TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

DIAGNOSTIC CHALLENGES IN RHEUMATOLOGY – IMPROVING VALIDITY OF REGISTRY-BASED STUDIES

Johanna Paltta

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1819 | MEDICA – ODONTOLOGICA | TURKU 2024

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

DIAGNOSTIC CHALLENGES IN RHEUMATOLOGY – IMPROVING VALIDITY OF REGISTRY-BASED STUDIES

Johanna Paltta

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. D OSA – TOM. 1819 | MEDICA – ODONTOLOGICA | TURKU 2024

University of Turku

Faculty of Medicine
Department of Internal Medicine
Doctoral Programme in Clinical Research

Supervised by

Adjunct Professor Laura Pirilä, MD, PhD
Centre for Rheumatology and Clinical Immunology
Turku University Hospital and
University of Turku
Turku, Finland

Adjunct Professor Antti Palomäki, MD, PhD
Centre for Rheumatology and Clinical Immunology
Turku University Hospital and
University of Turku
Turku, Finland

Reviewed by

Adjunct Professor Kari Puolakka, MD, PhD
Terveystalo Healthcare
Turku and Lappeenranta, Finland

Adjunct Professor Paula Vähäsalo, MD, PhD
Research Unit of Clinical Medicine
University of Oulu and
Department of Children and Adolescents
Oulu University Hospital
Oulu, Finland

Opponent

Adjunct Professor Dan Nordström, MD, PhD
Departments of Medicine and Rheumatology
Helsinki University Hospital and
University of Helsinki
Helsinki, Finland

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

Cover Image: Roosa Ruuskanen, Cover Design: Veera Ruuskanen

ISBN 978-951-29-9875-3 (PRINT)
ISBN 978-951-29-9876-0 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)
Painosalama, Turku, Finland 2024

To my family

UNIVERSITY OF TURKU
Faculty of Medicine
Department of Internal Medicine
JOHANNA PALTTA: Diagnostic Challenges in Rheumatology – Improving Validity of Registry-Based Studies
Doctoral Dissertation, 133 pp.
Doctoral Programme in Clinical Research
August 2024

ABSTRACT

Finnish health care registers are frequently used for research, but little attention has been paid on the validity of diagnoses in these registers, especially in the field of rheumatology. This dissertation investigated the validity of rheumatoid arthritis (RA) biobank diagnoses that were supplemented with data from the Finnish Care Register for Health Care and the Drug Reimbursement Register. Additionally, this dissertation evaluated the validity of hospital registry diagnoses of systemic sclerosis (SSc) and polymyalgia rheumatica (PMR). The validity of the diagnoses was evaluated retrospectively by a thorough manual review of all data in the medical records from the follow-up period of patients from five university and central hospitals in Finland. The goal was to understand the strengths and limitations of Finnish registry data in the field of rheumatology and how to best use this data for rheumatological research. Additionally, we examined the most common reasons and risk factors for misdiagnosis, especially for PMR.

The validity of an RA diagnosis could be greatly improved by including data on number of visits and medication reimbursement rights. By utilising this additional data, validity was high, particularly for seropositive RA. On the contrary, the validity of a single RA diagnosis was only moderate, especially for seronegative RA.

The registry diagnosis of SSc was moderately accurate, although the validity clearly improved when the diagnosis was made in the rheumatology department. The validity of a more specific diagnosis of limited cutaneous SSc, which is the most common SSc type in Finland, was high even for diagnoses made in any departments.

Among patients diagnosed with PMR at a university hospital, a third of the diagnoses were found to be incorrect following a more comprehensive assessment and clinical follow-up. The risk of misdiagnosis was particularly high in patients with atypical disease patterns, inadequate response to medication, and those not meeting classification criteria. The most common conditions that mimicked PMR were other types of inflammatory arthritis, noninflammatory musculoskeletal conditions, infectious diseases and malignancy.

Combining data from different sources, especially information on special reimbursement for medication, markedly improves the validity of registry diagnoses. Careful consideration of differential diagnosis is essential when diagnosing PMR to avoid misdiagnosis.

KEYWORDS: Rheumatoid arthritis, systemic sclerosis, polymyalgia rheumatica, registry study, validity of a diagnosis, differential diagnosis, study design

TURUN YLIOPISTO
Lääketieteellinen tiedekunta
Sisätautioppi
JOHANNA PALTTA: Diagnostisia haasteita reumatologiassa – rekisteritutkimusten validiteetin parantaminen
Väitöskirja, 133 s.
Turun kliininen tohtoriohjelma
Elokuu 2024

TIIVISTELMÄ

Suomalaisia terveydenhuollon rekistereitä käytetään usein tutkimuksissa, mutta rekisteridiagnoosien paikkansapitävyyttä eli validiteettia on ylipäätään selvitetty melko vähän, ja tieto reumatologisten diagnoosien validiteetista on vielä niukempaa. Väitöskirjassa tarkastettiin yksityiskohtaisten seuranta-ajan sairaskertomustietojen perusteella biopankkien hoitoilmoitusrekisteri- ja lääkerekisteritiedoilla täydennettyjen nivelreumadiagnoosien sekä sairaalarekisterien systeemisen skleroosin ja polymyalgia rheumatican diagnoosien validiteettia retrospektiivisesti. Tavoitteena oli selvittää suomalaisen rekisteriaineiston vahvuudet ja rajoitukset reumatologian alalla, sekä miten tätä aineistoa voidaan parhaiten hyödyntää reumatologisissa tutkimuksissa. Lisäksi tarkastelimme virheellisen diagnoosin yleisimpiä syitä ja riskitekijöitä, erityisesti polymyalgia rheumatican osalta.

Nivelreumadiagnoosin validiteettia voitiin selvästi parantaa ottamalla huomioon tiedot käyntien määrästä ja lääkkeiden erityiskorvattavuusoikeuksista. Näitä lisätietoja hyödyntämällä diagnoosin validiteetti oli hyvä, erityisesti tutkittaessa seropositiivista nivelreumaa. Yhdellä käyntikerralla asetetun nivelreumadiagnoosin validiteetti oli muuten ainoastaan kohtalainen, etenkin kun kyseessä oli seronegatiivinen nivelreuma.

Systeemisen skleroosin rekisteridiagnoosien paikkansapitävyys oli kohtalaisen korkea, ja validiteetti oli vielä selvästi parempi, jos diagnoosi oli tehty reumatologian yksikössä. Suomessa yleisimmän systeemisen skleroosin alatyypin, eli rajoittuneen systeemisen skleroosin, diagnoosit olivat luotettavia silloinkin, kun diagnoosi oli tehty millä tahansa erikoisalalla.

Yliopistosairaalassa polymyalgia rheumatica –diagnoosin saaneista potilaista kolmasosalla diagnoosi osoittautui kattavamman arvioinnin ja kliinisen seurannan jälkeen virheelliseksi. Väärän diagnoosin riski oli erityisen suuri potilailla, joilla taudinkuva oli epätyypillinen, joilla vaste lääkitykseen oli riittämätön ja joilla luokittelukriteerit eivät täyttyneet. Yleisimpiä alun perin polymyalgiaksi väärin diagnosoituja sairauksia olivat muut niveltulehdussairaudet, ei-tulehdukselliset tuki- ja liikuntaelinsairaudet, infektiosairaudet ja pahanlaatuiset sairaudet.

Rekisteridiagnoosien validiteettia voidaan huomattavasti parantaa yhdistämällä tietoja eri lähteistä, ja erityisesti tiedot lääkkeiden erityiskorvattavuusoikeuksista ovat tärkeitä. Erotusdiagnoosien huolellinen tarkastelu on olennaista polymyalgia rheumatican diagnosoinnissa, jotta vältytään vääriltä diagnooseilta.

AVAINSANAT: Reumatologia, nivelreuma, systeeminen skleroosi, polymyalgia rheumatica, rekisteritutkimus, diagnoosin validiteetti, erotusdiagnostiikka

Table of Contents

Abbreviations

ACE	Angiotensin Converting Enzyme
ACPA	Anti-Citrullinated Protein Antibody
ACR	American College of Rheumatology
ANA	Antinuclear Antibody
ASAS	Assessment of Spondyloarthritis International Society
ATC	Anatomical Therapeutic Chemical
axSpA	Axial Spondyloarthritis
bDMARDs	Biological Disease-Modifying Anti-Rheumatic Drugs
CENP	Anti-Centromeric Protein Antibody
CI	Confidence Interval
CT	Computed Tomography
CTDs	Connective Tissue Diseases
CRHC	Care Register for Health Care
CRP	C-Reactive Protein
DANBIO	Danish Registry for Biologic Therapies in Rheumatology
dcSSc	Diffuse Cutaneous Systemic Sclerosis
DMARDs	Disease-Modifying Anti-Rheumatic Drugs
DNPR	Danish National Patient Register
DRG	Diagnosis Related Group
ESR	Erythrocyte Sedimentation Rate
EULAR	European League Against Rheumatism – The European Alliance of Associations for Rheumatology
FHDR	Finnish Hospital Discharge Register
FIMEA	Finnish Medicines Agency
FINDATA	Finnish Social and Health Data Permit Authority
GC	Glucocorticoids
GCA	Giant Cell Arteritis
HILA	Pharmaceuticals Pricing Board (Hintalautakunta)
HIV	Human Immunodeficiency Virus
HLA-B27	Human Leukocyte Antigen B27
HRCT	High Resolution Computed Tomography
ICD	International Classification of Diseases
ICD-10	International Classification of Diseases – 10th Revision

ILD	Interstitial Lung Disease
IQR	Interquartile Range
JAK	Janus Kinase
KELA	Social Insurance Institution of Finland (Kansaneläkelaitos)
lcSSc	Limited Cutaneous Systemic Sclerosis
MRI	Magnetic Resonance Imaging
mRSS	Modified Rodnan Skin Score
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
NLR	Negative Likelihood Ratio
NPV	Negative Predictive Value
NVC	Nailfold Videocapillaroscopy
OR	Odds Ratio
PAH	Pulmonary Arterial Hypertension
PET	Positron Emission Tomography
PLR	Positive Likelihood Ratio
PMR	Polymyalgia Rheumatica
PPV	Positive Predictive Value
RA	Rheumatoid Arthritis
RF	Rheumatoid Factor
RNP	Anti-U1 Ribonucleoprotein Antibody
SCL-70	Anti-DNA Topoisomerase 1 Antibody (Anti-Scleroderma 70 kDa Nuclear Protein Antibody)
SLE	Systemic Lupus Erythematosus
SSc	Systemic Sclerosis
STM	Ministry of Social Affairs and Health in Finland (Sosiaali- ja terveysministeriö)
THL	Finnish Institute for Health and Welfare (Terveyden ja Hyvinvoinnin Laitos)
ULN	Upper Limit of Normal
US	Ultrasound
WHO	World Health Organization

List of Original Publications

This dissertation is based on the following original publications, which are referred to in the text by their Roman numerals:

I Paltta J., Heikkilä H-K., Pirilä L., Eklund K K., Huhtakangas J., Isomäki P., Kaipiainen-Seppänen O., Kristiansson K., Havulinna A S., Sokka-Isler T., Palomäki A. for the FinnGen investigators. The validity of rheumatoid arthritis diagnoses in Finnish biobanks. *Scand J Rheumatol*, 2023;52(1):1–9.

II Paltta J., Kortelainen S., Käyrä M., Pirilä L., Huhtakangas J., Palomäki A. The validity of systemic sclerosis diagnoses in two university hospitals in Finland. *Scand J Rheumatol*, 2023;52(1):84–87.

III Paltta J., Suuronen S., Pirilä L., Palomäki A. Differential diagnostics of polymyalgia rheumatica in a university hospital in Finland. *Scand J Rheumatol*, 2023;52(6):689–695.

1 Introduction

Diagnosis is the foundation of medicine. It is a simple way to gather a lot of complex information under one concept. The doctor needs a diagnosis to organise all the patient's symptoms and findings into a manageable unit, consult with their colleagues more precisely, prescribe the right treatment and organise proper follow-up. The patient needs a diagnosis to understand what is wrong with them, tell their family about it, and get the financial support granted by the government in the form of sick leave and medication reimbursements, for example. (Jutel 2009)

However, a diagnosis is not an unchanging entity, and may change over time with new symptoms and findings emerging during follow-up. The progress of medicine may also change the definitions of diagnoses because they are ultimately a mutual agreement based on the current knowledge of the pathophysiology of diseases (Conrad et al. 2010; Pietikäinen et al. 2017). Some diagnoses have a long history, some old diagnoses have been superseded by subsequent developments in medical knowledge, and new diagnoses emerge to define previously unknown entities.

Historically, diseases have been categorised based mainly on the patient's symptoms and possible clinical status findings, but in the future, the classification of diseases will rely more on the actual pathogenesis of a disease, creating new diagnoses and changing the definitions of the previous ones (Petersen 2021; König et al. 2017). A deeper understanding of the origins of diseases is also seen in the development of precision medicine, with the potential to target the disease-causing abnormalities in the body even more specifically (Guthridge et al. 2022).

The scientific community and health care industry also require accurate diagnoses in order to create uniform patient cohorts. If the aim is to study the pathogenesis, treatment or prognosis of diseases, it is important that the diagnosis is clearly defined so that the research questions can be addressed for the right patients. If the patients in the study do not actually have the presumed disease, the results of the study will become inaccurate and unreliable.

In the field of rheumatology, many syndromes interconnect and overlap. Only a few diagnostic criteria exist, and the diagnosis is almost never made based on a single test, but rather as a diagnostic pattern of symptoms, findings, and laboratory or

imaging results. Many classification criteria are available, which are often used in clinical practice to help with a diagnosis, but they have been developed for research purposes rather than as diagnostic tools (Aggarwal et al. 2015; Porter et al. 2020). The diversity of diagnostic processes creates challenges in assessing the reliability of individual diagnoses in patient records.

The validity of register diagnoses in Finland has been previously studied in certain disease groups, and it has been shown to vary from moderate to high with the accuracy between 75% and 99% in common diagnoses, but lower in rare diseases (Sund 2012; Vuori et al. 2019). Clearly fewer studies have been conducted on the accuracy of diagnostic information in Finnish biobank patients (Haverinen et al. 2020; Vesterinen et al. 2020) and, to our knowledge, there are no validation studies of rheumatological diagnoses in Finnish biobank patients or in the Finnish Care Register for Health Care (CRHC).

Rheumatoid arthritis (RA) is a chronic autoimmune and inflammatory disorder primarily targeting the joints of a patient. RA most often affects smaller joints of the hands, wrists and feet, usually in a symmetric and polyarthritic pattern. Untreated RA causes painful swelling and stiffness of the joints and can eventually cause structural damage, bone erosion and joint deformity. The diagnosis of RA cannot be confirmed by a single test; instead, a physician makes the diagnosis based on the clinical presentation of the patient. Factors influencing this decision include the amount and location of tender and swollen joints, duration of symptoms, possible elevated inflammatory markers, possible presence of rheumatoid factor (RF) and anti-citrullinated peptide antibodies (ACPA), and possible erosive changes in radiograph images. (Sparks 2019; Scherer et al. 2020; Aletaha et al. 2018)

Systemic sclerosis (SSc) is a rare autoimmune disorder causing inflammation and fibrotic scarring in the connective tissue of the skin, joints, internal organs and blood vessels. There are different subsets of SSc and diagnosis is made by taking into account comprehensive clinical examinations as well as imaging and laboratory analyses. Common findings in SSc are thickening of the skin on the hands, lesions in the fingertips, telangiectasis, abnormal findings in nail fold capillaries, Raynaud's phenomenon, pulmonary arterial hypertension or interstitial lung disease, and the presence of disease-specific autoantibodies. The progression of SSc can range from mild to life-threatening, and a timely diagnosis is important in order to prevent permanent damage or even death. (Volkmann 2023; Denton et al. 2017)

Polymyalgia rheumatica (PMR) is a common inflammatory disease that causes muscle pain and morning stiffness, especially in the neck, shoulders, upper arms, and pelvic girdle. Other common symptoms include fatigue, weight loss, and fever. PMR usually and almost exclusively affects patients over the age of 50, with incidence increasing progressively with age. The diagnosis of PMR is based on symptoms, laboratory evidence of an acute-phase reaction and possible imaging. The

mainstay of PMR treatment is glucocorticoids, and a fast, sufficient response to glucocorticoids is a characteristic of the disease and is sometimes used as a factor confirming the diagnosis. The differential diagnostics of PMR is challenging because many other common conditions may have similar presentation, and it is essential for physicians to have a broad clinical view of possible mimics of PMR. (González-Gay et al. 2017; Buttgereit et al. 2016; Lundberg et al. 2022; Espígol-Frigolé et al. 2023)

The goal of this dissertation was to understand the strengths and limitations of Finnish registry data in the field of rheumatology and how to best utilise this data for rheumatological research. For PMR, we also sought to determine the most common diseases misdiagnosed as PMR and to define the most important reasons and risk factors for misdiagnosis.

2 Review of the Literature

2.1 What is a diagnosis?

Making a diagnosis is one of the most important skills of a doctor and the basis of modern medicine. The word diagnosis comes from the Greek words dia (through) and gnōsis (knowledge) (Dorland 2012). Diagnosis means either the determination of the nature of a disease according to a certain classification scheme, or the process of distinguishing one disease from another (Jutel 2009; Dorland 2012).

By making a diagnosis, the doctor gathers and organises a lot of complex information under one concept. Relevant data is gathered from the patient's lifelong history starting with family history and possible predisposing factors and taking into account the symptoms and ailments of the patient, findings from the clinical examination, any abnormalities in blood tests, imaging or other examinations, results of possible treatment trials, and many other things. All this is summed up into a single concept for the doctor to focus on, possibly consult other specialists, and search for additional information from literature and clinical guidelines. (Jutel 2009)

For a patient, it is important to identify their health problem. If no diagnosis can be made for the patient's symptoms, it can cause great uncertainty and lead to physiological and psychological consequences for the patient (Kornelsen et al. 2016). With a diagnosis, the patient can receive information on what is wrong with him or her, what are the treatment options and what can be expected in the future. They can educate their family and friends about the disease. With a diagnosis, a patient is able to connect with other patients and get important peer support in support groups, patient organisations and even in social media. In a sense, getting a diagnosis gives the patient the permission to be ill.

A diagnosis is also a matter of money. Many illnesses impair the patient's functional capacity and ability to work, cause absence from work and even disability, thus affecting his or her livelihood. In addition, expenses related to medications, other therapies, examinations and visits increase. These challenges threaten the patient's financial stability. The diagnosis provides an opportunity for the patient to receive support from society and their employer. In Finland, the national health insurance system grants all patients with certain diagnoses of chronic and severe diseases an entitlement to special reimbursement for the costs of medications (Kela.

Reimbursements for medicine expenses 2023). A diagnosis is also needed to be eligible for long-term paid sick leave, rehabilitation, disability allowance, pension, insurance reimbursement and other kinds of financial support (Kela. Life situations 2023).

A diagnosis is an administrative indicator used by a government to collect information about citizens, and many registers and statistics classify people and resource use based on diagnoses (THL. Statistics by topic 2023; Kela. Research and statistics 2023). This information is used, among other things, to monitor the health status of the population and to analyse changes in and possible causes of the population's health status. Based on this information, measures to improve health can be planned and implemented, and the limited resources of society can be allocated as efficiently as possible.

With the development of medicine, it has become possible to treat a disease and its causes, rather than just relieve the symptoms (Jessop et al. 2016; Koutsouris 2017). A correct diagnosis is needed to prescribe the right treatment and proper follow-up. An incorrect diagnosis may lead to patients not getting the appropriate help for their problems, the lack of proper treatment may cause pain and suffering or even worsen the prognosis, and the wrong kind of treatment may cause more harm than good. (Thammasitboon et al. 2013; Graber et al. 2005; Gunderson et al. 2020; Kuhn 2002)

Advancements in medical science may change the definitions of diagnoses because they are ultimately a mutual agreement on which set of symptoms and findings are considered to be a particular disease. Cultural and societal factors can also influence what is considered normal and abnormal symptoms, physical features or behaviour. (Jutel 2009; Conrad et al. 2010; Pietikäinen et al. 2017; Aronowitz 2001)

Some diagnoses have a long history, with examples including gout, which was first identified by the Egyptians in 2640 BC and more accurately described by Hippocrates in the fifth century BC (Nuki & Simkin 2006). On the contrary, some diagnoses, such as female hysteria, have been superseded by subsequent developments in medical knowledge (Tasca et al. 2018). As medical science develops, new diagnoses emerge to define previously unknown entities, and the definitions of old diagnoses may change.

Historically, diseases have mainly been categorised based on the patient’s symptoms and possible clinical findings, but in the future, the classification of diseases will rely more on the aetiological and pathophysiological aspects of the disease. In this new approach, the diagnosis aims to describe the root cause of the disease instead of looking at the end results. This change in perspective leads to shifting the existing boundaries of the diseases and creating new ones (Petersen 2021; König et al. 2017). The same trend of a deeper understanding of medical

science is also seen in the development of precision medicine, which targets the specific disease-causing abnormalities in the body (Guthridge et al. 2022).

Diagnosis is the basis for medical and pharmaceutical research and development with an increasing aspiration for evidence-based medicine (Djulbegovic et al. 2017). When designing a research study in order to study the pathogenesis, treatment or prognosis of diseases, it is important to consider how to verify that the diagnosis is correct so that the study questions can be applied to the appropriate patients and a uniform research cohort can be obtained. If the patients in the study do not actually have the presumed disease, the results of the study will become inaccurate and unreliable.

