI, et al. Genetic overlap between autoimmune diseases and non-Hodgkin lymphoma subtypes. Genet Epidemiol. 2019;43(7):844-63.
328. Miao J, Brismar K, Nyren O, Ugarph-Morawski A, Ye W. Elevated hip fracture risk in type 1 diabetic patients: a population-based cohort study in Sweden. Diabetes Care. 2005;28(12):2850-5.
329. Smedby KE, Akerman M, Hildebrand H, Glimelius B, Ekbom A, Askling J. Malignant lymphomas in coeliac disease: evidence of increased risks for lymphoma types other than enteropathy-type T cell lymphoma. Gut. 2005;54(1):54-9.
330. Neovius M, Simard JF, Askling J, group As. Nationwide prevalence of rheumatoid arthritis and penetration of disease-modifying drugs in Sweden. Ann Rheum Dis. 2011;70(4):624-9.
331. Arkema EV, Jonsen A, Ronnblom L, Svenungsson E, Sjowall C, Simard JF. Case definitions in Swedish register data to identify systemic lupus erythematosus. BMJ Open. 2016;6(1):e007769.
332. Jakobsson GL, Sternegard E, Olen O, Myrelid P, Ljung R, Strid H, et al. Validating inflammatory bowel disease (IBD) in the Swedish National Patient Register and the Swedish Quality Register for IBD (SWIBREG). Scand J Gastroenterol. 2017;52(2):216-21.
333. Murley C, Friberg E, Hillert J, Alexanderson K, Yang F. Validation of multiple sclerosis diagnoses in the Swedish National Patient Register. Eur J Epidemiol. 2019;34(12):1161-9.
334. Andreasson K, Saxne T, Bergknut C, Hesselstrand R, Englund M. Prevalence and incidence of systemic sclerosis in southern Sweden: population-based data with case ascertainment using the 1980 ARA criteria and the proposed ACR-EULAR classification criteria. Ann Rheum Dis. 2014;73(10):1788-92.
335. Vasaitis L, Nordmark G, Theander E, Backlin C, Smedby KE, Askling J, et al. Population-based study of patients with primary Sjogren's syndrome and lymphoma: lymphoma subtypes, clinical characteristics, and gender differences. Scand J Rheumatol. 2020;49(3):225-32.

336. Hudson JI, Javaras KN, Laird NM, VanderWeele TJ, Pope HG, Jr., Hernan MA. A structural approach to the familial coaggregation of disorders. Epidemiology. 2008;19(3):431-9.
337. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999;10(1):37-48.
338. VanderWeele TJ, Hernan MA, Robins JM. Causal directed acyclic graphs and the direction of unmeasured confounding bias. Epidemiology. 2008;19(5):720-8.
339. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. Int J Epidemiol. 2010;39(2):417-20.
340. Frisell T, Holmqvist M, Kallberg H, Klareskog L, Alfredsson L, Askling J. Familial risks and heritability of rheumatoid arthritis: role of rheumatoid factor/anti-citrullinated protein antibody status, number and type of affected relatives, sex, and age. Arthritis Rheum. 2013;65(11):2773-82.
341. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300.
342. Tam V, Patel N, Turcotte M, Bosse Y, Pare G, Meyre D. Benefits and limitations of genome-wide association studies. Nat Rev Genet. 2019;20(8):467-84.
343. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019;47(D1):D1005-D12.
344. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015;47(11):1236-41.
345. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet. 2015;47(11):1228-35.
346. Werme J, van der Sluis S, Posthuma D, de Leeuw CA. An integrated framework for local genetic correlation analysis. Nat Genet. 2022;54(3):274-82.
347. Lyon MS, Andrews SJ, Elsworth B, Gaunt TR, Hemani G, Marcora E. The variant call format provides efficient and robust storage of GWAS summary statistics. Genome Biol. 2021;22(1):32.
348. Bioconductor Package Maintainer. liftOver: Changing genomic coordinate systems with rtracklayer::liftOver. R package version 1.24.0; 2023.
349. Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
350. Murphy AE, Schilder BM, Skene NG. MungeSumstats: a Bioconductor package for the standardization and quality control of many GWAS summary statistics. Bioinformatics. 2021;37(23):4593-6.
351. Alan Murphy BSaNS. MungeSumstats: Standardise summary statistics from GWAS. R package version 1.2.4; 2022.
352. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47(3):291-5.
353. Leeuw JWaCd. LAVA: Local Analysis of [co]Variant Association. R package version 0.1.0; 2023.
354. Andreassen OA, Thompson WK, Schork AJ, Ripke S, Mattingsdal M, Kelsoe JR, et al. Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. PLoS Genet. 2013;9(4):e1003455.

355. Andersen PK, Esbjerg S, Sorensen TI. Multi-state models for bleeding episodes and mortality in liver cirrhosis. Stat Med. 2000;19(4):587-99.
356. Andersen PK, Keiding N. Multi-state models for event history analysis. Stat Methods Med Res. 2002;11(2):91-115.
357. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med. 2002;21(15):2175-97.
358. Hinchliffe SR, Lambert PC. Flexible parametric modelling of cause-specific hazards to estimate cumulative incidence functions. BMC Med Res Methodol. 2013;13:13.
359. Crowther MJ, Lambert PC. Parametric multistate survival models: Flexible modelling allowing transition-specific distributions with application to estimating clinically useful measures of effect differences. Stat Med. 2017;36(29):4719-42.
360. Hopper JL, Bishop DT, Easton DF. Population-based family studies in genetic epidemiology. Lancet. 2005;366(9494):1397-406.
361. Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. Nat Rev Genet. 2013;14(2):139-49.
362. Garber JE, Offit K. Hereditary cancer predisposition syndromes. J Clin Oncol. 2005;23(2):276-92.
363. Kardes S, Gupta L, Aggarwal R. Cancer and myositis: Who, when, and how to screen. Best Pract Res Clin Rheumatol. 2022;36(2):101771.
364. Hu S, Zhou D, Wu Y, Zhao Y, Wang S, Han B, et al. Autoimmune disease-associated non-Hodgkin's lymphoma-a large retrospective study from China. Ann Hematol. 2019;98(2):445-55.
365. Berndt SI, Skibola CF, Joseph V, Camp NJ, Nieters A, Wang Z, et al. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. Nature genetics. 2013;45(8):868-76.
366. Sampson JN, Wheeler WA, Yeager M, Panagiotou O, Wang Z, Berndt SI, et al. Analysis of heritability and shared heritability based on genome-wide association studies for 13 cancer types. Journal of the National Cancer Institute. 2015;107(12):djv279.
367. Bergman D, King J, Lebwohl B, Clements MS, Roelstraete B, Kaplan GG, et al. Two waves of coeliac disease incidence in Sweden: a nationwide population-based cohort study from 1990 to 2015. Gut. 2021.