2.2 Diagnostic process – making a diagnosis

A diagnosis is a formal statement by a doctor that defines a patient's health problem in one concept. Due to the autonomy of the doctor, the doctor can make a diagnosis based on their own decision, but the decision is based on a pre-existing set of categories agreed upon by the medical profession. (Jutel 2009)

Forming a diagnosis is like solving a puzzle – combining a lot of data from different sources in order to create a uniform picture. The difficulty of diagnostics is related to the fact that not all the pieces available belong to this puzzle, some pieces are bigger than others, and some pieces are missing.

Solving a diagnostic puzzle and making a diagnosis is a process of gathering the most relevant information about a patient's symptoms and conditions, outlining possible differential diagnostic options and interpreting the best available information to determine the final diagnosis. This process usually requires a systematic approach as well as problem-solving and decision-making skills. However, with a more experienced physician, the process may change from a simple systematic review of differential diagnosis lists to a more automated process based on pattern recognition. (Renko et al. 2010)

So where does the doctor get the pieces of the diagnostic puzzle? It is an interactive process between the doctor, the patient, other health care professionals and even the patient's family. A big part of the diagnosis is based on the patient's history, so it is important for the doctor to gather a detailed history by letting the patient tell their story and asking supplementary questions. The history is interpreted and integrated with the results of a physical examination and different kinds of investigations such as blood tests, imaging procedures, and sometimes biopsy results. Sometimes this process leads directly to the final diagnosis, but often a working diagnosis or a hypothesis is formed first, and the accuracy is assessed later based on additional information collected during follow-up as well as results of a possible drug treatment trial. Even if at some point the diagnosis seems certain, it

can later become uncertain or even wrong if the follow-up reveals facts that are more indicative of another diagnosis. (Summerton 2004; Balogh et al. 2015)

2.2.1 International Classification of Diseases

The World Health Organization (WHO), which is an agency of the United Nations responsible for international public health, maintains an International Classification of Diseases (ICD) (WHO. International Classification of Diseases 2023). ICD contains a list of codes classifying different diseases, as well as various signs and symptoms, abnormal findings, and causes of injuries and illnesses. ICD-10 is the 10th revision of this list and is the revision in use in Finland since 1996. ICD-11, the 11th and latest revision, officially came into effect on 1 January 2022 and is being implemented in Finland during 2023–2026. The new ICD-11 diagnosis classification has a different structure and is broader compared to the previous ICD-10 classification. With more detailed codes and by combining codes and terms, it is possible to describe the patient's condition in more detail than before. ICD-11 is also built to be used completely electronically. (Harrison et al. 2021)

2.2.2 Diagnostic criteria and classification criteria

Some diseases have diagnostic criteria, which are a set of signs, symptoms, and tests for use in routine clinical care to make the diagnosis and guide the treatment of the patient. For example, a patient has diabetes if their fasting blood glucose has been more than 7 mmol/l in two separate measurements, or a patient has human immunodeficiency virus (HIV) infection if a test for HIV is positive. In the field of rheumatology, only a few diagnostic criteria exist, and the diagnosis is almost never made based on a single test, but rather based on a diagnostic pattern of symptoms, findings, and laboratory or imaging results. Many syndromes interconnect and overlap, and two different diseases can have similar findings and test results.

Classification criteria are a set of disease characteristics that define a group of patients that are a relatively homogeneous population having similar clinical disease features (Felson et al. 1995; Johnson et al. 2007). Classification criteria are often used in medical education to describe the disease in question and are also used regularly in clinical practice to help make the diagnosis (June et al. 2014). The fulfilment of classification criteria is often considered a requirement for a diagnosis. This may be useful especially in the beginning of the rheumatologist’s career when the clinical experience on the various features of the disease is still limited. But classification criteria are not meant for diagnostic purposes (June et al. 2014). They are meant to be used for research purposes to obtain well-defined and uniform patient cohorts in studies. This way, the results of these studies can be put into the context

of previous study results, and the results can be compared to the results of other studies because it is likely that the studies have examined similar kinds of patients who are suffering from the disease in question (Aggarwal et al. 2015; Porter et al. 2020; June et al. 2014).

Classification criteria are not designed to identify all the patients with the disease because the disease characteristics can vary quite a lot among different patients. In other words, classification criteria aim for high specificity in order to ensure that all the diagnosed patients do have the disease, which often comes at the cost of lower sensitivity, meaning not all patients with the disease meet the criteria. Diagnostic criteria, however, is different because diagnostic criteria aim for high sensitivity but with a lower specificity (Aggarwal et al. 2015; Porter et al. 2020; Felson et al. 1995; Johnson et al. 2007). It is possible that a patient with a rheumatic disease does not fulfil the classification criteria due to an early phase of the disease or an atypical presentation.

Classification criteria used in rheumatology have been developed at different times in various ways (Felson et al. 1995; Johnson et al. 2007). The criteria are often based on a consensus of expert opinions. For example, the 2010 rheumatoid arthritis classification criteria from the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) were made in three phases. First, factors from arthritis patients were identified that would prompt the physician to initiate methotrexate treatment. Second, these factors and their relative weights, as well other factors that might be clinically relevant, were refined using a series of patient case scenarios. Finally, the most important factors were assigned points, and the optimal cutoff point to a definite diagnosis of RA was decided (Aletaha et al. 2010).

In a review study from 2007, it was found that 50% of the criteria used in rheumatology at the time were based on expert opinion rather than patient data, and when patient data was used, the number of cases analysed ranged from 20 to 588 and the number of controls from 50 to 787 (Johnson et al. 2007).

In the development of classification criteria, differential diagnostics is usually done before the inclusion of patients into cohorts from which the classification criteria are derived and patients who have been included in the disease cohort have already been diagnosed (Funovits et al. 2010; Van Den Hoogen et al. 2013). The classification criteria usually state that the criteria are not to be applied to patients having some other disease that better explains their manifestations. If classification criteria are applied to unselected patients for whom the differential diagnostics has not been made, one patient can fulfil classification criteria for more than one disease.

Because the classification criteria are often made based on the preselected and prediagnosed patient material in rheumatology clinics, they may not work so well in other patient populations such as primary health care patients. It is possible that

symptoms and findings of many common ailments can be combined so that they fulfil the classification criteria. (June et al. 2014)

The genetic background of the population can also affect how easily the criteria are fulfilled (June et al. 2014). For example, in the Assessment of Spondyloarthritis International Society (ASAS) classification criteria for axial spondyloarthritis (axSpA) (Rudwaleit et al. 2009), human leukocyte antigen B27 (HLA-B27) positive genotype has a significant impact. If a patient under 45 years with chronic back pain is HLA-B27 positive, they fulfil the criteria with, for example, a good response to nonsteroidal anti-inflammatory drugs (NSAIDs) and psoriasis with no need for sacroiliitis in magnetic resonance imaging (MRI). In Finland, with 15% of the population being HLA-B27 positive (Jaakkola et al. 2006), this is much easier to fulfil than, for example, in Japan where HLA-B27 positivity is 0.4% (Kameda et al. 2021).

The advancement of medical science can also affect the diagnostics of diseases by developing new tests and imaging modalities that will require updating the classification criteria (June et al. 2014). For example, the diagnostics of rheumatoid arthritis (RA) changed with the introduction of anti-citrullinated protein antibody (ACPA), with its over 95% specificity and 70% sensitivity for RA (Schellekens et al. 2000). After this change, the ACPA serology was included in the new ACR/EULAR 2010 classification criteria for RA (Aletaha et al. 2010). Furthermore, the boundaries of diagnostic criteria may also undergo modification over time, in accordance with advancements in medical understanding of diseases and their clinical manifestations and underlying pathological processes.

So why do physicians still use the classification criteria as though they are diagnostic criteria instead of just developing diagnostic criteria? The answer lies in the complexity of rheumatic conditions. It is difficult to create a set of criteria that would be both sensitive and specific, account for the wide range of disease phenotypes and be reliable in populations with different ethnic backgrounds. For example, ACR has stated that it will only provide approval for classification criteria and will no longer consider funding or endorsement of diagnostic criteria (Aggarwal et al. 2015). It remains the responsibility and privilege of the rheumatologist to make the decision on the diagnosis based on the diverse symptoms and findings of the individual patient. The patient has the right to receive a diagnosis, and appropriate treatment should not be denied if some classification criteria are not yet met. If the more precise nature of the illness is not yet clear at the first appointment, an unspecified diagnosis e.g. unclassified arthritis, can be established first and refined later if follow-up reveals more detailed information.

2.2.3 Diagnostic error

Diagnostic error, which causes the patient to get an incorrect diagnosis, is common. It has been estimated that in the USA, most people will experience a diagnostic error in their lifetime (Balogh et al. 2015) and that 10–15% of all diagnoses are erroneous (Graber 2013). Not every diagnostic error is harmful to the patient, but in a systematic review and meta-analysis on the prevalence of harmful diagnostic errors in hospitalised adults, it was shown that a minimum of 0.7% of hospitalised adults suffer harm due to a diagnostic error (Gunderson et al. 2020). In a review on autopsy studies, it was estimated that in a hospital in the United States in 2000, 4% of the patients who died in the hospital had a diagnostic error that might have contributed to the patient's death (Shojania et al. 2003).

Diagnostic errors can be classified into three distinct categories according to the information that is obtained later. First, as a diagnosis that was unintentionally delayed even if sufficient information was available earlier. Second, as entirely wrong diagnosis, when another diagnosis was made before the correct one. Third, a diagnosis may be missed entirely, with no diagnosis ever being made. (Graber et al. 2005).

Diagnostic error usually happens as a result of an unfortunate coincidence of several factors, which can be divided into system-related errors, cognitive errors by the physician and errors due to other reasons (Graber et al. 2005; Balogh et al. 2015).

Cognitive reasons are the biggest group of reasons for diagnostic error. In a study by Graber et al. 2005, cognitive reasons were behind 74% of diagnostic errors. Cognitive flaws come in many forms, but the most common are due to errors in combining information from different sources. Data combining error can mean failure to collect all the relevant data, failure to prioritise the data and set the data into the right context, or to make conclusions that are not supported by the data. A big problem of data integration is the premature closure, which means that after an initial diagnosis is established, other reasonable diagnoses are no longer considered. This problem can happen together with anchoring, which means that if the physician has decided on the diagnosis, subsequent or contrasting information is not taken into consideration. Graber et al. 2005 showed that faulty or inadequate knowledge was an uncommon (4%) reason for diagnostic error, and it was often in the context of a rare disease. (Kuhn 2002; Balogh et al. 2015).

System flaws are the second biggest cause (65%) of diagnostic error (Graber et al. 2005). Most often these system-related reasons are due to organizational problems like faulty policies and procedures, inefficient or unclear processes, unclear division of responsibilities, teamwork difficulties, lack of communication and difficulties for the patient to access health care. Technical failures and equipment problems are rare. (Graber et al. 2005; Balogh et al. 2015)

Not all errors are caused by the physician or the system. Some may be related to the patient if they are uncooperative, unintentionally misleading, deceptive, or simply do not recognise that they have a problem or seek help for it. Diagnostic errors can happen even if there are no errors in the diagnostic process; for example, the disease can present atypically or be in such an early stage that it cannot realistically be detected at the time. (Thammasitboon et al. 2013; Graber et al. 2005; Kuhn 2002; Balogh et al. 2015)

It should be kept in mind that uncertainty is always present in medicine and it can be minimised, but not eliminated completely. Confirmatory examinations and tests cannot be continued indefinitely considering the inconvenience and possible harm caused to the patient, as well as the limited resources of the health care system. All errors in diagnosis cannot be prevented, but it is important to be aware of their existence.

2.3 What does validity mean?

The concept of validity refers to the degree of correspondence between a piece of information and the truth, or how close this piece of information is to real world. Validity can be classified into three principal categories: test validity, experimental validity and diagnostic validity.

2.3.1 Test validity

Validity and reliability are both concepts used in scientific research to describe the quality of the object in question. Validity refers to how well a method measures what it is supposed to measure. Reliability refers to consistency, meaning that the result does not change over time if the conditions remain the same. For example, a reliable measurement is something that gives the same result every time an object is measured if the object itself does not change. A reliable measurement can still have poor validity, meaning that the results are consistent and reproducible but not necessarily correct. Measurements with good validity usually have good reliability. (Hazra et al. 2017)

Validity of a recorded diagnosis refers to the probability of the recorded diagnosis being correct, which means it fits the generally accepted consensus on that diagnosis. For example, the perception of RA is different today than it was years ago, in particular in the light of the new classification criteria, and a diagnosis of RA made in the past may be wrong according to current consensus. Validity is often reported numerically either as a percentage (0–100%) or a proportion (0–1) of diagnosed individuals who have the disease. This proportion is called a positive predictive value (PPV). A negative predictive value (NPV) indicates how often a

person who is not thought to have a particular disease actually does not have the disease in question. (Hazra et al. 2017; Olliaro et al. 2021)

2.3.2 Experimental validity

Experimental validity describes the validity of the design of experimental research study, and it covers statistical conclusion validity as well as internal and external validity.

Statistical conclusion validity refers to the extent to which the conclusions about the variables are accurate. It is based on the use of appropriate sampling, right measurement techniques and the use of suitable statistical tests.

Internal validity refers to the extent to which evidence supports a claim about cause and effect in a particular study. External validity refers to the extent to which the results of a particular study are generalizable outside the specific circumstances of the study, for example in a different population.

2.3.3 Diagnostic validity

Diagnostic validity describes the ability of a given test or set of diagnostic criteria to distinguish between individuals who are afflicted with a particular disease and those who are not, or who are afflicted with a different disease. If diagnostic criteria have good diagnostic validity, the proportion of false positives and false negatives is low. This means that only individuals with the disease in question fulfil the criteria, and those with another disease or no disease at all will not fulfil the criteria.

2.4 Registry-based studies

Registry-based studies are studies that utilise data that is usually collected for other purposes. Registry studies typically combine register data from different authorities in a new way or complement the data with interview and survey data collected during the study. (Ludvigsson et al. 2015; Edwards et al. 2012; Hafström et al. 2019; Soussi et al. 2022)

There are many clear benefits to utilising registry data in studies. The data already exist, often with long follow-up periods, which often make registry studies quicker and cheaper to perform compared to studies recruiting patients from the clinical practice (Anderson et al. 2020). It is possible to identify a large number of patients from the registries, making it easier to identify statistically significant differences. The ability to search for subjects from a large number of patients is especially important when studying rare diseases or health events, or when there may be a long period between exposure and a health event. Patients dropping out of the

study during follow-up is also minimised in nationwide registries. Some studies such as drug exposure during pregnancy would not be possible to do due to ethical reasons. Real-world data from registries is more generalizable than data from clinical studies with tightly selected patient cohorts. Registry data also minimises different types of bias such as volunteer bias, selection bias, recall bias and the influence of the diagnostic process determined by the study. It is also possible to improve data quality by combining data from different registries. (Thygesen et al. 2014; Maret-Ouda et al. 2017; Rubinger et al. 2023)

2.4.1 Problems and error sources of registry-based studies

Registry studies are not without problems. Some of the error sources are presented in Table 1.

Table 1. Error sources of registry studies.

Error sources of registry studies
Registry data not generated for study purposes • Limited number of variables • Varying quality and coverage, differences among professionals and institutions • Different or differently coded variables in different registries • Missing data • Lack of confounder information
Administrative or economic circumstances and guidelines • Diagnosis recorded before it is medically confirmed • Diagnoses influence the funding allocated to hospitals on the basis of benchmarking data • Raised threshold for hospitalization and resource constraints may cause a disease to be diagnosed at a more advanced state
Large data sets • Statistically significant findings with no practical importance in real life • Data dredging
Left truncation
Poor validity of diagnoses

2.4.1.1 Data not generated for study purposes

The data in the registries are not generated by researchers to be used in registry studies, but rather by other professionals, such as clinical practitioners, to be used in patient care and for administrative purposes. This practice limits the number of possible variables to be used, and there may be differences among professionals and institutions in the generation of the data. The quality and coverage of the data can vary between different institutions if the coding process and administrative

guidelines are not the same, and validating the quality of registries can be difficult. When combining multiple registries, the same variables may not be found in all registries, and the variables may be coded differently. Because of these discrepancies, combining registry data requires deeper knowledge of the different variables in the registries, how and for what purpose the data was originally generated, and whether interpretation of the variables is the same. A clinician familiar with the routines of the clinic can have a very different view of the quality and interpretation of variables than a statistician would. Cleaning and merging the data also takes time and effort when combining data from different registries. (Thygesen et al. 2014; Maret-Ouda et al. 2017)

In registry-based studies, handling of missing data may be difficult. Important data might be unavailable or misclassified, and it may be hard to interpret the meaning of missing data; for example, does the lack of a recorded positive test result indicate a negative finding, or does it mean that the test was not performed at all? Low coverage of registry data, with only a part of patients registered, can potentially indicate a selective enrolment. Selective enrolment causes bias in the study conclusions. In registry intervention or safety studies, adverse events may be underreported compared to randomised controlled trials (RCT), because in RCTs the patients are actively encouraged to report adverse events. Confounder information is often missing from registries and is only available at irrelevant times or through proxy data. For instance, a patient's socio-economic status may be determined by registry data on level of education. (Thygesen et al. 2014; Rubinger et al. 2023)

2.4.1.2 The effect of economic and administrative factors on diagnoses

Nonmedical administrative or economic circumstances and guidelines may lead to a diagnosis being recorded before it is medically confirmed. For example, a suspicion of an illness is often set as a true diagnosis in hospital administrative records even if the follow-up proves this suspicion wrong. This situation can happen because the administrative information system requires a coded diagnosis from the physician on every contact with the patient, and the physician may feel that a symptom code alone is not sufficient. (Maret-Ouda et al. 2017)

The Diagnosis Related Group (DRG) system is a classification system where treatments and patient cases are placed into groups that are similar medically and in terms of costs based on diagnoses and other recorded information. In DRG, some diagnoses are more expensive than others, and a patient with many diagnoses of comorbidities is more expensive than a patient with no comorbidities. Many countries use the DRG system in the pricing and invoicing of medical care services, as well as to analyse and compare hospitals' operational, financial and productivity data – also known as benchmarking data. In Finland, this benchmarking data of

public hospitals contain information by common Nordic DRG system (NordDRG Full) patient groups (THL. Hospital Benchmarking 2023). In the Law on the Financing of Welfare Areas, the funding of welfare areas was mainly determined by need-based state funding, which in turn is partly based on the calculated DRG group costs from the hospital productivity statistics (Finlex. Act on the funding of welfare regions 2023; THL. Needs assessment of health and social services funding 2022). Because of this funding formula, physicians are incentivised to record several diagnoses of a patient's chronic comorbidities to ensure future funding for the hospital, even if they are not the reason for the visit. (Thygesen et al. 2014; Maret-Ouda et al. 2017; Böcking et al. 2005)

Societal and financial factors can also affect the results of prevalence and incidence studies. Economic problems or political efforts to increase outpatient care may reduce patient occupancy in hospitals, thus raising the threshold for hospitalization, and resource constraints can cause a disease to be diagnosed at a more advanced stage. Both of these situations may lead to a decrease in prevalence or incidence in some diseases. (Thygesen et al. 2014)

2.4.1.3 Problems from large amount of data

In registry studies with large data sets, it is possible to get statistically significant results from small differences that have no practical importance in real life. Due to the lack of confounder information together with the significant statistical power of big data sets, registry studies can be prone to over- or underestimation of the observed association. (Thygesen et al. 2014; Maret-Ouda et al. 2017; Rubinger et al. 2023; Olsen 2011)

The availability of huge amounts of data in large registries may influence a researcher into making different kinds of analyses from the data and then afterwards choosing the study topic based on the results that yielded the largest statistical impact, especially if the scientist's financial stability or reputation are dependent on frequent publication. However, this kind of data dredging is not a valid way to perform science because the study design should always be based on a hypothesis and not the other way around. However, it is useful for identifying hypotheses for future studies through explorative research, as long as basic scientific principles are followed. (Thygesen et al. 2014; Olsen 2011)

2.4.1.4 Missing data

Missing data is a common problem in registry studies. In some cases, data may be missing completely at random, which should not have a significant impact on the results. However, data can also be missing not at random, in which case the missing

values are related to some other unknown factor, potentially leading to bias in the results. When assessing the validity and completeness of register data, it is important to understand the type and extent of the missing data. Although the researcher cannot know with certainty why data are missing, it is still worthwhile to attempt to ascertain the reason for the missing data in order to determine how the missing data should be handled in the study.

2.4.1.5 Left truncation

The incidence and prevalence of diseases can also be difficult to differentiate at the start of the registry period, when old cases that were diagnosed years before the start of the register appear as new cases on the first year of the register – a phenomenon called left truncation. Left truncation can cause overly high incidence rates and overly low prevalence rates in the first years of the register, especially with diseases that rarely require contact with the hospital.

2.4.1.6 The effect of poor validity of registry data

If a registry study is based on a register with a poor validity of diagnoses, misclassification bias can cause the results of the study to become inaccurate and unreliable, because patients are assigned to the wrong study category (Pham et al. 2019). In a study designed to determine the effectiveness of a treatment for a specific disease, a potentially effective treatment may falsely appear to be ineffective if the patients selected in the study do not actually suffer from the disease in question. In incidence and prevalence studies, a large amount of false positive diagnoses in the population falsely increases the incidence or prevalence. In studies investigating the use of health care resources to treat a chronic disease, a large amount of false positive diagnoses can make the results erroneously small because too many healthy individuals were mixed in the study cohort. In studies investigating genetic background or cytokine profile of a disease, relevant findings are lost in the confusion caused by genetic and cytokine profiles of patients with erroneous diagnoses. (Benchimol et al. 2011)

2.4.2 Investigating the validity of register data

Before starting a registry study, it is important to find out the validity of the diagnoses of the subject under investigation in order to avoid the misclassification bias resulting from incorrect diagnoses (Benchimol et al. 2011). An international consortium has identified studies where medical records are reviewed to determine the validity of hospital abstract data, as one of the ten most highly ranked priorities

of methodological research (De Coster et al. 2006). The validity of register diagnoses can be investigated many different ways, and some of these ways might be more effective than others.

The completeness or validity of the register diagnoses can be assessed by comparing the register data with another register that is believed to be more accurate (Thygesen et al. 2014). One possible method for conducting this comparison would be to compare claims data with hospital administrative registers (Katz et al. 1997; Losina et al. 2003). This method is fast and inexpensive, but it is not the most accurate because it does not consider potential problems like wrong or missing diagnoses in the register the comparison is being made to. In some studies, researchers have checked the validity by asking the diagnosed patients themselves whether their diagnosis is true (MacLean et al. 2001; Callhoff et al. 2023). Naturally, this does not confirm a true diagnosis, only the patient's perception of it, but it finds the most obvious recording errors. The completeness of the register can also be approximated by comparing the number of found cases to expected numbers calculated from rates in the same kind of register from another area that is demographically similar (Thygesen et al. 2014).

When studying the validity of register diagnoses by comparing registers or by asking the patients, the PPVs and percentages of correct diagnoses are generally high. The sensitivity of diagnoses is higher because the probability of wrongly recalling a diagnosis false is small, but the specificity is lower because the criteria for what is considered true diagnosis are loosely defined.

A more specific but also more expensive and time-consuming way of validating registry diagnoses is by evaluating patient records and assessing the correctness of the diagnosis based on the information recorded (Thygesen et al. 2014). However, the criteria that must be fulfilled in order for the diagnosis to be assessed as correct may vary. For example, it may have been enough that the diagnosis was originally set by a rheumatologist, or that the clinical picture of the patient's condition suited that of the disease in question based on the expert opinion of the researcher (Klein et al. 2010; Ng et al. 2012; Waldenlind et al. 2014; Poulsen et al. 2017; Carroll et al. 2012; Hanly et al. 2015). The strictest criterion for a true diagnosis is that the fulfilment of classification criteria for the disease and sufficient differential diagnostics must be verified from the patient documents (Bili et al. 2011; Kim et al. 2011). This way the specificity of a diagnosis is high because it is unlikely that the diagnoses stated as true are not correct, but it comes with the cost of lower sensitivity because some disease cases may not fulfil the classification criteria simply due to deficiencies in patient records, and classification criteria may also change over time.

Validating a lack of registry diagnosis by calculating the NPV, meaning a true absence of a diagnosis, is harder and requires the evaluation of controls that do not have a registered diagnosis.

The aim of the planned research defines the method by which the validity of the register diagnoses should be investigated. A study investigating the influence of genetic factors on a disease requires high specificity from the diagnosis instead of finding all the possible cases of the disease in question, whereas epidemiological studies require high completeness and sensitivity in addition to good specificity. (Thygesen et al. 2014)

2.5 Finnish health care registries

Finland has a comprehensive range of health and social welfare registries maintained by different health care authorities such as The Finnish Institute for Health and Welfare (Terveyden ja Hyvinvoinnin Laitos, THL) and the Social Insurance Institution of Finland (Kansaneläkelaitos, KELA) (Gissler et al. 2004). The main Finnish health care registers used and discussed in this dissertation are presented in table 2.

Table 2. The main Finnish health care registers used and discussed in this dissertation.

Register authority	Register
THL	NordDRG Full, Hospital benchmarking data
	Care Register for Health Care (CRHC)
	Register of Primary Health Care Visits
	Care Register for Social Welfare
	Quality registers, including The Finnish Rheumatology Quality Register (FinRheuma)
	Cancer Register
	Implant Register
	Finnish National Infectious Diseases Register
	Medical Birth Register
	Register of Congenital Malformations
	Finnish National Vaccination Register and monitoring of the vaccination programme
Kela	Kanta Archiving Services Kanta Prescription Centre and Prescription Archive
	Drug Reimbursement Register
	Drug Purchase Register
Statistics Finland	The Cause of Death Register

THL, The Finnish Institute for Health and Welfare (Terveyden ja Hyvinvoinnin Laitos); Kela, the Social Insurance Institution of Finland (Kansaneläkelaitos).

THL is an independent, state-owned, research institute operating under the Ministry of Social Affairs and Health (Sosiaali- ja terveysministeriö, STM) (THL. About THL 2023). Duties of THL are established by the Act on the National Institute for Health and Welfare (Finlex. 2023). THL promotes the welfare, health and safety of the population and serves as an authority that maintains a number of statistics and registers in health care, including the CRHC (THL. Statistics on health care services 2023). THL's other registers contain information on a range of topics, including statistics on the use and coverage of preventive services, primary health care visits, the service needs of children, families and the elderly, disability services, cancers, infectious diseases, visual impairment and other morbidity, accidents, implants, substance abuse and addiction data, sexual and reproductive health, medical births, congenital malformations, vaccinations, access to services and treatments, social and health care resources and home care activities. These registers are used as part of international statistical cooperation. THL also maintains nine quality registers of national health care which were defined by STM decree. One of these registers is The Finnish Rheumatology Quality Register (FinRheuma), which collects national information on the treatment outcomes of rheumatic diseases in specialised medical care (THL. The Finnish Rheumatology Quality Register 2023).

KELA is an independent social security institution of Finland, which has its own administration and funding, and is supervised by the Finnish Parliament (Kela. Kela’s organisation 2023). KELA maintains numerous social security registers, including the Drug Reimbursement Register containing data on medicine reimbursement entitlements and the Drug Purchase Register.

Statistics Finland is Finland's national statistical institute, collecting data from a wide range of sources. Statistics Finland maintains a register of causes of death in Finland. (Statistics Finland 2024)

A personal identification number was introduced in Finland in 1964, and all Finnish citizens and permanent residents have this unique identification code. The code is included in all administrative registers and can be used to combine data from different registries. (Sund 2012)

The majority of the data produced by Finnish health care and social welfare institutions during recent years is in electronic form. Finnish legislation, namely the Act on the Electronic Processing of Client Data in Healthcare and Social Welfare (the Client Data Act) (Finlex. Act on the Electronic Processing of Customer Data in Social Welfare and Health Care 2021), requires all public health care and social welfare services providers to enter patient and client data in national Kanta Archiving Services, a database maintained by Kela (Kanta Services 2023). Kanta Services are also mandatory for private sector health care and social welfare providers if they have an information system that stores patient and client data. Health care providers store medical records in the Kanta archives. This means that

the Kanta archives contain information on the patient's diagnoses and visits, allergies and other risks, laboratory results, vaccinations, procedures, medication data, physiological measurements, imaging examinations recorded with a procedure code, data on booked appointments as well as the plan for the patient's examinations, treatment and possible rehabilitation. Patient data has been stored in Kanta Services since 2014 and social welfare client data since 2018. The data, however, is seldom in structured, easily useable form.

The Electronic Prescription Act, which entered into force on 1 January 2017, makes it mandatory for all prescriptions in Finland to be processed electronically at the national Prescription Centre and Prescription Archive, which are a part of Kanta Archiving Services and are maintained by Kela. A paper prescription or telephone prescription can only be used in exceptional cases, and even then, the pharmacy supplying the medicine will convert the prescription into an electronic prescription and save it in the Prescription Centre and Prescription Archive. (Finlex. Act on electronic prescription 2023)

It is possible to obtain data from these health care registers for research purposes. Permission to use the data can be requested from The Finnish Social and Health Data Permit Authority (Findata) (Finnish Social and Health Data Permit Authority Findata 2023). The use of the data is strictly regulated and defined by the Act on the Secondary Use of Health and Social Data (Finlex. Act on the secondary use of social and health data 2019).

2.5.1 The Care Register for Health Care (CRHC)

The Care Register for Health Care (CRHC) is one of the most important of many registries maintained by THL. CRHC collects information about the activities and patients of health centres, hospitals and other institutions that provide inpatient, outpatient or home-nursing care. The data is collected in CRHC for statistical, research and planning purposes. (THL. Care Register for Health Care 2023)

CRHC is a continuation of the Finnish Hospital Discharge Register (FHDR) and replaced it in 1994. CRHC contains nationwide data, connected through a personal identification code, on all hospital inpatient discharges since 1969, specialised outpatient care and day surgery visits to hospitals since 1998, and a count of patients in inpatient care in health centres and hospitals on 31 December each year, when the previous FHDR only included data regarding inpatient care. The diagnoses of these contacts have been recorded using ICD-8 during 1969–1986, ICD-9 during 1987–1995 and ICD-10 since 1996. (Sund 2012; THL. Statistics on health care services 2023)

In addition to the discharge diagnoses, the CRHC data contains numerous other variables including service branch and speciality of the service provider,

municipality of residence of the patient, and admission information such as the referring institution, as well as detailed data on the treatment received and procedures performed during inpatient stays or outpatient and day surgery visits. (THL. Care Register for Health Care 2023)

2.5.2 The Drug Reimbursement Register and the Drug Purchase Register in Finland

The Finnish national health insurance system entitles all patients with certain chronic and severe diseases, such as RA, to special reimbursement for the cost of medications. Information about these reimbursement entitlements and purchases of the medications is recorded in registries maintained by KELA (Kela. Reimbursements for medicine expenses 2023).

The KELA Drug Reimbursement Register includes data on reimbursement entitlements for medicine, which includes the number of reimbursement entitlements during the year and at the end of the year, and the number of entitlements that started that year, the starting date expressed in months, as well as the number of entitlements that ended due to death that year. These reimbursement entitlement statistics are available starting in 1986. (Kela. Statistics on entitlements to reimbursement of medicines 2023)

The reimbursement entitlements for medicine expenses are divided into three reimbursement levels: diseases or disease groups with a basic rate of reimbursement (40%), a lower special rate of reimbursement (65%), and a higher special rate of reimbursement (100%) in year 2023. All the different diseases or disease groups in these three levels have unique codes. For example, code 202 refers to diffuse connective tissue diseases (CTDs), rheumatic arthritis and similar conditions, and a patient granted this reimbursement right gets a 65% discount on the cost of medicine approved in the code 202 category (Kela. Reimbursements for medicine expenses 2023). Most medicines are covered by basic rate of reimbursement if a doctor has issued a prescription for them. For medicines with limited basic reimbursement and medicines with lower or higher special rate of reimbursement, the right to reimbursement must be applied for from Kela and these reimbursement rights are recorded in the Drug Reimbursement Register. The Pharmaceuticals Pricing Board (Hintalautakunta, Hila), operating in connection with the STM, decides whether a medicine is to be reimbursed in a particular reimbursement category.

The Drug Reimbursement Register contains both the code of the disease group and the diagnosis code with which the reimbursement was originally applied for and granted, all connected through a patient's personal identification number. Regrettably, some diseases, such as CTDs, are coded in crude 3-character form, which constrains the use in epidemiological research.

The Drug Purchase Register is collected from pharmacies and based on purchases of reimbursed medicine (Kela. Statistics on purchased medicines reimbursed by health insurance 2023). The register includes information on the reimbursed purchases, including the cost of the medication before reimbursement and the amount of reimbursement. The purchased medicines are coded in the register using the anatomical therapeutic chemical (ATC) code, an international classification system for medicines maintained by the WHO (WHO. ATC Structure and principles 2024).

2.5.3 Finnish biobanks

Biobanks are collections of biological samples of human origin and information on the donor's health data, collected with the donor's permission (Paskal et al. 2018). In biobank-based studies, data obtained from the collected biological samples are combined with related data from electronic health records (Bycroft et al. 2018; Li et al. 2020). Biobank-based research often relies on diagnostic information recorded in health care registers (Li et al. 2020). To select patients for research cohorts, biobank studies also rely on diagnostic information from medical records and registers (Bycroft et al. 2018; Kurki et al. 2023; Sudlow et al. 2015; Rämö et al. 2023; Mishra et al. 2022; Vuorinen et al. 2021).

In Finland there are eleven nationally registered biobanks. These biobanks and their catchment areas and target populations are listed in Table 3. Seven of them are regional biobanks established by universities, health care districts and catchment areas. Four biobanks are nationwide and collect samples from the whole of Finland. The Biobank Act (688/2012) regulates all biobank activities (Finlex. Biobank Act 2012), and a register of Finnish biobanks is kept by the Finnish Medicines Agency (Fimea) (Fimea. Biobanks 2023).

Patients treated in the biobank hospitals have the option to include all their medical data in the biobank during any hospital visit. According to Finnish biobank legislation, written consent is obtained from each patient before their data are included (Finlex. Biobank Act 2012).

After the patient has given consent, a biobank blood sample is taken as part of other laboratory tests related to treatment or follow-up. All available clinical data from hospital electronic medical records are combined, including diagnoses recorded during a patient's outpatient and inpatient visits, medical procedures performed, medicines the patient is taking, and results of laboratory tests and imaging. Information and samples may also be collected as part of separate biobank studies. A biobank consent and blood sample can also be provided by healthy people with no other contact with the hospital. Potential study patients can therefore be searched from the biobank using a variety of criteria.

Table 3. Regional and nationwide biobanks in Finland.

Regional biobanks and their catchment areas	
Auria Biobank, Turku	The Wellbeing Services Counties of Southwest Finland, Satakunta and Ostrobothnia.
Borealis Biobank, Oulu	The Wellbeing Services Counties of North Ostrobothnia, Lapland, Kainuu and Central Ostrobothnia. In addition, infection screening samples collected at maternity clinics across Finland from 1983 to 2016 and early pregnancy screening samples from 2017 to 2021.
Central Finland Biobank, Jyväskylä	The Wellbeing Services County of Central Finland
Biobank of Eastern Finland, Kuopio	The Wellbeing Services Counties of North Savo, North Karelia and South Savo.
Finnish Clinical Biobank Tampere	The Wellbeing Services Counties of Pirkanmaa, South Ostrobothnia and Kanta-Häme
Helsinki Biobank	Hospital District of Helsinki and Uusimaa (HUS), The Wellbeing Services Counties of South Karelia, Kymenlaakso and Päijät-Häme.
Arctic Biobank, University of Oulu	The Northern Finland Birth Cohorts from provinces of Oulu and Lapland including two longitudinal and prospective birth cohorts from years 1966 and 1986 and from two cohorts on aging populations with individuals born in 1935 and 1945. One-year military call-up sample from provinces of Oulu and Lapland in the autumn 2014.
Nationwide biobanks and their target populations	
Blood Service Biobank	Blood donors in Finland
THL Biobank	Several collections from national population studies and disease-specific studies.
The Finnish Hematology Registry and Clinical Biobank	Patients with haematological diseases in Finland
Terveystalo Biobank (run by a private health care provider)	Patients of Terveystalo in Finland

THL, The Finnish Institute for Health and Welfare (Terveyden ja Hyvinvoinnin Laitos).

2.5.4 The FinnGen study

Three country-wide and seven regional biobanks in Finland participate in the FinnGen study, which combines genome information with digital health care data. The project was launched in 2017 and has collected genome and health data from 500,000 biobank participants in Finland. The FinnGen project aims to study the genetic background of numerous diseases and ultimately aims to identify new therapeutic targets and diagnostic possibilities. (Kurki et al. 2023; Finngen: an expedition into genomics and medicine 2023)

2.5.5 Validity of register diagnoses in Finland

The validity of register diagnoses in Finland has been previously studied in certain disease groups like cardiovascular diseases, psychiatric diagnoses, injuries, miscarriages, dementia and Alzheimer's disease (Sund 2012; Vuori et al. 2019; Helle et al. 2022; Solomon et al. 2014), with reasonably good accuracy in the diagnoses. In a meta-analysis by Sund et al. of 32 studies analysing the quality of diagnoses in the Finnish Hospital Discharge Register, which was later replaced by the CRHC, the validity of diagnoses in these registries was shown to vary between 75% and 99% for common diagnoses, but for rare diseases, false-positive diagnoses were more likely (Sund 2012). It should be noted that in several of these studies, the register data was validated by comparing it with data from another register and not from patient medical records (Sund 2012). There are also some diseases, like the division between ST-elevation myocardial infarction and non-ST-elevation myocardial infarction, in which the validity of the diagnosis has been shown to be poor, and these diagnoses should not be used as research outcomes (Okkonen et al. 2020). Only a few studies have been conducted on the accuracy of diagnostic information in Finnish biobank patients. Diagnoses of psoriasis recorded in Finnish biobanks were validated with a PPV of 88.0% (95% CI 82.7–92.2) (Haverinen et al. 2020). To our knowledge, before our study, there were no previous validation studies of rheumatologic diagnoses in Finnish biobank patients or in the CRHC register.

Further, no studies on the validity of RA diagnoses in health care registers have been done in Finland in recent years. Three decades ago, Hakala et al. studied a cohort of 220 subjects with drug reimbursement for chronic rheumatic diseases extracted from the National Sickness Insurance Register. Of these patients, 56% (109/193) fulfilled the ACR 1987 revised criteria for the classification of RA (Arnett et al. 1988) and were 16 years or older at the disease onset, but 6% of these 109 patients had some disease other than RA. Of the 193 patients, 22% did not fulfil the ACR 1987 criteria despite a clinician’s primary diagnosis of RA (Hakala et al. 1993).

2.6 Rheumatoid Arthritis

2.6.1 RA as a disease

Rheumatoid arthritis (RA) is the most common autoimmune inflammatory arthritis in the world, affecting almost 1% of adults worldwide. The incidence and prevalence of RA varies around the world and is more common in industrialised countries (Finckh et al. 2022). In Northern Europe, the prevalence of RA has been 0.4–0.8% and the annual incidence is 20–40/100,000. Following the year 2000, the incidence of seropositive RA in Finland has remained around 29/100,000 for adults, while the

incidence of seronegative RA has decreased from 18/10,000 between the years 2000 and 2004 to 14/100,000 between years 2010 and 2014, due to the introduction of new classification criteria (Muilu et al. 2019). RA is two to three times more common in females than in males (Gravallese et al. 2023).

Seropositive RA is an autoimmune disease that starts with predisposing genetic factors and progresses over the years into an overt disease due to the influence of external factors. Smoking is the single biggest risk factor, and obesity, low vitamin D levels, low alcohol consumption, poor dental health, unhealthy diet, and use of oral contraceptives also increase the risk. Several infections have been proposed as aetiologic and contributing factors. (Sparks 2019; Scherer et al. 2020; Gravallese et al. 2023)

RA most often affects the smaller joints of the hands, wrists and feet, usually in a symmetric and polyarthritic pattern, but can occur in any joint. Current diagnostic methods have somewhat changed the classic clinical picture of RA, and a recent study from Finland showed that 25% of patients with newly diagnosed RA did not have symmetric swelling at baseline (Weman et al. 2023). Untreated RA causes painful swelling and stiffness of the joints and can eventually cause structural damage, bone erosion and joint deformity. Inflammatory values c-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) may be elevated. Rheumatoid factor (RF) and anti-citrullinated peptide antibodies (ACPA), autoantibodies related to RA, are often but not always positive, and RA is divided into seropositive RA and seronegative RA based on the serological status of the patient. RA can cause numerous but nowadays rare manifestations outside joints, such as interstitial lung disease, pericarditis, pleuritis, vasculitis, secondary Sjögren syndrome and amyloidosis. (Sparks 2019)

Currently, treatment of RA aims for remission. The mainstream treatment in Finland is methotrexate in combination with other disease-modifying anti-rheumatic drugs (DMARDs) (Duodecim. Rheumatoid Arthritis: Current Care Guidelines 2022). Biologic DMARDs (bDMARDs) or janus kinase (JAK) inhibitors are used if remission or low disease activity is not reached with conventional DMARDs. NSAIDs and glucocorticoids, either intra-articularly or perorally, are used in the beginning and during possible flare-ups of high disease activity. (Sparks 2019; Scherer et al. 2020)

2.6.2 Diagnosing RA

Diagnosing RA is complex because there is no single test to confirm the diagnosis. Diagnostic criteria do not exist, but there are some classification criteria that have been developed primarily for research purposes to enable clinical studies to have uniform cohorts. ACR 1987 and ACR/EULAR 2010 classification criteria for RA

(Aletaha et al. 2010; Arnett et al. 1988) are the criteria used most often in clinical settings to classify arthritis patients as having RA. A diagnosis of RA, however, is ultimately an opinion of the rheumatologist (Aggarwal et al. 2015; Aletaha et al. 2010; Arnett et al. 1988). This opinion is based on a subjective assessment of a combination of clinical signs and symptoms such as the duration of symptoms, amount and location of tender and swollen joints, possible extra-articular manifestations, imaging results, potentially elevated inflammatory markers as well as the presence of RF or ACPA. The role of ACPA in the diagnosis of RA has been highlighted by the new ACR/EULAR 2010 classification criteria. Relevant differential diagnostics should be considered, and knowledge about the epidemiology of the rheumatologist's geographical area should be considered (Aggarwal et al. 2015; Aletaha et al. 2018). Different imaging modalities are used in the diagnosis of RA. US imaging and MRI can assess the presence and activity of inflammation, synovitis, or tenosynovitis, as well as structural damage like erosions in the joints and tendon injuries. Radiographic imaging can show erosions, periarticular osteoporosis and joint space narrowing. (Baker et al. 2018; Tan et al. 2011)

Because of the complexity of diagnosing RA, individuals in registers with diagnosis of especially seronegative RA may actually have another arthritis or even non-rheumatic disease.

Seronegative RA is a heterogeneous disease entity that is challenging to diagnose, and differential diagnostics between RA and other inflammatory arthritides can be difficult. Many patients present with a competing diagnosis during longer follow-up. In a register study by Paalanen et al., of the 9784 patients with seronegative RA, 5.7% of the patients were diagnosed with either psoriatic arthritis, nonradiographical axial spondyloarthritis, ankylosing spondylitis, or inflammatory bowel disease-related arthritis during 15 years of follow-up (Paalanen et al. 2021). In another study by Paalanen et al., 435 patients with a diagnosis of seronegative RA were followed for up to 10 years in the department of rheumatology of the Central Finland Central Hospital. During follow-up, 13 patients (3%) could be reclassified as seropositive or erosive RA, 68 (16%) as polymyalgia rheumatica, 46 (11%) psoriatic arthritis, 45 (10%) osteoarthritis, 38 (8.7%) spondyloarthritis, 15 (3.4%) plausible reactive arthritis, 10 (2.3%) gout, 17 (3.9%) pseudogout, 6 (1.4%) paraneoplastic arthritis, 6 (1.4%) juvenile idiopathic arthritis, 2 (0.5%) haemochromatosis, 3 (0.7%) ankylosing spondylitis, 2 (0.5%) giant cell arteritis, and 8 other miscellaneous diagnoses. 140 patients (32%) could not be reclassified in any one diagnosis and had features of transient arthritis or seronegative spondyloarthritis, while 49 (11%) remained unspecified (Paalanen et al. 2019).

2.6.3 Validity of RA diagnoses

The validity of RA diagnoses has specifically been examined in many studies, and algorithms have been developed to identify patients with RA in the registers. Requirements for a correct diagnosis in some of these previous studies are presented in Table 4.

Table 4. Previous studies on the validity of RA diagnoses.

Study	Year	Country	Study population	Validity	Requirement for correct diagnosis
Singh et al.	2004	USA	Rheum.	66%	A diagnosis of RA made by a rheumatologist on 2 separate occasions > 6 weeks apart
Bili et al.	2011	USA	Rheum.	97%	Fulfilment of ACR -87 classification criteria
Widdifield et al.	2013	Canada	Rheum.	55%	RA documented in patient records, clinical diagnosis
Waldenlind et al.	2014	Sweden	Rheum.	90%	Fulfilment of ACR -87 or ACR/EULAR classification criteria or having a clinical diagnosis of RA.
Carrara et al.	2015	Italy	Rheum.	77%	Clinical diagnosis or fulfilment of classification criteria
Kim et al.	2011	USA	Rheum., general	Rheum. 67%, general 56%	A diagnosis of RA made by a rheumatologist
Ng et al.	2012	USA	Rheum., general	Rheum. 40%, general 31%	Clinical diagnosis
Hanly et al.	2015	Canada	Rheum., general	Rheum. 47%, general 39%	A diagnosis of RA made by a rheumatologist
Poulsen et al.	2017	Denmark	Rheum., general	Rheum. 96%, general 79%	Expert opinion
Thomas et al.	2008	UK	General	56 %	Fulfilment of ACR -87 criteria and clinical diagnosis
Carroll et al.	2012	USA	General	22–49%	Clinical diagnosis
Callhoff et al.	2023	Germany	General	81%	Confirmed by the patient
MacLean et al.	2001	USA	General	92%	Confirmed by the patient

Rheum., study patients from a department of rheumatology; General, study patients from any department or primary health care; RA, rheumatoid arthritis; ACR, American College of Rheumatology; EULAR, European League Against Rheumatism.

To the best of our knowledge, no studies on the validity of RA diagnoses in health care registers in Finland have been done. In the other Nordic countries, the structure, statistical methods and treatment practices of the health care system are similar to those in Finland. In Sweden, Waldenlind et al. studied patients diagnosed

with RA by a rheumatology clinic on at least two visits (Waldenlind et al. 2014). Approximately 90% of these patients had a definite diagnosis of RA defined by the fulfilment of ACR -87 or ACR/EULAR 2010 classification criteria or having a clinical diagnosis of RA. Either of the two classification criteria were fulfilled in 94% of the seropositive patients. In Denmark, Poulsen et al. studied the validity of the diagnoses of RA in the Danish Registry for Biologic Therapies in Rheumatology (DANBIO), which is a register of inflammatory arthritis diseases, and in the Danish National Patient Register (DNPR) (Poulsen et al. 2017). The inclusion criteria for the patients were either a diagnosis of RA in the DANBIO register or a diagnosis of RA in the DNPR register that was obtained during at least two visits at a rheumatology clinic. A correct diagnosis of RA was determined by expert opinion. In the DANBIO register, the accuracy of the diagnoses was 96%, and in the DNPR register, it was slightly lower at 79%.

In other parts of Europe outside Nordic countries, RA diagnoses in data from a large, German, statutory health insurance database has been studied by Callhoff et al. The patients with a database diagnosis of RA on at least two visits were asked whether they had RA. The diagnosis was confirmed by the patient in 81% of the cases, 94% for seropositive RA and 76% for seronegative RA. A correct diagnosis was more likely if the patient had elevated inflammatory markers (82%), had visited the rheumatology department (85%) or used specific medication (89%) (Callhoff et al. 2023).

Most of the previous studies on RA have validated diagnoses set at rheumatology clinics, and a minority of the studies have included diagnoses set at any speciality or primary health care appointments. The percentage of the correct RA diagnosis has varied between 40%–97% in rheumatology clinic patients (Ng et al. 2012; Waldenlind et al. 2014; Poulsen et al. 2017; Hanly et al. 2015; Bili et al. 2011; Kim et al. 2011; Carrara et al. 2015; Widdifield et al. 2013; Chung et al. 2013; Singh et al. 2004) and between 22%–56% in patients diagnosed in any speciality or in primary health care (Ng et al. 2012; Carroll et al. 2012; Hanly et al. 2015; Kim et al. 2011; Thomas et al. 2008; Chung et al. 2013).

In the aforementioned studies, many algorithms have been suggested to better identify patients with RA from the registers. These algorithms have usually included the number and the location of the diagnoses (for example, at a rheumatology clinic or elsewhere), anti-rheumatic medications prescribed for the patient, and other rheumatic diagnoses (MacLean et al. 2001; Callhoff et al. 2023; Ng et al. 2012; Waldenlind et al. 2014; Carroll et al. 2012; Kim et al. 2011; Carrara et al. 2015; Thomas et al. 2008; Widdifield et al. 2013; Chung et al. 2013; Singh et al. 2004; Liao et al. 2010).

2.7 Systemic Sclerosis

2.7.1 SSc as a disease

Systemic sclerosis (SSc), which is also called scleroderma, is a rare autoimmune disorder affecting 0.01% of the people in the world (Denton et al. 2017). The incidence of SSc varies between 0.77/100,000 in the Netherlands to 5.6/100,000 in the United States (Zhong et al. 2019). The incidence in Finland during 2014–2018 was 2.8/100,000 in people over 16 years old (Kortelainen et al. 2024). SSc is three to eight times more common in females than in males (de Angelis et al. 2022). Even though SSc is rare, it has high morbidity, and mortality in SSc is the highest of all rheumatic diseases (Volkmann et al. 2023; Calderon et al. 2021).

Aetiology and pathogenesis of SSc is complex and still partly unknown. The onset of the disease is likely caused by a combination of genetic susceptibility and several environmental factors. The early stage of the disease is characterised by microvascular dysfunction and autoimmune phenomena. SSc causes inflammation and fibrotic scarring in the connective tissues of the skin, joints, internal organs and blood vessels, leading to several organ-based manifestations. (Denton et al. 2017)

The first symptom of SSc is often Raynaud's phenomenon (RP), which is a sign of digital vasculopathy. Vasculopathy can also cause ulceration, gangrene and even autoamputation. The skin is nearly always affected, starting from the distal parts and progressing proximally. Thickening and fibrosis of the skin can lead to joint contractures. Other possible manifestations include calcinosis of the skin, arthritis, interstitial lung disease (ILD), pulmonary arterial hypertension (PAH), gastrointestinal tract manifestations, renal crisis and cardiac problems. (Volkmann et al. 2023; Denton et al. 2017)

SSc is divided into different subcategories based on the extent of skin involvement. Diffuse cutaneous systemic sclerosis (dcSSc) patients have proximal skin involvement, and limited cutaneous systemic sclerosis (lcSSc) patients have skin involvement in the limbs distal to the elbows or knees, with or without changes in the skin of face and neck. Systemic sclerosis with no skin manifestations is called SSc sine scleroderma (ssSSc). The disease can also occur as SSc overlap syndrome, when there are features of both SSc and some other autoimmune disease like RA, polymyositis, dermatomyositis, systemic lupus erythematosus (SLE) or Sjögren's syndrome, and criteria for both diseases are fulfilled. (Volkmann et al. 2023; Denton et al. 2017)

Several autoantibodies are associated with SSc. Antinuclear antibody (ANA) is elevated in up to 95% of the patients. Several autoantibodies are SSc-specific, including anti-centromeric protein (CENP), anti-DNA topoisomerase 1 (ScL-70), Anti-RNA polymerase 3, Anti-U1 ribonucleoprotein (RNP), anti-fibrillarin, and

anti-Th/To -antibodies. These SSc-specific antibodies can be used in predicting organ involvement and severity of the SSc. Inflammatory markers like CRP can be elevated. (Volkmann et al. 2023; Denton et al. 2017)

The treatment of SSc is targeted according to the organ manifestations and major complications and is often a collaborative effort by several specialties. Treatment ranges from immunosuppressive medication, DMARDs and bDMARDs, to anti-fibrotic agents, angiotensin-converting enzyme (ACE) -inhibitors, vasodilators, endothelin receptor antagonists, phosphodiesterase isoenzyme 5 inhibitors, prostaglandin and even to haematopoietic autologous stem-cell transplantation. (Denton et al. 2017; Pope et al. 2023)

2.7.2 Diagnosing SSc

Early clinical signs of SSc can be symptoms that are common in the population, and due to the rareness of SSc, diagnosis can often be delayed. No diagnostic criteria exist for SSc, and the diagnosis is made by a rheumatologist considering clinical signs and symptoms, as well as a wide range of imaging, function tests and laboratory analyses like autoantibodies, inflammatory values and organ-specific blood tests. At the time of diagnosis and later at certain intervals during follow-up, the presence of possible organ manifestations should be specifically and systematically screened (Volkmann et al. 2023).

Due to a wide range of possible affected organs, many different imaging modalities are used in the diagnosis of SSc. Conventional radiograph images are used to evaluate manifestations like acro-osteolysis, calcifications, and soft tissue thinning. Computed tomography (CT) scan can reveal a dilated oesophagus, pericardial effusion and pericardial fibrosis, and high-resolution computed tomography (HRCT) of the lungs can reveal ILD. Ultrasound (US) can help to detect synovitis and problems in tendons. US of the heart, echocardiography, is the most frequently used initial method in assessing possible PAH. MRI is a useful method to evaluate musculoskeletal manifestations like synovitis, tendinitis and joint erosions. MRI can evaluate morphology, function and vitality of cardiac muscle and reveal heart manifestations like cardiac and pericardial fibrosis and myositis of the heart. (Rutka et al. 2021)

Nailfold videocapillaroscopy (NVC) is a noninvasive technique to evaluate the microvascular damage. The scleroderma pattern, including giant capillaries, microhaemorrhages, capillary loss, and abnormal shapes of capillaries, can be seen in the majority of SSc patients. (Smith et al. 2023)

Other examinations typically performed are oesophagogastroscopy and colonoscopy to evaluate gastrointestinal manifestations, spirometry and diffusion capacity of the lung for carbon monoxide to evaluate pulmonary function, modified

Rodnan Skin Score (mRSS) to evaluate the extent of skin involvement, and right heart catheterisation to confirm PAH. (Volkmann et al. 2023; Denton et al. 2017)

Several classification criteria have been developed for identifying patients with a similar clinical entity for research cohorts (Van Den Hoogen et al. 2013; LeRoy et al. 2001; Masi 1980). These classification criteria are presented in table 5. Even if these classification criteria are not meant as diagnostic criteria, they are often used in clinical practice to help diagnose SSc. The 2013 revised ACR/EULAR classification criteria for SSc (Van Den Hoogen et al. 2013) and the 2001 LeRoy and Medsger classification criteria for early SSc (LeRoy et al. 2001) are currently the most commonly used criteria.

The patients with SSc usually fulfil the 2013 ACR/EULAR classification criteria (Jordan et al. 2015; Araújo et al. 2017), but nonfulfillment of these criteria should not prevent the diagnosis of SSc, especially in the disease's early stages where the criteria are less sensitive. In a study by Araújo et al., 28% of the patients with early SSc fulfilling the LeRoy and Medsger criteria, fulfilled the 2013 ACR/EULAR criteria (Araújo et al. 2017).

As treatment options for SSc improve, it is increasingly important to diagnose the disease in its early stages, before the manifestation of organ complications when the treatments are most effective. The increase in and better availability of diagnostic possibilities have made early diagnosis easier. For example, the increasing availability of NVC has made it possible to better identify those patients with early SSc and ensure that they receive proper follow-up (Araújo et al. 2017). On the other hand, if the diagnosis of SSc is made too easily due to not considering the classification criteria, it may lead to heterogeneity of patient cohorts and make it difficult to compare the data between countries.

Table 5. Classification criteria for diagnosis of systemic sclerosis.

Criteria	Clinical and laboratory characteristics	Requirement for diagnosis
ACR criteria, 1980	Major criteria: • Proximal scleroderma Minor criteria: • Sclerodactyly • Digital pitting scars of fingertips or loss of substance of the distal finger pad • Bibasilar pulmonary fibrosis	One major or two or more minor criteria required for diagnosis
Revised ACR/EULAR, 2013	• Skin thickening of the fingers of both hands extending proximal to the metacarpophalangeal joints (9 points) • Puffy fingers (2 points) or sclerodactyly of the fingers (distal to MCP but proximal to the PIPs) (4 points) • Digital tip ulcers (2 points) or fingertip pitting scars (3 points) • Telangiectasia (2 points) • Abnormal nailfold capillaries (2 points) • Pulmonary arterial hypertension and/or Interstitial lung disease (2 points) • Raynaud's phenomenon (3 points) • Scleroderma related antibodies (any of anti-centromere, anti-topoisomerase I, anti-RNA polymerase III) (3 points)	A total score of 9 or more required for definite diagnosis
LeRoy and Medsger criteria for early SSc, 2001	**Limited SSc (lSSc):** • Raynaud's phenomenon (objective documentation) plus SSc-type nailfold capillary pattern or SSc selective autoantibodies or • Raynaud's phenomenon (subjective only) plus SSc-type nailfold capillary pattern and SSc selective antibodies **Limited cutaneous SSc (lcSSc):** • Criteria for lSSc plus distal cutaneous changes **Diffuse cutaneous SSc:** • Criteria for lSSc plus proximal cutaneous changes **Diffuse fasciitis with eosinophilia:** • Proximal cutaneous changes without criteria for lSSc or lcSSc	

ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; SSc, systemic sclerosis.

2.7.3 Validity of SSc diagnoses

Compared to RA, the validity of SSc diagnoses has been studied to a much lesser extent. Some of the previous studies are presented in table 6. PPV of SSc diagnoses recorded with an ICD-9 code 710.1 in the United States has varied between 63% and 76% (Bernatsky et al. 2011; Valenzuela et al. 2015). For SSc diagnoses recorded with ICD-10 in Switzerland, the PPV has been found to be 61% (Ziswiler et al. 2007). More recently, the accuracy of ICD-10 codes of SSc patients in the French

hospital database for inpatient stays during 2010 to 2017 was assessed by Chaves et al. They found the diagnoses to be reliable with a PPV of 93% for SSc overall and 95% for lcSSc (Chaves et al. 2020). When studying prevalence and incidence of SSc in Sweden, Andréasson et al. noticed that 5.3% of all the M34 diagnoses of rheumatology clinic patients had been input incorrectly (Andréasson et al. 2014). To the best of our knowledge, the validity of SSc ICD-10 diagnoses has not been studied in recent years in a setting where the diagnoses were made in any speciality and in both inpatient and outpatient visits. Because of the rarity of SSc, it is important that these patients are appropriately identified from health care registers.

Table 6. Previous studies on the validity of SSc diagnoses.

Study	Year	Country	Study population	Validity	Requirement for correct diagnosis
Ziswiler et al.	2007	Switzerland	Rheum.	61%	Fulfilment of 1980 ACR classification criteria
Bernatsky et al.	2011	USA	Rheum.	63%	A diagnosis of SSc made by a rheumatologist
Valenzuela et al.	2015	USA	General	76%	Fulfilment of one of the following classification criteria: 1980 ACR, CREST or the 2013 revised ACR/EULAR.
Chaves et al.	2020	France	General	93% for SSc, 95% for lcSSc	Fulfilment of 2013 revised ACR/EULAR classification criteria

Rheum., study patients from a department of rheumatology; General, study patients from any department or primary health care; SSc systemic sclerosis; lcSSc, limited cutaneous systemic sclerosis; ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; CREST, Calcinosis, Raynaud's phenomenon, Esophageal dysmotility, Sclerodactyly and Telangiectasias.

2.8 Polymyalgia Rheumatica

2.8.1 PMR as a disease

Polymyalgia rheumatica (PMR) is a common inflammatory disease affecting patients over the age of 50 almost exclusively with incidence increasing progressively with age. PMR is one of the most common rheumatological conditions in people over 50, and it affects females three times more than males. PMR is most common in Scandinavian countries, where the incidence ranges between 41–113/100,000 in people 50 years and older, which increases the clinical importance of this disorder in Northern Europe. (Lundberg et al. 2022; Espígol-Frigolé et al. 2023)

Like several other rheumatological conditions, PMR is likely caused by a combination of genetic susceptibility, aging and environmental factors. Viral infections and aging of the immune system have been proposed to be causative agents. (Lundberg et al. 2022)

PMR is characterised by muscle pain and morning stiffness, especially in the neck, shoulders, upper arms, and pelvic girdle, usually bilaterally. Other common symptoms include fatigue, weight loss, depression and low-grade fever. Bilateral synovitis of shoulders and hips can occur. Some patients have also distal manifestations, like pain and swelling in wrists and knees. Up to 25% of the patients can have peripheral synovitis, but unlike in RA, the synovitis is asymmetric and nonerosive. The onset of symptoms is often sudden, but it can take a couple of weeks to reach their maximum severity. (Lundberg et al. 2022; Espígol-Frigolé et al. 2023)

PMR is closely associated with giant cell arteritis (GCA), which is a vasculitis that is most often located in the arteries of the head and scalp, especially the temporal arteries. PMR and GCA can be considered to be different ends of the same disease spectrum. PMR can precede or follow GCA, 50% of patients with GCA also have symptoms of PMR, and 15–20% of patients with PMR have or will develop vasculitis. It is believed that the pathogenic mechanisms of these two diseases are common but ultimately target different tissues – synovial structures in PMR and arterial structures in GCA. (Lundberg et al. 2022; Espígol-Frigolé et al. 2023; Tomelleri et al. 2023)

US examination and MRI imaging can show subacromial or subdeltoid bursitis and tendinitis in the long head of the biceps. Positron emission tomography (PET) can show inflammation in interspinous bursa and adjacent to the ischial tuberosities. If a PMR patient has subclinical GCA, PET may also show inflammation in the arteries. Inflammatory values, CRP and ESR, are almost always elevated. (Lundberg et al. 2022; Espígol-Frigolé et al. 2023)

The usual treatment of PMR is oral glucocorticoids, starting with a prednisolone dose of approximately 12.5–25 mg per day and then tapering the dosage until discontinuation around one to two years from diagnosis (Dejaco et al. 2015). Usually, glucocorticoids give a rapid and full response to symptoms and inflammatory values, and this response is considered diagnostic for PMR. A relapse can happen during tapering or later after treatment is discontinued. In these cases, DMARD treatment, usually methotrexate, has been used. the role of bDMARDs, like tocilizumab or sarilumab, is increasing in the treatment of more resistant cases and when the risk of serious side effects of glucocorticoids is high. (Espígol-Frigolé et al. 2023)

2.8.2 Diagnosing PMR

The diagnosis of PMR is mainly clinical and based on the pattern of symptoms and status findings as well as laboratory evidence of an inflammatory acute-phase reaction in a person over 50 years old. Imaging with US, MRI and PET is sometimes used to make the diagnosis. (Lundberg et al. 2022; Espígol-Frigolé et al. 2023)

Several sets of classification criteria for PMR have been created for research purposes (Bird et al. 1979; Jones et al. 1981; Chuang et al. 1982; Healey 1984; Dasgubta et al. 2012). As in RA and SSc, the classification criteria are often used as if they were diagnostic criteria. These criteria are presented in Table 7.

Even if these criteria have common elements like shoulder pain, morning stiffness and elevated inflammatory values, they differ somewhat from each other and thus cover slightly different disease patterns. The four older criteria are based on anamnesis and abnormal inflammatory values, whereas the ACR/EULAR criteria involve the use of US in the diagnosis. ACR/EULAR is also the first criteria made by prospective longitudinal analysis, compared to the older criteria where a retrospective chart review of patients already diagnosed with PMR had been performed. In addition, the response to glucocorticoid treatment is not included in the ACR/EULAR criteria due to the unspecific effect of glucocorticoids on many different conditions and symptoms. (Espígol-Frigolé et al. 2023)

Diagnosing PMR can be difficult, as there is no gold standard to confirm the diagnosis, and many symptoms and findings of PMR may be present in other conditions. Therefore, it is essential to rule out other conditions with similar presentation. Misdiagnosis can lead to unnecessary harm and risks caused by long-term glucocorticoid treatment. Additionally, the real underlying undiagnosed disease may be left untreated, and this may potentially result in irreversible complications like loss of vision due to posterior ciliary or retinal arteritis. Conditions mimicking PMR include a wide variety of disorders, including other inflammatory arthritides like RA, polymyositis, and other myopathies, infections, malignancies, musculoskeletal disorders due to repetitive strain or degeneration, endocrinological diseases and chronic pain syndromes. (Lundberg et al. 2022; Gonzalez-Gay et al. 2000; Michet et al. 2008; Nothnagl et al. 2006)

Table 7. Classification criteria for diagnosis of polymyalgia rheumatica

Criteria	Clinical and laboratory characteristics	Requirement for diagnosis
Bird, 1979	• Shoulder pain and/or stiffness bilaterally • Onset of illness of < 2 weeks duration • Initial ESR ≥40mm/h • Morning stiffness duration >1 h • Age > 65 years • Depression and/or loss of weight • Upper arm tenderness bilaterally	Three or more of these characteristics required for probable diagnosis
Jones and Hazleman, 1981	• Shoulder and pelvic girdle pain without muscle weakness • Morning stiffness • Duration of at least 2 months unless treated • ESR >30 mm/h or CRP >6 mg/l • Absence of rheumatoid or inflammatory arthritis or malignant disease • Absence of objective signs of muscle disease • Prompt and dramatic response to systemic corticosteroids	All the characteristics required for diagnosis
Chuang and Hunder, 1982	• Bilateral aching and stiffness ≥ 1 month in two of the following: neck or torso, shoulders or upper arms, hips or thighs • Age ≥ 50 years • ESR >40 mm/h • Exclusion of other diagnosis, with the exception of GCA	All the characteristics required for diagnosis
Healey, 1984	• Persistent pain >1 month in two of the following: neck, shoulders, pelvic girdle • Morning stiffness >1 h • ESR >40 mm/h • Absence of other joint or musculoskeletal diseases • Rapid response to prednisolone (≤20 mg/day)	Age of >50 years and at least three of these characteristics required for diagnosis
ACR/EULAR, 2012	• Morning stiffness >45 min (two points) • Hip pain or limited range of motion (one point) • Absence of RF or ACPA (two points) • Absence of other joint involvement (one point) • If ultrasonography available, at least one shoulder with subdeltoid bursitis, biceps tenosynovitis or glenohumeral synovitis (posterior or axillary); and at least one hip with synovitis or trochanteric bursitis (one point) • If ultrasonography available, both shoulders with subdeltoid bursitis, biceps tenosynovitis or glenohumeral synovitis (one point)	Age ≥50 years, bilateral shoulder aching, abnormal CRP and/or ESR and at least four points (without ultrasound) or at least five points (with ultrasound) required for diagnosis

ESR, erythrocyte sedimentation rate; CRP, c-reactive protein; RA, rheumatoid arthritis; GCA, giant cell arteritis; RF, rheumatoid factor; ACPA, anti-citrullinated protein antibody.

2.8.3 Validity of PMR diagnoses

Compared to the prevalence of PMR, the validity of PMR diagnoses has been studied relatively little. Some of the previous studies are presented in table 8. In a Swedish study by Fors et al., the validity of PMR diagnoses set in primary health care between 2000 and 2013 has been shown to be 60% (Fors et al. 2019). In a hospital setting, the persistence of PMR diagnoses has varied between 48% to 79% (Caporali et al. 2001; Falsetti et al. 2011; Bernatsky et al. 2011). In a follow-up of UK rheumatology clinic patients with PMR diagnosed on the first visit, 5% of these diagnoses changed into RA and 12% into GCA, and 83% stayed as PMR (Pease et al. 2005). In many studies, the proportion of correct diagnoses has been found to be higher if the diagnosis was made in the department of rheumatology compared to other specialities in the hospital.

Table 8. Previous studies on the validity of PMR diagnoses.

Study	Year	Country	Study population	Validity	Requirement for correct diagnosis
Caporali et al.	2001	Italy	Rheumatology and internal medicine	72%	Fulfilment of Jones and Hazleman classification criteria
Pease et al.	2005	UK	Rheumatology	83%	Fulfilment of 2013 revised ACR/EULAR classification criteria
Falsetti et al.	2011	Italy	Rheumatology	48%	Fulfilment of Bird classification criteria
Bernatsky et al.	2011	USA	Rheumatology	79%	A diagnosis of PMR made by a rheumatologist
Fors et al.	2019	Sweden	Primary health care	60%	Clinical diagnosis

PMR, polymyalgia rheumatica; ACR, American College of Rheumatology; EULAR, European League Against Rheumatism

In the UK, a prospective glucocorticoid treatment study followed 122 rheumatology clinic patients who were diagnosed with PMR during 2001 to 2003. These patients all met the Jones and Hazleman classification criteria for PMR at the time of diagnosis, and vigorous differential diagnostics had already been made. By month 12, seven of the patients (6%) were considered unlikely to have had PMR, even though the original diagnosis had been considered carefully (Hutchings et al. 2007).

3 Aims

The aims of this dissertation were to study the validity of the diagnoses of three rheumatological diseases (RA, SSc, PMR) in Finnish health care registers and the factors contributing to the validity of these diagnoses in order to ensure the quality of future register-based studies. For PMR, we also sought to find out the most common diseases misdiagnosed as PMR and which factors predicted a change of diagnosis during follow-up. We wanted to investigate the accuracy of the data in the registers and how to best utilise the registry data. Improving validity is important in all registry-based studies, and especially when examining the genetic basis of the diseases, such as FinnGen and other biobank studies.

The specific aims of dissertation studies were:

1. To analyse the accuracy of Finnish biobank diagnoses of RA supplemented with data recorded in the CRHC and in the Drug Reimbursement Register.
2. To analyse the validity of the diagnoses of systemic sclerosis in two Finnish university hospitals.
3. To analyse how often a competing diagnosis was found during a clinical follow-up in patients diagnosed with polymyalgia rheumatica in a Finnish university hospital. Additionally, the study aimed to determine the most common conditions misdiagnosed as polymyalgia rheumatica, and to investigate which factors predicted a change of diagnosis during follow-up.

4 Materials and Methods

4.1 The validity of rheumatoid arthritis diagnoses in Finnish biobanks (Study I)

4.1.1 Study population

The patients included in this study were randomly selected from the electronic patient records of five hospital biobanks in Finland listed in Table 9. The study sample included 250 patients with a diagnosis of RA, divided into 125 patients registered with ICD-10 codes M05.8 and M05.9 for seropositive RA, and 125 patients registered with ICD-10 code M06.0 for seronegative RA. The control group consisted of 250 age- and gender-matched controls, who had no diagnosis of RA registered in the patient records of the participating hospitals. All five biobanks selected an equal number of patients for the study: 25 patients with seropositive RA, 25 with seronegative RA and 50 controls.

Table 9. Biobanks participating in study I.

Biobank	Location
Auria Biobank	Turku
Borealis Biobank	Oulu
Biobank of Eastern Finland	Kuopio
Central Finland Biobank	Jyväskylä
Finnish Clinical Biobank Tampere	Tampere

The biobank of each participating hospital performed the initial selection of patients by random sampling from the electronic patient records of that hospital using the predefined diagnosis codes. The patients selected for the study were required to have the inclusion diagnosis registered in the electronic medical records of the hospital on at least one visit to any speciality during 2007 to 2018. The

diagnoses registered on hospital visits are transferred from the hospital records to the CRHC, so it was assumed that the diagnoses in hospital records would correspond to those in the CRHC. However, when the CRHC data on the number of visits with the diagnosis were obtained, it was discovered that the data on diagnoses in CRHC did not fully match the data in the hospital registry. At this stage, it was decided to change the focus of the study by supplementing the biobank data from the hospital records with CRHC data and analysing only those patients who had at least one visit with a diagnosis of RA also recorded in the CRHC.

Figure 1 presents the inclusion and exclusion of patients in the study. Five patients were excluded from the control group because they were later found to have CRHC-registered visits with RA diagnosis at a hospital not participating in the study. Three patients with diagnosed seropositive RA and two patients with diagnosed seronegative RA were ultimately analysed as part of the control group because they were later found to have no visits with these diagnoses in the CRHC despite the diagnoses being recorded in local hospital patient records. A total of 12 patients, four in the seropositive and eight in the seronegative group, were excluded from the final

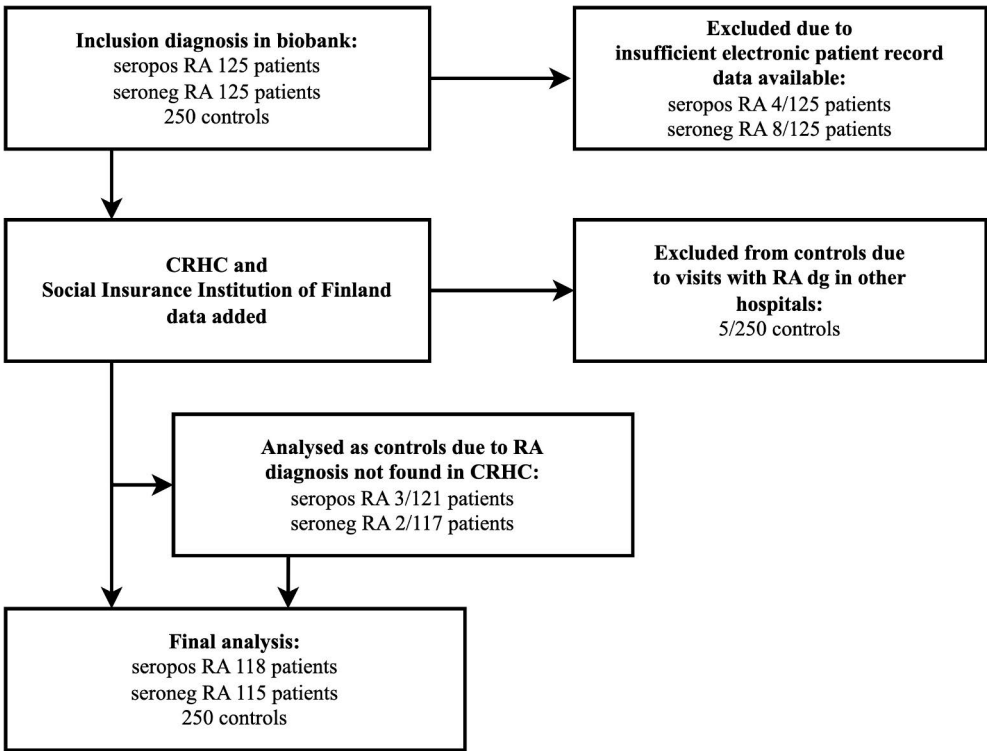

Figure 1. Study flowchart with inclusion and exclusion of the patients. CRHC, Finnish Care Register for Health Care; dg, diagnosis; RA, rheumatoid arthritis; seroneg, seronegative; seropos, seropositive. Modified from Paltta et al., 2023 (I).

analysis because the reviewer was not able to confirm the correctness of the diagnosis due to insufficient electronic patient record data available for analysis in the participating biobank hospital.

The initial diagnosis of RA was not required to be made during the study years but could have been made earlier or in another hospital. The inclusion visit could therefore be at any stage of RA or in connection with the treatment of another medical condition. The reviewer collected the year of the initial diagnosis of RA from the patient records. Length of follow-up was calculated as the time from initial diagnosis to the last contact at the department of rheumatology.

4.1.2 Register data

For each patient and control, data was retrieved from the CRHC on the number of visits registered with a diagnosis of either seropositive or seronegative RA. This data retrieval was done after the initial selection of patients and controls from participating biobanks. This sequence resulted in some people being assigned to the control group even though they had been diagnosed with RA at another hospital, and some people were assigned to the RA group even though there was no diagnosis of RA in the CRHC.

When analysing the number of visits for patients with an inclusion diagnosis of seropositive RA, only data on the number of visits with a diagnosis of seropositive RA were used, and for patients with an inclusion diagnosis of seronegative RA, only visits for seronegative RA were analysed. When all the patients with an inclusion diagnosis of either seronegative or seropositive RA were analysed as a common group of RA, data on both seropositive and seronegative RA visits were used.

From the Drug Reimbursement Register of KELA, we searched for information on whether the patients and controls were entitled to special reimbursement of medicine expenses. The specific entitlements searched for were entitlement to reimbursement of the cost of DMARDs with code 202 for CTDs, RA, and comparable diseases, and entitlement to reimbursement of the cost of bDMARDs with the codes 281 or 313 for RA, juvenile idiopathic arthritis, psoriatic arthritis, ankylosing spondylitis, and comparable diseases. We also checked whether these reimbursement entitlements were originally granted with ICD-10 code M05 for seropositive RA, ICD-10 code M06 for seronegative RA, or for some other disease.

4.1.3 Clinical data

The clinical data were collected by a systematic review of the electronic medical records of the biobank hospitals in the study. The review was performed by a rheumatologist or an experienced resident in rheumatology participating in the study.

Study data were collected and managed using REDCap electronic data capture tools, hosted at the University of Turku (Harris et al. 2009; Harris et al. 2019).

The data were collected according to a predefined data form, and it included symptoms of the patient, clinical findings, laboratory results, imaging findings, and the medication used by the patient. The exact numeric values of RF and ACPA laboratory results were recorded when available. For some patients, RF and ACPA results were not available because the diagnosis had been made elsewhere or the diagnosis was made before the start of electronic medical records, but the serological status for these patients could be confirmed otherwise from the written medical records. When evaluating the results of imaging studies, the radiologist's initial opinion was relied upon. If this opinion was not available for older imaging studies, the physician's text was relied upon if it specifically mentioned the results. Images were not scored or systematically re-evaluated. The fulfilment of the ACR 1987 and ACR/EULAR 2010 classification criteria for RA were evaluated for each patient. It was also recorded whether the patient was visiting a department of rheumatology or any other speciality when the diagnosis of RA was made.

4.1.3.1 Assessing the accuracy of RA diagnosis

The reviewer assessed the correctness of the RA diagnosis based on a thorough review of the medical records and considering all the information collected during the follow-up period. The reviewer confirmed a diagnosis of RA as true positive based on their opinion of the clinical picture of the disease. It was assessed whether the clinical follow-up was fully compatible with RA and whether the patient had been treated with DMARDs for RA. Also taken into account was whether the patient had been diagnosed with RA by a rheumatologist, a resident working at a department of rheumatology, or an internist, or whether a rheumatologist had later confirmed a diagnosis made elsewhere. The patient was not required to fulfil the ACR 1987 or ACR/EULAR 2010 classification criteria for RA but meeting these criteria was considered in making the decision on the diagnosis (Arnett et al. 1988; Aletaha et al. 2010).

Sometimes the diagnosis clearly deviated from the physician’s record and was considered incorrectly input. Sometimes the follow-up revealed facts suggesting that another disease better explained the patient's symptoms and findings, even if the initial presentation had been consistent with rheumatoid arthritis.

For control patients, the absence of a diagnosis was categorised as a false negative if there was evidence in the medical records that the patient had RA, even though there was no ICD-10 code for RA recorded in the CRHC.

4.2 The validity of systemic sclerosis diagnoses in two university hospitals in Finland (Study II)

4.2.1 Study population

All patients with at least one inpatient or outpatient visit at any speciality with a diagnosis of SSc during 2008–2018 in the discharge registers of two Finnish tertiary referral centres, Turku and Oulu University Hospitals, were included in the study. For a diagnosis of SSc, ICD-10 codes beginning with M34 were used. The study sample included 412 patients. 27 patients were excluded from the final analysis due to insufficient patient chart data available to confirm the diagnosis because the original diagnosis had been made in another hospital district. Final analysis included 385 patients with a diagnosis of SSc. Of these 385 patients, 226 patients had a more specific diagnosis of lcSSc with an ICD-10 code of M34.1. (Figure 2)

The initial diagnosis of SSc was allowed to be made before the first year of the study or in another hospital. The inclusion visit could be at any stage of diagnosis or treatment of SSc or in the context of a visit for another medical condition. The year of the initial SSc diagnosis was collected from the patient records. Length of follow-up was calculated as the time from diagnosis to the last contact at the department of rheumatology.

The patients included in the study were searched from the electronic databases of the participating hospitals. The search was conducted with data mining carried out by the information services of participating hospitals.

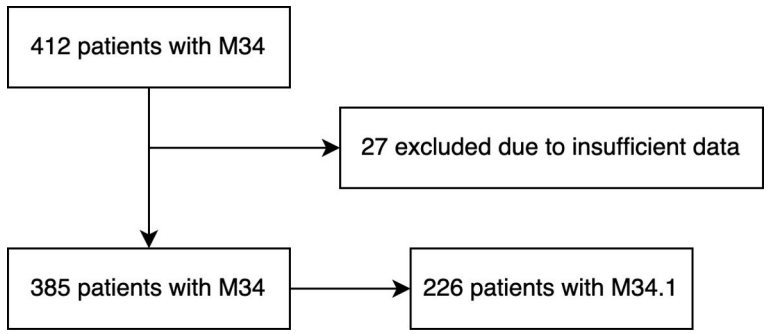

Figure 2. Study flowchart with inclusion and exclusion of patients. M34, ICD-10 code for systemic sclerosis (SSc); M34.1, ICD-10 code for limited cutaneous SSc. Modified from Paltta et al., 2023 (II).

4.2.2 Clinical data

Thorough chart review was performed by a rheumatologist and an experienced resident in internal medicine who collected the data on a predefined form. Study data were collected and managed using REDCap electronic data capture tools, hosted at the University of Turku (Harris et al. 2009; Harris et al. 2019).

Data was gathered on symptoms, clinical findings, NVC changes and other imaging, investigations of specific organ manifestations, autoantibodies of SSc, and other laboratory test results as well as comorbidities. The patient was considered to have Raynaud's phenomenon if the condition was documented in the medical records as self-reported by the patient or witnessed by a physician. For each patient, the number of visits with the inclusion diagnosis was recorded. It was also noted whether the visit when the diagnosis was made was to the department of rheumatology or to some other speciality.

4.2.2.1 Assessing the accuracy of SSc diagnosis

The reviewers evaluated the correctness of the SSc diagnosis, taking into account all the patient record data available from the follow-up period. The requirements for a correct final diagnosis of SSc and early SSc are shown in Table 10.

Table 10. Requirements for a correct final diagnosis of systemic sclerosis (SSc) and early SSc.

Final diagnosis	Requirement
SSc	Patient fulfilled the 2013 revised ACR/EULAR classification criteria for SSc with a score of 9 or higher
Early SSc	Patient fulfilled the 2001 LeRoy and Medsger classification criteria for early SSc

SSc, systemic sclerosis; ACR, American College of Rheumatology; EULAR, European League Against Rheumatism

The final diagnosis of SSc was further divided into dcSSc, lcSSc, SSc overlap syndrome, and ssSSc. A final diagnosis of ssSSc was used if the patient fulfilled the ACR/EULAR criteria but there were no skin manifestations. A diagnosis of SSc was considered incorrectly input if the ICD-10 code of SSc was clearly different from the information and written diagnosis recorded by the physician.

4.3 Differential diagnostics of polymyalgia rheumatica in a university hospital in Finland (Study III)

4.3.1 Study population

All patients with a new diagnosis of PMR recorded as a primary diagnosis with an ICD-10 code M35.3 on at least one inpatient or outpatient visit at any speciality during 2016–2019 were identified from the electronic hospital discharge register of Turku University Hospital in Finland. The identification of the patients was carried out by the information services of the hospital. The study did not include patients with ICD-10 code M31.5 for GCA with PMR.

Turku University Hospital is in the Hospital District of southwest Finland and is a tertiary referral centre for a catchment area of 480,000 people. The patients are admitted to hospital wards or outpatient clinics either via the emergency department or with a referral from other primary or specialist care units.

Figure 3 presents the inclusion and exclusion of patients in the study. The study cohort included 455 patients. For 73 patients, the reviewers were not able to evaluate the correctness of the diagnosis due to insufficient patient record data available, and these patients were excluded from final analysis. For eight patients, the diagnosis was considered incorrectly input in the hospital register, because it was a clear deviation from the written information in the physician's record, and these eight patients were also excluded from the final analysis.

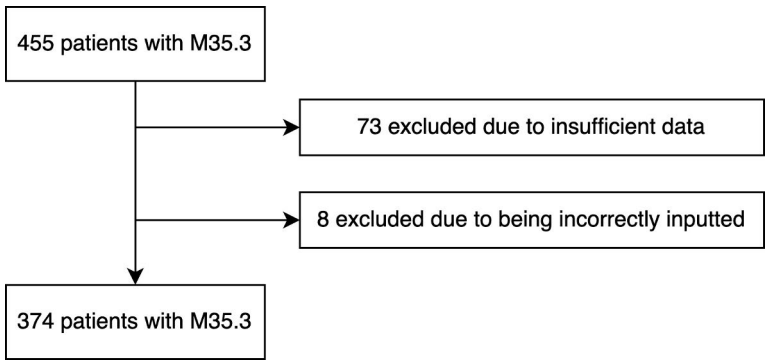

Figure 3. Study flowchart with inclusion and exclusion of the patients. M35.3, ICD-10 code for polymyalgia rheumatica (PMR). Modified from Paltta et al., 2023 (III).

4.3.2 Clinical data

The clinical data were collected by a systematic chart review by a rheumatologist and a medical student having a Bachelor of Medicine, and a standardised protocol was used in the collection process. Data were collected and managed using REDCap electronic data capture tools, hosted at the University of Turku (Harris et al. 2009; Harris et al. 2019).

Information on patients' past and present comorbidities, current symptoms and clinical findings, imaging and laboratory results and maximum dosage of glucocorticoid used were collected. The number of hospital visits with the diagnosis of PMR for a patient were recorded and whether the diagnosis of PMR had been made at a department of rheumatology or in another speciality in the hospital.

When assessing the fulfilment of classification criteria, if an item was not marked as positive in the patient charts, it was counted as negative. However, a negative test result was required to confirm the absence of RF and/or ACPA. The evaluation of ACR/EULAR classification criteria fulfilment was performed without using US, because not all physicians in the Turku University Hospital used US routinely for the evaluation of PMR during the years of the study, and it is possible that negative findings were not reported. We considered typical symptoms for PMR to be pain, tenderness and morning stiffness in the neck, shoulders, upper arms, hips and thighs. The duration of morning stiffness was seldom recorded, so morning stiffness was considered relevant if stated in the records.

In our study, we calculated the duration of symptoms from the start of the symptoms to the inclusion visit with a first recorded PMR diagnosis in the hospital, and the period from the inclusion visit to the last recorded contact in the hospital was the duration of follow-up, even if the diagnosis had been made earlier in some other health care unit.

4.3.2.1 Assessing the accuracy of PMR diagnosis

The validity of every PMR diagnosis was evaluated considering the full clinical follow-up period. The length of this follow-up period had a median of 34 months. A diagnosis of PMR was confirmed to be a true positive if the patient fulfilled at least one set of the five classification criteria (Dasgupta et al. 2012; Bird et al. 1979; Jones et al. 1981; Chuang et al. 1982; Healey 1984), if the thorough clinical follow-up was consistent with PMR, and if the patients' condition was not better explained by some other diagnosis (Table 7). In the case that some other diagnosis was more likely to explain the patient's condition considering the data from the follow-up period, the diagnosis of PMR was not considered correct, even if the original diagnosis of PMR seemed correct at the time and the classification criteria of PMR had been met. If there was uncertainty about the patient's condition or if the reviewers' opinions on

the patient's diagnosis differed, the final decision on the diagnosis was made by the senior rheumatologist. It was recorded how often the initial PMR diagnosis changed after a more thorough diagnostic assessment and during the follow-up period. The study also listed the most common final diagnoses that were initially misdiagnosed as PMR.

4.4 Definitions of an incorrectly input diagnosis and a misdiagnosis

In this dissertation, an incorrectly input diagnosis was defined as a registry diagnosis that had been recorded incorrectly by mistake and was clearly different from the written records of the physician.

A misdiagnosis was defined as a diagnosis that was not considered to be correct, meaning that the data available fit another condition better.

4.5 Ethical considerations and study permissions

All the studies in this dissertation were retrospective and noninterventional, and the patients in the studies were not contacted directly. Therefore, according to Finnish legislation, patient consent or ethical committee approval was not needed. For Study I, The Ethical Committee of Hospital District of Southwest Finland was still consulted, and the committee did not find any possible ethical problems (Dnro 62/1804/2019).

The legal basis for processing personal data was public interest and scientific research (EU General Data Protection Regulation 2016/679 (GDPR), Article 6(1)(e) and Article 9(2)(j); Data Protection Act, Sections 4 and 6) (EUR-Lex. General Data Protection Regulation 2016).

Permissions for Study I were obtained from the Social Insurance Institution of Finland (no. 77/522/2019), the Finnish Institute for Health and Welfare (permission no. THL/1233/ 5.05.00/2019), and the biobank and hospital district of every hospital contributing to the study.

Permissions for Study II were obtained from The Hospital District of Southwest Finland for Turku University Hospital and the Northern Ostrobothnia Hospital District for Oulu University Hospital.

Permissions for Study III were obtained from The Hospital District of Southwest Finland.

4.6 Statistical methods

The results are expressed as medians with interquartile ranges (IQR) for continuous variables and counts with percentages for categorical variables. When comparing differences and calculating p values between different categories, the nonparametric Mann–Whitney U test was used for continuous variables, and Pearson's Chi-square test for categorical variables. Differences were considered statistically significant with two-sided p values <0.05. For all of the predictive statistics, exact 95% confidence intervals (CIs) were calculated.

All statistical analyses were performed using R version 3.6.2 with The R base, descr, dplyr, epiR, ggfortify, stats, stringr, survival, tidyr and vcd packages. (The R Project for Statistical Computing 2024)

4.6.1 Statistical methods – Study I

In this dissertation, PPV represents the proportion of individuals with a recorded diagnosis who truly have the disease and NPV represents the proportion of individuals with no recorded diagnosis who truly do not have the disease. The PPV for a biobank diagnosis of RA supplemented with CRHC and Drug Reimbursement Register data was calculated as a proportion of RA diagnoses confirmed by the researchers out of all investigated registry diagnoses of RA. The NPV for no diagnoses of RA in the registry was calculated as a proportion of people with no signs of RA in patient charts out of all investigated controls with no registry diagnoses of RA.

In this study, likelihood ratios were used to compare the probability of a person with RA having a register diagnosis of RA as compared to someone without RA. These are represented with positive likelihood ratio (PLR) for a registry diagnosis of RA and negative likelihood ratio (NLR) for a lack of a registry diagnosis of RA. The positive likelihood ratio (PLR) and negative likelihood ratio (NLR) were calculated using sensitivity and specificity. Sensitivity was calculated by dividing the number of people who were considered to have RA and had a recorded RA diagnosis (true positives) by the total number of people who were considered to have RA regardless of whether they had a recorded diagnosis of RA or not. Specificity was calculated by dividing the number of people who were considered to not have RA and had no recorded RA diagnosis (true negatives) by the total number of people who were considered to not have RA regardless of whether they had a recorded diagnosis of RA or not. PLR was calculated as sensitivity divided by (1 - specificity) and NLR as (1 - sensitivity) divided by specificity.

In this study, diagnostic accuracy represents the proportion of individuals correctly classified as having or not having RA based on the presence or lack of RA registry diagnosis. Diagnostic accuracy was calculated by dividing the total number

of true positives and true negatives by the total number of all the patients and controls analysed. It should be kept in mind, that diagnostic accuracy is affected by the disease prevalence, and with the same sensitivity and specificity, diagnostic accuracy of a registry diagnosis of RA increases if the RA prevalence decreases.

When available, the information about the patients' positive ACPA test and the right to reimbursement of medicine expenses was also taken into account. In order to measure the agreement between the CRHC diagnosis of RA and the clinical diagnosis of RA, the Cohen's kappa was calculated.

All patients with a diagnosis of either seropositive or seronegative RA were analysed together as one group. In addition, patients with a diagnosis of seropositive RA and patients with a diagnosis of seronegative RA were also analysed as separate groups.

4.6.2 Statistical methods – Study II

The PPV for a registry diagnosis of SSc was calculated as a proportion of registry diagnoses of SSc concurring with diagnoses of SSc confirmed by the researchers.

The PPV results were calculated separately for a registry diagnosis of SSc and for a more specific diagnosis of lcSSc. For both diagnoses, PPV results were calculated separately for diagnoses made in any speciality and for diagnoses made in a department of rheumatology. All the aforementioned PPV results were also calculated in two ways, depending on whether a patient with early scleroderma was considered to have a true scleroderma or not.

4.6.3 Statistical methods – Study III

The PPV for a registry diagnosis of PMR was calculated as a proportion of registry diagnoses of PMR concurring with diagnoses of PMR confirmed by the researchers.

A Kaplan-Meier time-to-event analysis was performed to investigate survival of PMR diagnosis as a function of time, with 95% confidence intervals.

In order to investigate predictors of the change of PMR diagnosis during follow-up, univariate and multivariable logistic regression analyses were performed. These regression analyses consisted of six variables that were selected based on their clinical relevance and because these variables are often included in different classification criteria of PMR.

5 Results

5.1 The validity of rheumatoid arthritis diagnoses in Finnish biobanks (Study I)

Figure 4 presents the patients included in the final analysis and the main results of the study. A total of 233 patients diagnosed with RA were included in the final study analysis. Out of these 233 patients, 118 had a diagnosis of seropositive RA and 115 patients had a diagnosis of seronegative RA. The final analysis also included 250 controls who had no diagnosis of RA recorded in CRHC.

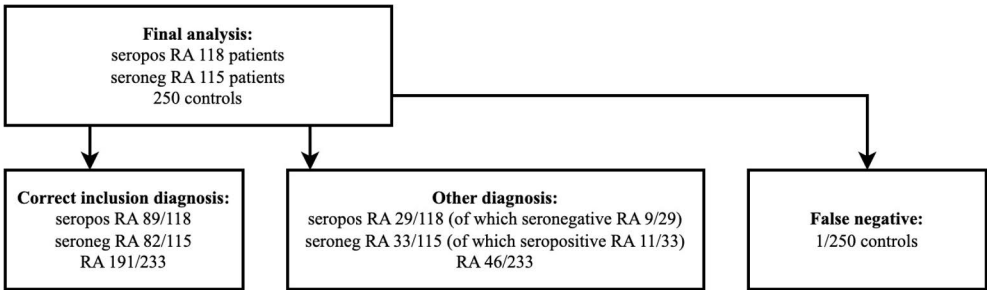

Figure 4. Number of patients in final analysis and results of the study. RA, rheumatoid arthritis; seroneg, seronegative; seropos, seropositive. Modified from Paltta et al., 2023 (I).

The demographic and clinical characteristics of the study population are presented in Table 11. The patients with seropositive RA were 46 years old at the time of diagnosis, and the patients with seronegative RA were a few years older – 54 years. The diagnosis had been made in a department of rheumatology for 90% of the patients with RA, and a clear majority of the patients, 95%, had also been treated in rheumatology department. DMARDs were being used to treat all patients with seropositive RA and 95% of patients with seronegative RA, meaning that 97% of all RA patients were on DMARDs. The cost of DMARDs was reimbursed for 95% (219/230) of all RA patients, but only 66% (152/230) of the RA patients had been granted this entitlement specifically for a diagnosis of RA, and for the 118 patients with seropositive RA, only 52% (61/118) were granted the entitlement specifically for seropositive RA. Out of patients with seropositive RA, 63% had radiographic

changes suggestive of RA at the end of follow-up compared with 30% of patients with seronegative RA. Development of radiographic changes of RA was therefore more common in seropositive RA.

Table 11. Demographic and clinical characteristics of the study patients with RA.

	N with data	All RA	Seropositive RA	Seronegative RA
Number of patients	233	233	118	115
Female (%)	233	160 (69%)	75 (64%)	85 (74%)
Year of diagnosis (range)	222	2005 (1960–2018)	2001 (1960–2018)	2008 (1969–2018)
Age at diagnosis in years [IQR]	224	50.0 [40.0–59.0]	46.0 [36.0–56.0]	54.0 [45.0–61.1]
Follow-up in years [IQR]	220	11.0 [4.3–22.0]	16.0 [4.5–30.0]	9.0 [4.0–16.0]
Diagnosed in rheumatology department (%)	188	169 (90%)	79 (91%)	90 (89%)
Treated in rheumatology department (%)	228	209 (95%)	103 (94%)	106 (95%)
Nr. of visits seropositive RA [IQR]	233	5.0 [0.0–25.0]	20.0 [6.0–31.5]	0 [0.0–3.0]
Nr. of visits seronegative RA [IQR]	233	2.0 [0.0–13.0]	0.0 [0.0–0.0]	11.0 [4.5–23.0]
Treated with DMARDs (%)	221	215 (97%)	110 (100%)	105 (95%)
Reimbursement for DMARDs (inclusion diagnosis specific) (%)	230	152 (66%)	61 (52%)	75 (66%)
Reimbursement for DMARDs (%)	230	219 (95%)	111 (95%)	108 (96%)
EULAR classification criteria positive at diagnosis (%)	136	79 (58%)	47 (72%)	32 (45%)
ACR classification criteria positive at diagnosis (%)	128	78 (61%)	41 (67%)	37 (55%)
ACR or EULAR classification criteria positive at diagnosis (%)	128	95 (74%)	49 (80%)	46 (69%)
ACR or EULAR classification criteria positive ever (%)	195	164 (84%)	90 (89%)	74 (79%)
Highest RF ever [IQR]	189	13.0 [6.0–73.0]	75.5 [16.5–205.0]	9.0 [0.0–13.5]
Highest ACPA ever [IQR]	166	1.45 [0.0–126.0]	129.0 [7.0–340.0]	0.8 [0.0–1.6]
ACPA positive ever (%)	166	63 (38%)	55 (75%)	8 (9%)
Erosions in radiographs at diagnosis (%)	143	31 (22%)	19 (32%)	12 (14%)
Erosions in radiographs ever (%)	210	98 (47%)	66 (63%)	32 (30%)

Continuous variables are expressed as medians with interquartile ranges, and categorical variables are described as counts with percentages. ACPA, anti-citrullinated protein antibody; ACR, American College of Rheumatology; DMARDs, disease-modifying anti-rheumatic drugs; EULAR, European League Against Rheumatism; RA, rheumatoid arthritis; RF, rheumatoid factor. Modified from Paltta et al, 2023 (I).

5.1.1 Patients with RA, seropositive and seronegative diagnoses combined.

If a patient was diagnosed with either seropositive or seronegative RA on at least one visit in the hospital, the PPV for this diagnosis of RA recorded in CRHC was 82% (191/233). The PPV depended on the number of visits on which an RA diagnosis was recorded, and it increased with the number of visits. For at least two visits the PPV was 85% (189/ 222), for at least five visits 89% (175/197), and for at least 10 visits 90% (137/152) (Figure 5, Table 12).

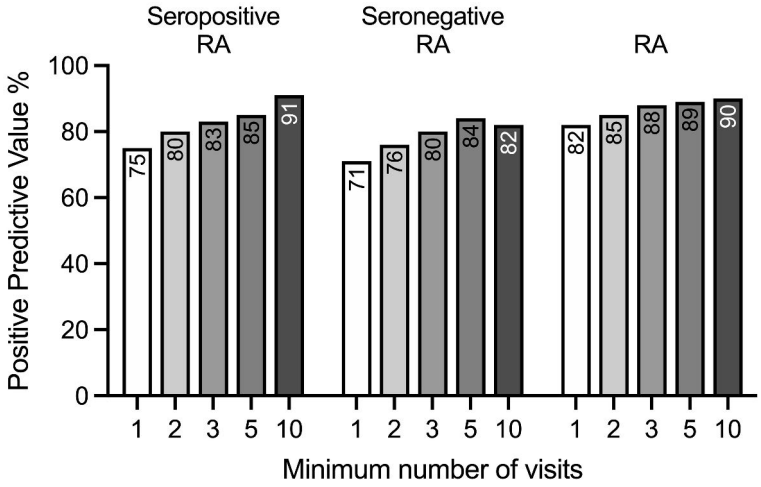

Figure 5. Positive predictive value of rheumatoid arthritis (RA) diagnoses in the Finnish Care Register for Health Care compared to chart review, based on minimum number of visits with the diagnosis. Modified from Paltta et al., 2023 (I).

Although the PPV for a diagnosis of RA increased with the number of visits, longer follow-up in years did not increase the PPV. For RA patients with a follow-up period of at least one year, the PPV was 77% (159/207), for at least two years follow-up the PPV was 77% (146/190), for at least five years 75% (120/161) and for at least ten years 78% (93/120) (Figure 6).

Table 12. Agreement between CRHC diagnosis of RA in biobank patients and diagnoses according to chart review.

	All RA	Seropositive RA	Seronegative RA
At least one visit in CRHC with RA			
PPV (95% CI)	82% (76%–87%)	75% (67%–83%)	71% (62%–79%)
NPV (95% CI)	100% (98%–100%)	100% (98%–100%)	100% (98%–100%)
PLR (95% CI)	6.89 (5.21–9.12)	9.48 (6.71–13.39)	8.44 (6.12–11.64)
NLR (95% CI)	0.01 (0.00–0.04)	0.01 (0.00–0.09)	0.01 (0.00–0.10)
Accuracy (95% CI)	0.91 (0.88–0.93)	0.92 (0.89–0.94)	0.91 (0.87–0.93)
Kappa (95% CI)	0.82 (0.77–0.87)	0.80 (0.73–0.87)	0.77 (0.69–0.84)
At least one visit in CRHC and entitlement for reimbursement of DMARDs with inclusion diagnosis			
PPV (95% CI)	89% (83%–94%)	93% (84%–98%)	79% (68%–87%)
NPV (95% CI)	100% (98%–100%)	100% (98%–100%)	100% (98%–100%)
PLR (95% CI)	16.44 (10.22–26.44)	62.16 (23.50–164.43)	16.29 (10.12–26.22)
NLR (95% CI)	0.01 (0.00–0.05)	0.02 (0.00–0.12)	0.02 (0.00–0.12)
Accuracy (95% CI)	0.96 (0.93–0.98)	0.98 (0.96–0.99)	0.95 (0.92–0.97)
Kappa (95% CI)	0.91 (0.87–0.95)	0.95 (0.90–0.99)	0.84 (0.77–0.91)
At least one visit in CRHC and positive ACPA			
PPV (95% CI)	98% (91%–100%)	96% (87%–100%)	
NPV (95% CI)	100% (98%–100%)	100% (98%–100%)	
PLR (95% CI)	246.03 (34.78–1740.19)	123.18 (30.96–490.03)	
NLR (95% CI)	0.02 (0.00–0.11)	0.02 (0.00–0.13)	
Accuracy (95% CI)	0.99 (0.98–1.00)	0.99 (0.97–1.00)	
Kappa (95% CI)	0.98 (0.95–1.01)	0.97 (0.93–1.00)	

ACPA, anti-citrullinated protein antibody; CRHC, the Finnish Care Register for Health Care; DMARDs, disease-modifying anti-rheumatic drugs; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value; RA, rheumatoid arthritis. Modified from Paltta et al., 2023 (I).

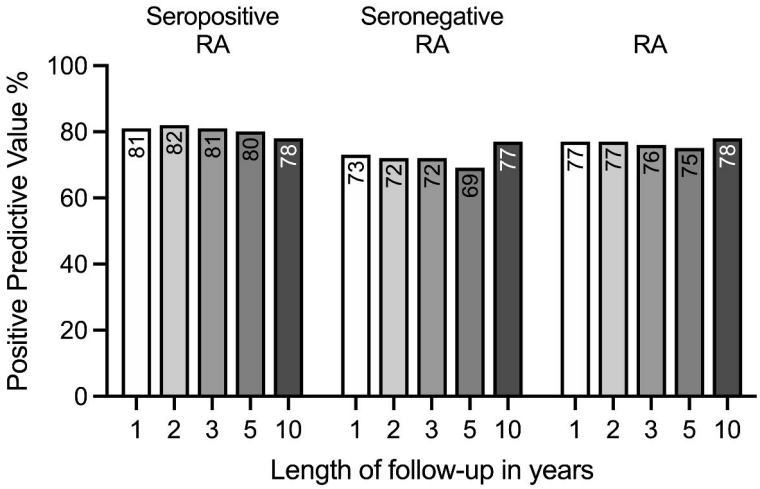

Figure 6. Positive predictive value of rheumatoid arthritis (RA) diagnoses in the Finnish Care Register for Health Care, based on length of follow-up in years. Previously unpublished.

For those patients who had been diagnosed in a department of rheumatology, the PPV was 89% (150/169). The PPV rose slightly with the number of visits. For at least two visits the PPV was 89% (149/167), for at least five visits 91% (138/152), and for at least 10 visits 91% (102/112) (Figure 7).

Laboratory test results have been included in the patient selection criteria for some biobank studies. In our study, we also analysed the effect of ACPA status of the patient for the validity of the RA diagnosis. For ACPA-positive patients, the PPV for a diagnosis of RA was 98% (62/63) (Table 12).

The fulfilment of the ACR 1987 and ACR/EULAR 2010 classification criteria for RA was analysed for those patients who were categorised as having RA. There was enough data for analysis for 173 out of 191 patients, and 92% (160/173) of these patients fulfilled either one or both of these classification criteria.

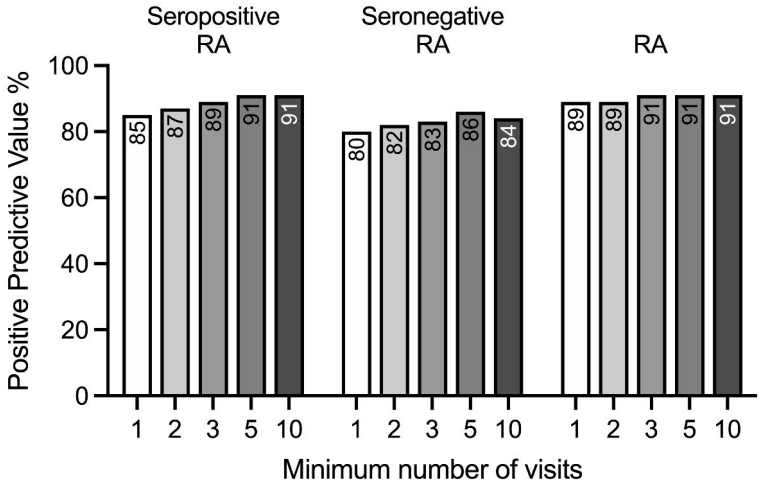

Figure 7. Positive predictive value of rheumatoid arthritis (RA) diagnoses in the Finnish Care Register for Health Care for patients diagnosed in a department of rheumatology. Previously unpublished.

5.1.2 Patients with seropositive RA

If a patient was diagnosed with seropositive RA on at least one visit to the hospital, the PPV for this diagnosis of seropositive RA recorded in CRHC was 75% (89/118). The PPV depended on the number of visits in which a seropositive RA diagnosis was recorded, and it increased with the number of visits. For at least two visits the PPV was 80% (88/110), for at least five visits 85% (82/96), and for at least 10 visits 91% (70/77) (Figure 6, Table 12).

For those seropositive RA patients who had been diagnosed in a department of rheumatology, the PPV was 85% (67/79). The PPV rose with the number of visits. The PPV was 87% (67/77) for at least two visits, 91% (62/68) for at least five visits, and 91% (51/56) for at least 10 visits (Figure 7).

Among patients who had a diagnosis of seropositive RA and who were granted entitlement for the reimbursement of DMARDs specifically for seropositive RA, the PPV for at least one visit was 93% (57/61), for at least two visits 93% (57/61), for at least five visits 96% (53/55), and for at least 10 visits 96% (44/46) (Figure 8, Table 12).

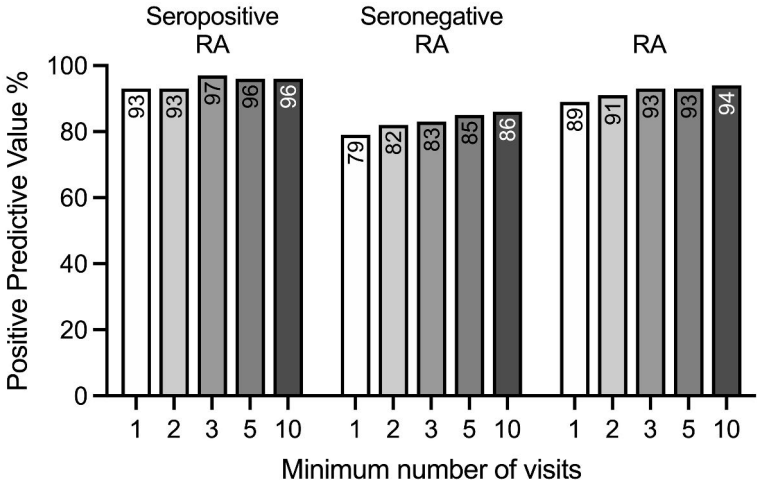

Figure 8. Positive predictive value of rheumatoid arthritis (RA) diagnosis in patients with reimbursement for disease-modifying antirheumatic drugs (specifically for inclusion diagnosis). Modified from Paltta et al., 2023 (I).

Among patients who had a diagnosis of seropositive RA and who were granted entitlement for reimbursement of DMARDs with a less specific reimbursement code of 202, including CTDs, RA, and comparable diseases, the PPV for at least one visit was 79% (88/111), for at least two visits 84% (87/104), for at least five visits 89% (81/91), and for at least 10 visits 92% (69/75) (Figure 9).

Patients were considered ACPA-positive if ACPA was greater than the upper limit of normal (ULN) and high ACPA-positive if ACPA was three or more times the ULN. For ACPA-positive patients the PPV for seropositive RA diagnosis was 96% (53/55) (Table 12), and for high ACPA positive patients the PPV was 98% (52/53). If a patient with a diagnosis of seropositive RA had RF greater than the ULN, the PPV for seropositive RA diagnosis was 92% (58/63), and the PPV was 94% (49/52) if RF was three or more times the ULN.

The fulfilment of the ACR 1987 and ACR/EULAR 2010 classification criteria for RA was also analysed for patients categorised as having seropositive RA, and 97% (83/86) of these patients fulfilled either one or both of these classification criteria.

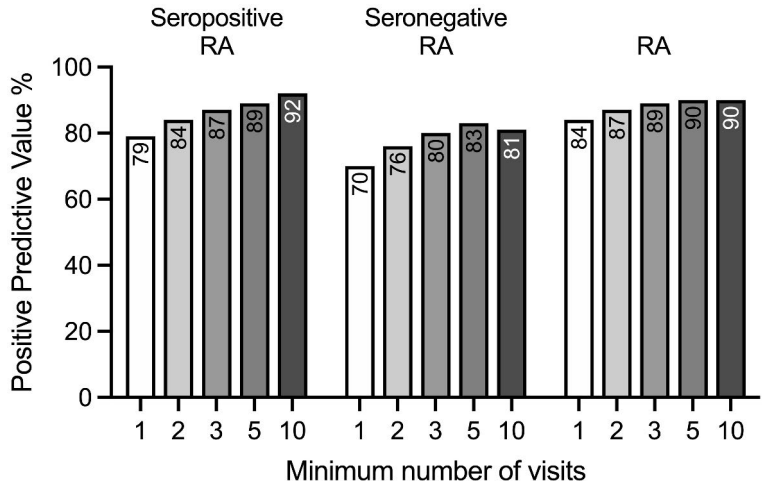

Figure 9. Positive predictive value of RA diagnosis in patients with reimbursement for DMARDs (reimbursement code 202, the Social Insurance Institution of Finland reimbursement code for CTDs, RA and comparable diseases). Modified from Paltta et al., 2023 (I).

5.1.3 Patients with seronegative RA

If a patient was diagnosed with seronegative RA on at least one visit in the hospital, the PPV for this diagnosis of seronegative RA recorded in CRHC was 71% (82/ 115). The PPV rose with the number of visits: PPV 76% (81/106) for at least two visits, PPV 84% (72/86) for at least five visits, and PPV 82% (49/60) for at least 10 visits (Figure 6, Table 12).

For diagnoses made in a department of rheumatology, the PPV was 80% (72/90). The PPV rose with the number of visits. For at least two visits the PPV was 82% (71/87), for at least five visits 86% (63/73), and for at least 10 visits 84% (41/49) (Figure 7).

Among patients who had a diagnosis of seronegative RA and were granted entitlement for reimbursement of DMARDs specifically for seronegative RA, the PPV for at least one visit was 79% (59/75), for at least two visits 82% (59/72), for at least five visits 85% (53/62), and for at least 10 visits 86% (37/43) (Figure 8, Table 12).

Among seronegative RA patients who were granted entitlement for reimbursement of DMARDs with a less specific reimbursement code of 202, which includes CTDs, RA, and comparable diseases, the PPV values for these patients with at least one, two, five, and 10 visits were 70% (76/108), 76% (76/100), 83% (67/81), and 81% (47/58), respectively (Figure 9).

For patients categorised as having seronegative RA, 88% (61/69) of the patients fulfilled either one or both of the ACR 1987 and ACR/EULAR 2010 classification criteria.

5.1.4 Control group

After a thorough review of the patient medical records, out of the 250 controls in final analysis, only one was found to have information suggesting a diagnosis of RA for the patient (NPV 100%, 249/250). This patient had been diagnosed with seronegative RA in the local hospital, but this diagnosis was not recorded in the CRHC database.

5.1.5 Incorrect diagnoses

An incorrect CRHC diagnosis of RA was found upon follow-up in 62 out of the 233 biobank patients in final analysis. In 20 of these 62 patients, the only diagnostic error was seropositive RA being recorded as seronegative or vice versa. Out of these 62 patients, 42 were determined to have no RA at all, but they had a variety of other medical conditions, rheumatological and unrelated, which are specified in Table 13.

A diagnosis was considered to be incorrectly input and not misdiagnosed if it clearly deviated from the physician's written record in the patient charts. All 29 of the incorrect seropositive RA diagnoses and 67% (22/33) of the incorrect seronegative RA diagnoses were incorrectly input. Rheumatology department recordings accounted for 39% (20/51) of these incorrect entries, surgery department recordings for 24% (12/51), physiotherapy department for 8% (4/51) and miscellaneous departments accounted for the remaining 29% (15/51). Misdiagnoses accounted for 18% (6/33) of the diagnoses of seronegative RA, and 15% (5/33) of seronegative RA diagnoses changed during the follow-up period even if they had seemed valid at the time the diagnosis was made (Figure 4).

Table 13. Final diagnosis of the patients if other than RA and reason for the change of diagnosis. Modified from Paltta et al, 2023 (I).

Number of patients	Final diagnosis	Reason for the change of diagnosis
6	Juvenile idiopathic arthritis	6 incorrectly input
5	Gout (with arthrosis in one patient)	3 incorrectly input, 2 new data during follow-up
5	Unspecified oligoarthritis	3 misdiagnoses, 2 incorrectly input
5	Unspecified polyarthritis (with arthrosis in one patient)	3 misdiagnoses, 2 incorrectly input,
3	Fibromyalgia and arthrosis	2 misdiagnoses, Incorrect input
3	Psoriatic arthritis	3 incorrectly input
2	Adult-onset Still disease	2 incorrectly input
2	Polymyalgia rheumatica	1 misdiagnosis and 1 incorrectly input
1	Ankylosing spondylitis	Incorrectly input
1	Arthropathy of Crohn disease	Incorrectly input
1	Complex regional pain syndrome of the hand	Incorrectly input
1	Dermatopolymyositis and Sjögren's syndrome	Incorrectly input
1	Erosive arthrosis	Incorrectly input
1	Fibrotic dysplasia of the bone	Incorrectly input
1	Idiopathic inflammation of the orbita	Incorrectly input
1	Juvenile ankylosing spondylitis	Incorrectly input
1	Mixed connective tissue disease	Incorrectly input
1	Myalgia	Misdiagnosis
1	Reactive arthritis	Incorrectly input
1	Sjögren's syndrome and arthrosis	Incorrectly input

5.2 The validity of systemic sclerosis diagnoses in two university hospitals in Finland (Study II)

Study population characteristics are presented in Table 14 and Figure 10. A total of 385 patients with a diagnosis of SSc (M34) were analysed, and a subanalysis was made for 226 patients with a more specific diagnosis of lcSSc (M34.1). Most of the patients were women: 79% of the patients with SSc and 86% of the patients with lcSSc. The SSc patients were diagnosed at the median age of 53 years and lcSSc patients at 55 years, and follow-up time was a median of 5 years. Out of the SSc patients, 91% had symptoms of Raynaud's syndrome and 68% had changes in videocapillaroscopy at the time of diagnosis. For the lcSSc patients, these percentages were a little higher with 94% of the patients having had Raynaud's symptoms and 75% of the patients having had videocapillaroscopy changes. 99% of

the SSc patients and all the lcSSc patients were ANA positive, and 80.9% of the SSc patients and 89% of the lcSSc patients had specific autoantibodies of SSc. 82% of the patients were ACR/EULAR classification criteria positive, and 12% fulfilled only the LeRoy and Medsger criteria for early SSc.

Table 14. Demographic and clinical characteristics of the study patients.

	M34 patients				M34.1 patients			
	N	Any department	N	Rheumato-logy	N	Any department	N	Rheumato-logy
Number of patients	385	385	279	279	226	226	198	198
Female (%)	385	304 (79%)	279	227 (81%)	226	194 (86%)	198	172 (87%)
Year of diagnosis [IQR]	381	2011 [2005–2015]	276	2012 [2007–2015]	226	2012 [2006–2015]	198	2013 [2008–2015]
Age at diagnosis in years [IQR]	381	53 [43–65]	276	54 [45–66]	226	55 [46–66]	198	55 [46–67]
Positive antinuclear antibodies (%)	267	265 (99%)	235	235 (100%)	204	204 (100%)	185	185 (100%)
Autoantibodies of SSc (%)	303	245 (81%)	262	215 (82%)	219	196 (89%)	196	175 (89%)
Videocapillaroscopy changes	248	168 (68%)	218	156 (72%)	179	134 (75%)	161	126 (78%)
Length of follow-up at rheumatology department in years [IQR]	300	5.0 [3.0–11.0]	261	5.0 [3.0–9.0]	219	5.0 [3.0–11.0]	197	5.0 [3.0–9.0]
Raynauds symptom at diagnosis (%)	305	276 (91%)	262	239 (91%)	223	210 (94%)	198	188 (95%)
Smoking before diagnosis (%)	253	98 (39%)	219	90 (41%)	180	71 (39%)	162	66 (41%)
ACR/EULAR classification criteria positive (%)	308	253 (82%)	265	217 (82%)	223	180 (81%)	198	160 (81%)
Deceased during follow-up	384	94 (25%)	279	70 (25%)	226	46 (21%)	198	41 (21%)

M34, ICD-10 code for systemic sclerosis; M34.1, ICD-10 code for limited cutaneous systemic sclerosis; ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; N, number with data; Autoantibodies of SSc: anti-centromere, anti-scl 70 and anti-RNA polymerase III antibodies. Modified from Paltta et al, 2023 (II).

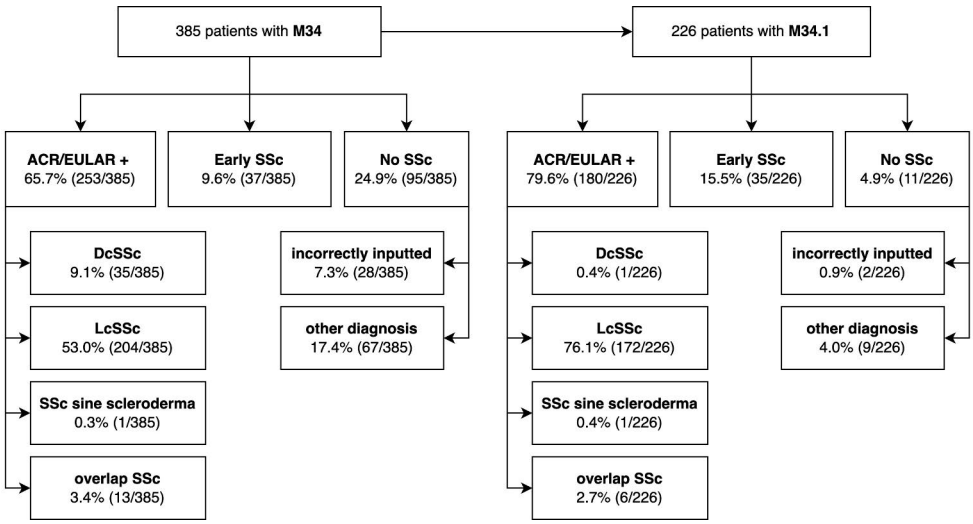

Figure 10. Final diagnoses for patients with an inclusion diagnosis of SSc or LcSSc. SSc, systemic sclerosis; LcSSc, limited cutaneous systemic sclerosis; DcSSc, diffuse cutaneous systemic sclerosis; M34, ICD-10 code for SSc; M34.1, ICD-10 code for LcSSc; ACR/EULAR classification criteria for SSc; early SSc, patients meeting LeRoy and Medsger classification criteria for early SSc. Modified from Paltta et al., 2023 (II).

Of the 385 patients with an inclusion diagnosis of SSc made in any department, 9.1% were classified as dcSSc, 53.0% as lcSSc, 0.3% as SSc sine scleroderma, 3.4% as SSc overlap syndrome, and 9.6% as early SSc. Some other diagnosis was found in 17.4% of the patients, and in 7.3% the diagnosis clearly differed from the physician's written record and was classified as incorrectly input. Other diagnoses that were the most frequent included systemic lupus erythematosus in 11% of the patients with some other diagnosis, morphea of the skin in 9%, RA in 9%, graft-versus-host disease of the skin in 7%, Sjögren's syndrome in 7%, mixed connective tissue disease in 7%, vasculitis in 7%, and polymyalgia rheumatica in 4% of the patients. The PPV for SSc diagnosis was 66% (95% CI 61%–70%), which increased to 75% (95% CI 71%–80%) if also early SSc was considered true SSc (Figure 10, Figure 11).

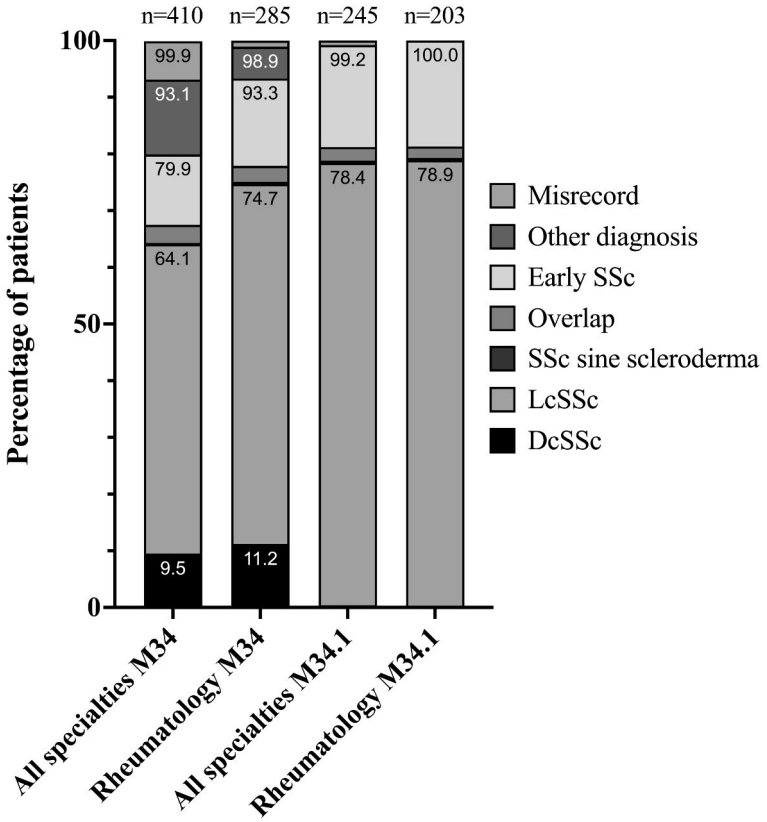

Figure 11. Percentages of the final diagnoses, divided by the inclusion diagnosis and by the speciality where the diagnosis was recorded. SSc, systemic sclerosis; LcSSc, limited cutaneous systemic sclerosis; DcSSc, diffuse cutaneous systemic sclerosis; M34, ICD-10 code for SSc; M34.1, ICD-10 code for LcSSc. Previously unpublished.

From an analysis of the 279 patients who had been diagnosed with SSc in the rheumatology department (Table 14), we found that 10.8% had dcSSc, 63.8% had lcSSc, 0.4% had SSc sine scleroderma, 2.9% had SSc overlap syndrome, and 12.2% had early SSc. 9.0% of the patients were considered to have another diagnosis, and 1.1% of the patients had an incorrectly input diagnosis. The PPV for SSc diagnosis was 78% (95% CI 73%–83%), and 90% (95% CI 86%–93%) if early SSc was also counted as SSc (Figure 10, Figure 11).

There were 226 patients diagnosed with lcSSc (ICD-10 code M34.1) (Table 14). Of these patients, 0.4% had dcSSc, 76.1% had lcSSc, 0.4% had SSc sine scleroderma, 2.7% had SSc overlap syndrome, and 15.5% had early SSc. Additionally, 4.0% of the patients had a different diagnosis, and in 0.9% of cases, the diagnosis was input incorrectly. The PPV of a SSc diagnosis was 80% (95% CI 74%–85%), and when early SSc was included, the PPV increased to 95% (95% CI

92%–98%) (Figure 10, Figure 11). For rheumatology-diagnosed lcSSc, the PPV was 81% (95% CI 75%–86%), and with early SSc included, the PPV was 97% (95% CI 95%–99%).

For patients with specific autoantibodies of SSc, the PPV of a diagnosis of M34 was 83% (204/245) for at least one visit with this diagnosis, and 81% (158/196) for M34.1. These PPV values were 97% (237/245) and 97% (191/196), respectively, if early SSc was also considered true SSc.

When diagnoses made solely in a department of rheumatology for patients with specific autoantibodies of SSc were analysed, PPV was 83% (179/215) for M34 and 81% (141/175) for M34.1. With early SSc also included in true SSc, PPV values were 97% (209/215) for M34 and 98% (171/175) for M34.1.

5.3 Differential diagnostics of polymyalgia rheumatica in a university hospital in Finland (Study III)

The final analysis included 374 patients. Table 15 presents demographic and clinical characteristics of the final study population. 57.2% of the study population were female with a median age of 70 at diagnosis. The majority of patients, 79.0%, were diagnosed at a department of rheumatology, and 81.0% were treated there. The patients were followed for a median duration of 34.0 months (IQR 21.0–50.0).

Figure 12 depicts the main results of the study. Of the 374 patients, there were 245 (65.5) patients whose diagnosis of PMR was confirmed upon follow-up, while for 129 patients (34.5%), the diagnosis of PMR was not supported.

The most frequent conditions initially misdiagnosed as PMR were inflammatory arthritides 34.9% (45/129) and musculoskeletal disorders caused by repetitive strain or degeneration 13.2% (17/129). Among other diagnoses, infection accounted for 9.3% (12/129), malignancy for 9.3% (12/129), giant cell vasculitis for 6.2% (8/129), other vasculitis for 6.2% (8/129), other rheumatological disease for 5.4% (7/129), fibromyalgia or other chronic pain syndromes for 3.9% (5/129), gout or other crystal arthropathies for 1.6% (2/129), endocrinological disease for 1.6% (2/129) and other or unknown diagnoses for 10.9% (14/129) (Figure 12).

Table 15. Demographic and clinical characteristics of the study patients.

	N with data	All	PMR diagnosis supported	PMR diagnosis not supported	p
Number of patients (%)	374	374	245 (65.5)	129 (34.5)	
Female (%)	374	214 (57.2)	139 (56.7)	75 (58.1)	0.794
Age at diagnosis, years [IQR]	374	70.0 [64.0–77.0]	71.0 [64.0–78.0]	69.0 [62.0–77.0]	0.286
Year of first diagnosis [IQR]	370	2017 [2016–2018]	2017 [2016–2018]	2017 [2016–2018]	0.082
Diagnosed in rheumatology department (%)	373	294 (79.0)	211 (86.0)	83 (65.0)	< 0.0001*
Treated in rheumatology department (%)	374	304 (81.0)	205 (84.0)	99 (77.0)	0.102
Number of visits [IQR]	374	6 [3–12]	7 [3–11]	6 [2–13]	0.489
Symptom duration, weeks [IQR]	282	10.5 [5.0–20.0]	12.0 [6.0–20.0]	7.0 [4.0–16.0]	0.016*
Elevated CRP or ESR (%)	369	334 (90.5)	230 (94.3)	104 (83.2)	< 0.001*
Full symptom response to GCs (%)	341	195 (57.2)	154 (66.4)	41 (37.6)	< 0.0001*
Full inflammatory value response to GCs (%)	325	190 (58.5)	152 (67.9)	38 (37.6)	< 0.0001*
Length of follow up, months [IQR]	374	34.0 [21.0–50.0]	33.0 [20.0–48.0]	36.0 [23.0–54.0]	0.011*
Bird criteria + (%)	365	274 (75.0)	208 (84.9)	66 (55.0)	< 0.0001*
Jones and Hazleman criteria + (%)	365	99 (27.1)	88 (35.9)	11 (9.2)	< 0.0001*
Chuang and Hunder criteria + (%)	365	113 (31.0)	91 (37.1)	22 (18.3)	< 0.001*
Healey criteria + (%)	365	248 (67.9)	198 (80.8)	50 (41.7)	< 0.0001*
ACR/EULAR criteria + (%)	365	209 (57.3)	170 (69.4)	39 (32.5)	< 0.0001*

Data are shown as count (percentage) for categorical variables and median (interquartile range) for continuous variables. PMR, polymyalgia rheumatica; GC, glucocorticoids; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ACR, American College of Rheumatology; EULAR, European League Against Rheumatism; +, criteria fulfilled. *Statistically significant difference ($p<0.05$). Modified from Paltta et al, 2023 (III).

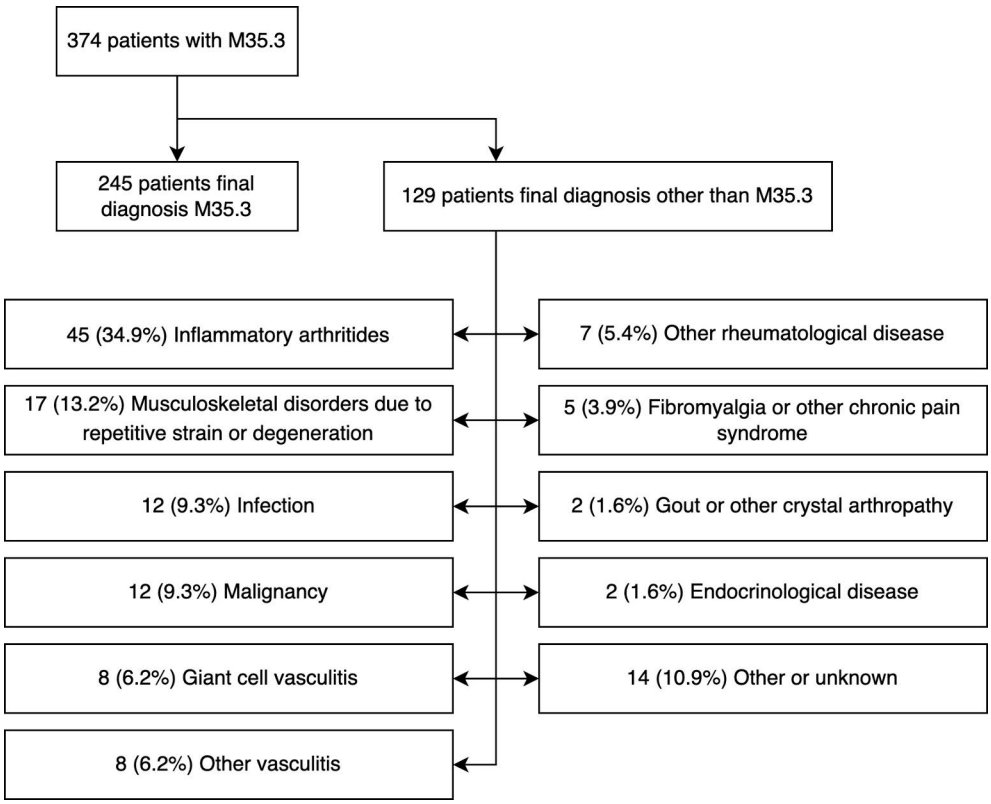

Figure 12. Final diagnosis for patients with an inclusion diagnosis of polymyalgia rheumatica (PMR). M35.3, ICD-10 code for PMR. Modified from Paltta et al., 2023 (III).

The most common group of inflammatory arthritides found in study patients was RA with 69% (31/45) of the patients with inflammatory arthritides, and 32% (10/31) of the patients with RA had seropositive RA. Unspecified arthritis was found in 20% (9/45) of the patients with inflammatory arthritides and the remaining 11% (5/45) of the patients were individual cases of systemic lupus erythematosus, psoriatic arthritis, juvenile idiopathic arthritis, reactive arthritis and arthritis due to Crohn's disease.

Colon carcinomas were the largest group of malignancies with 33% (4/12) of the patients, lymphomas accounted for 25% (3/12), and the remaining cases were individual instances of other malignancies: breast, uterus, kidney, peritoneal carcinosis, and chronic lymphocytic leukaemia. In patients with malignancies, for 75% (9/12) the symptoms were deemed paraneoplastic and not caused by the malignancy itself. Four patients with paraneoplastic symptoms received active treatment for their malignancy, and after treatment, they all became asymptomatic.

The patients had experienced symptoms for a median of 10.5 weeks before their first visit to the university hospital where they were diagnosed with PMR. If the diagnosis changed during follow-up, within six months 76.6% (95/124) of the diagnoses had changed, within one year 86.3% (107/124) and within two years of follow-up 94.3% (117/124) of the diagnoses had changed (Figure 13). Malignancies were discovered within a median time of 8.5 weeks (IQR 1.0-16.0) from the PMR diagnosis.

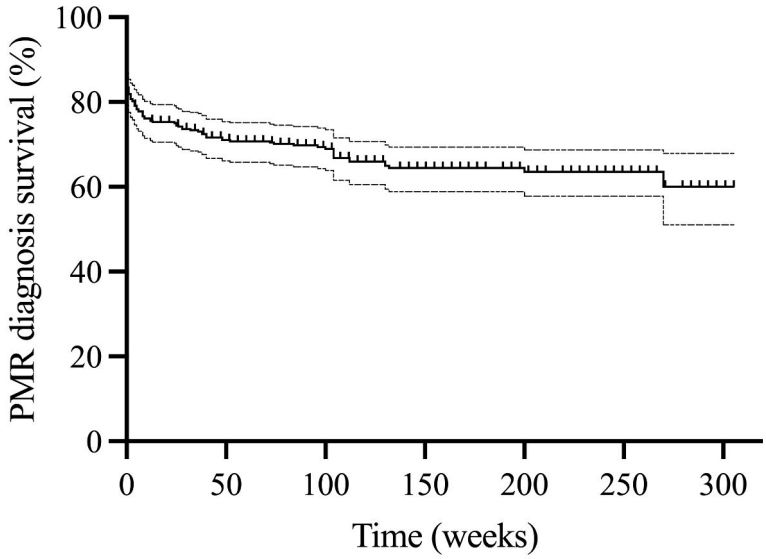

Figure 13. Kaplan-Meier survival curve for change of PMR diagnosis as a function of time, with 95% confidence intervals. Previously unpublished.

If the patient met the 2012 ACR/EULAR PMR classification criteria, the diagnosis was deemed correct in 81.3% (170/209) of the patients compared to 45.5% (75/165) if the criteria were not fulfilled ($p<0.0001$). 75.9% (208/274) of the patients meeting 1979 Bird criteria, 88.9% (88/99) meeting 1981 Jones and Hazleman criteria, 80.5% (91/113) meeting 1982 Chuang and Hunder criteria and 79.8% (198/248) meeting 1984 Healey criteria had PMR as a final diagnosis (Figure 14).

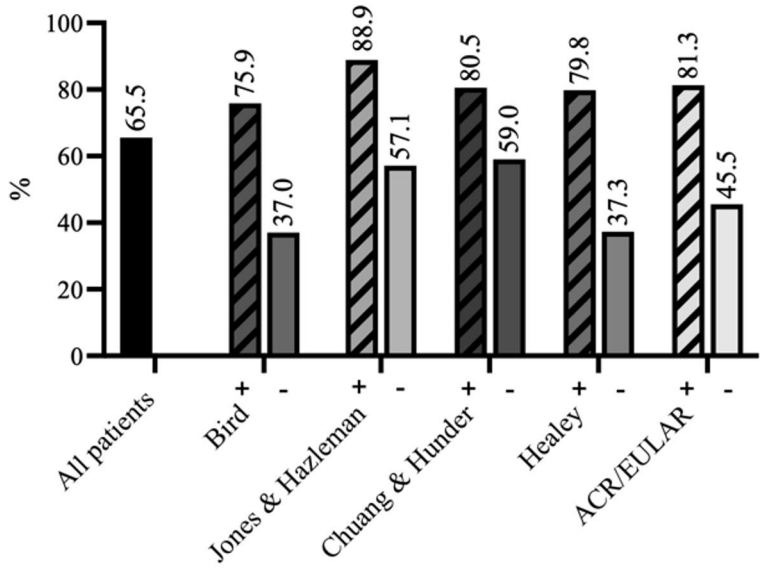

Figure 14. Percentage of patients accurately diagnosed with polymyalgia rheumatica (PMR), out of all patients and divided by fulfilment of classification criteria. +, Classification criteria fulfilled; -, classification criteria not fulfilled. Reprinted from the original publication III with permission from the copyright holders.

If the patient did not fulfil the 2012 ACR/EULAR PMR classification criteria, 9.7% (16/165) had a disease requiring a completely different treatment than PMR; 6.7% (11/165) of the patients had an infection, and 3.0% (5/165) had a malignancy. Among the patients who did meet the classification criteria, infection was found in only 0.5% (1/209) of the patients and malignancy in 3.3% (7/209), making a total of 3.8% (8/209) of the patients.

The PMR diagnoses persisted in 69% (230/334) of the patients with elevated inflammatory values at diagnosis and in 40% (14/35) of the patients with normal inflammatory values.

The accuracy of the PMR diagnosis was 65.5% (245/374) for patients diagnosed in any department compared to 71.8% (211/294) for those diagnosed in the rheumatology department. In contrast to our study on the validity of RA diagnoses in Finnish biobank patients, the accuracy of the diagnosis of PMR did not improve with the number of visits during which the diagnosis was made. The accuracy for one visit was 65.5% (245/374), for two visits 67.2% (219/326), for five visits 68.4% (158/231), and for 10 visits 66.1% (82/124) (Figure 15).

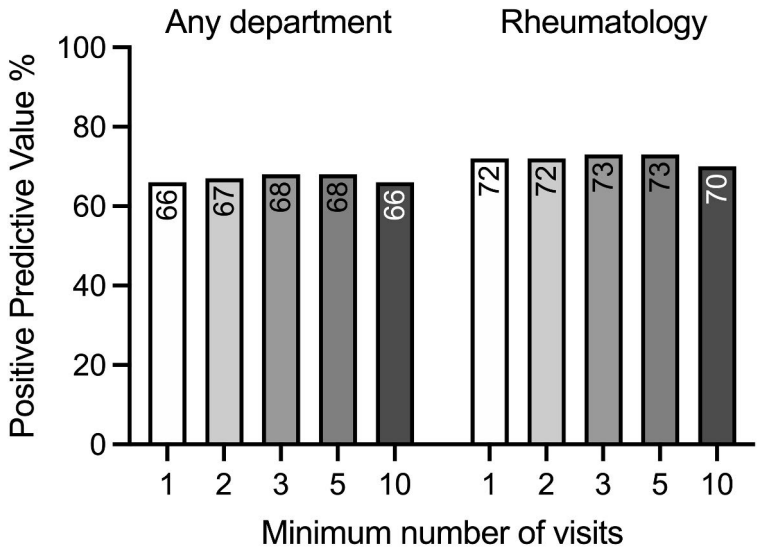

Figure 15. Percentage of patients accurately diagnosed with polymyalgia rheumatica (PMR) at different numbers of visits, by department in which the diagnosis was made. Previously unpublished.

Compared to the patients whose diagnosis changed during follow-up, those whose diagnosis remained unchanged had a longer duration of symptoms before diagnosis ($p=0.016$), their CRP or ESR values were elevated more frequently ($p<0.001$), and they had achieved a full response to symptoms ($p<0.0001$) and inflammatory values ($p<0.0001$) with glucocorticoid treatment more frequently. These patients with a confirmed diagnosis of PMR were also more frequently diagnosed in the rheumatology department ($p<0.0001$) and had longer follow-up periods ($p=0.011$) (Table 15).

The results of a multivariable logistic regression analysis indicate that a change in the diagnosis of PMR was statistically significantly predicted by the absence of morning stiffness (odds ratio (OR) of 1.75, 95% confidence interval (CI) 1.03–2.94), the absence of any other typical PMR symptoms (OR 7.45, 95% CI 2.65–24.53), and incomplete resolution of symptoms with glucocorticoid treatment (OR 3.01, 95% CI 1.83–5.00). When the patient presented with normal inflammatory values, OR for the change of PMR diagnosis was 2.27 (95% CI 0.98–5.24). Gender or younger age at diagnosis did not predict a change in the diagnosis. (Table 16).

Table 16. Logistic regression analyses to predict the change in the diagnosis of polymyalgia rheumatica (PMR) during follow-up.

	Univariate analyses OR (95% CI)	Multivariable analysis OR (95% CI)
Female	0.94 (0.61–1.45)	1.06 (0.64–1.75)
Age at diagnosis < 65 years	1.42 (0.89–2.27)	1.38 (0.80–2.37)
Normal CRP and ESR values at diagnosis	3.32 (1.64–6.92)*	2.27 (0.98–5.24)
No morning stiffness	2.34 (1.51–3.69)*	1.75 (1.03–2.94)*
Absence of typical symptoms of PMR†	12.12 (4.90–36.60)*	7.45 (2.65–24.53)*
Less than full symptom response to GC treatment	3.27 (2.05–5.29)*	3.01 (1.83–5.00)*

† Patient had none of the typical symptoms: pain and tenderness in the neck, shoulder, upper arm, hip and thigh areas. OR, odds ratio; CI, confidence interval; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; GC, glucocorticoid. *Statistically significant finding. Modified from Paltta et al, 2023 (III).

6 Discussion

In registry studies, it is crucial to assess the quality and potential weaknesses of the register data used to avoid sources of error. If a study relies on a register with a poor diagnostic validity, the results of the study become inaccurate and unreliable. It is also important to be familiar with the data's content to assess this accurately. A clinician who is familiar with the daily routines of the clinic may have a different perspective on how the data is produced and thus on data quality than a statistician might. Although a single diagnosis may have poor validity, certain methods can improve the quality of the study and the validity of its results.

6.1 Accuracy of diagnosis in a whole hospital or in rheumatology department – The impact of the diagnosing speciality

For the diseases studied in this dissertation, the proportion of correct registry diagnoses was higher if only the diagnoses made in the department of rheumatology were considered compared to the diagnoses made in all departments of the hospital.

Our study covered diagnoses recorded in any speciality for all the diseases in this dissertation. The data on the speciality of the department where the diagnoses occurred was collected in all studies. Thus, the results could be calculated not only for the department of rheumatology alone but also for all specialties, thus improving the generalizability of our results.

6.1.1 The impact of the diagnosing speciality on the validity of RA diagnoses

Previous studies on the validity of RA diagnoses have typically validated only the diagnoses made at rheumatology clinics. The diagnoses of RA set for rheumatology clinic patients on a minimum of two visits have been studied in Sweden by Waldenlind et al. 2014, and in Denmark in DANBIO and DNPR registers by Poulsen et al. 2017. A definite diagnosis of RA was found in approximately 90% of the study patients in Sweden, 96% in the DANBIO register, and slightly lower at 79% in the DNPR register. In Minneapolis, Minnesota, the validity of RA diagnosis made in a

rheumatology department was 55% in the Veterans Administration database (Singh et al. 2004).

In our study, a definitive diagnosis of RA made in any speciality was found in 82% of patients with RA, a diagnosis of seronegative RA on 71% of patients with recorded seronegative RA, and a diagnosis of seropositive RA on 75% of patients with seropositive RA. For patients who were diagnosed in a department of rheumatology, these same results were higher, with the accuracy of 89%, 80% and 85%, respectively.

6.1.2 The impact of the diagnosing speciality on the validity of SSc diagnoses

To our knowledge, no comparison has been made in recent years of the accuracy of ICD-10 diagnoses of SSc made in rheumatology department or in any speciality. SSc diagnoses made in any speciality have been validated with the ICD-9 code 710.1 in the United States (Valenzuela et al. 2015), and with the ICD-10 code in France (Chaves et al. 2020), and the PPVs were 76% and 93%, respectively. Chaves et al. also studied the PPV for lcSSc which was 95% (Chaves et al. 2020). Early SSc-diagnoses were not considered true SSc in these studies.

Our study revealed that when the diagnosis of SSc was made in any speciality, the PPV was 66% for a general SSc diagnosis and 80% for lcSSc, the most common subgroup of SSc in Finland. These results are slightly lower than the results of Valenzuela et al. 2015 and Chaves et al. 2020. With the inclusion of early SSc as a correct diagnosis of SSc in our study, the PPVs were 75% for SSc and 95% for lcSSc. Most of the diagnoses in our study had been made in the rheumatology department; for these diagnoses, the PPVs were higher than for those made in any speciality, 78% for SSc and 81% for lcSSc, and including early SSc, 90% and 97%, respectively.

6.1.3 The impact of the diagnosing speciality on the validity of PMR diagnoses

In earlier studies, PMR diagnoses set in primary health care have been investigated in Sweden, with the validity of 60% (Fors et al. 2019). Out of the PMR diagnoses made by rheumatology departments, the correct diagnoses have previously accounted for 79.1% in Alaska (Bernatsky et al. 2011) and from 69.0% to 94% in the United Kingdom (Pease et al. 2005; Hutchings et al. 2007). In two studies in Italy, the persistence of PMR diagnoses was 72% in rheumatology and internal medicine clinics (Caporali et al. 2001), and 48% in a rheumatology clinic (Falsetti et al. 2011). In our study, the proportion of the PMR diagnoses persisting even after the follow-up period was 66% for diagnoses made in any speciality and 72% for

diagnoses made in rheumatology department which corresponds to the results of the previous studies.

6.1.4 The impact of different levels of care for rheumatological diseases on patient cohorts

It is natural that the accuracy of the diagnoses of rheumatological conditions is better when the diagnosis is made by a rheumatologist. However, at least for certain rheumatological diseases like PMR, milder cases rarely come to a rheumatological consultation and therefore are not found in registries of rheumatology department diagnoses. In Finland, a majority of the SSc patients visit a rheumatology department at least during some part of the course of their disease, and a majority of the RA patients are diagnosed and treated in rheumatology or internal medicine departments, but patients with PMR only come in for a consultation if there is uncertainty in the diagnosis or problems with treatment. For registry studies, the levels of care for different diseases is something to be aware of. When planning for a study, it should be considered whether it is more important to analyse just the patients with a certain diagnosis or to find all the patients with the disease as comprehensively as possible.

6.2 Improving results by combining data from different registries

Algorithms have been developed to better identify patients with specific rheumatological conditions in registries. Most of these have investigated RA (Carrara et al. 2015; Thomas et al. 2008; Widdifield et al. 2013; Liao et al. 2010), and they have typically included the number of diagnoses and the locations where they are made (in a department of rheumatology or in any department), prescribed anti-rheumatic medications, and diagnoses of other rheumatic conditions. Our study supports the idea that combining data from various registers can significantly enhance the validity of registry diagnoses.

6.2.1 Number of visits with the diagnosis

Our study found a correlation between the validity of the diagnosis of RA and the number of visits with this diagnosis. The results showed that the PPV was 82% for an RA diagnosis recorded on at least one hospital visit. The validity of the diagnoses was higher if the patient had the same diagnosis recorded on more than one visit, with a PPV of 85% for at least two, 89% for at least five and 90% for at least ten visits. A diagnosis based on multiple visits reduces the incidence of incorrectly input diagnoses and refines the diagnosis with data collected at follow-up visits.

Somewhat surprisingly, in contrast to RA, the increase in the number of visits with a diagnosis of PMR had much less effect on the accuracy of PMR diagnoses. The accuracy for at least one visit was 65.5%, for at least two visits 67.2%, for at least five visits 68.4%, and for at least 10 visits 66.1%. This might be due to the fact that at least in Finland, PMR is usually treated in primary health care, and only the more complex cases are sent for a rheumatology consultation. This approach means that patients may be followed up more frequently if there is a lack of response to treatment or if the accuracy of the diagnosis is in doubt. These kinds of cases may require extra visits due to differential diagnostics investigations. Moreover, a lengthy follow-up to catch late-changing initial diagnoses of PMR is beneficial.

In the study on the validity of SSc diagnoses, we did not aim to investigate the impact of the number of visits on SSc diagnoses, and thus did not collect data on the number of visits the diagnosis was made on.

6.2.2 Entitlements to reimbursement of medicine expenses

With RA, there was also a correlation between the validity of the diagnosis and the entitlement granted for special reimbursement for medication costs. The results of our study showed that the PPV was 82% for an RA diagnosis. The validity of the diagnoses was better with a PPV of 84% if the patient was entitled to special reimbursement for the cost of DMARDs, and if the entitlement was granted specifically for the inclusion diagnosis (RA, seropositive RA or seronegative RA), the PPV was 89%.

In the studies on SSc and PMR, we did not collect data on entitlements to special reimbursement for the cost of DMARDs or on medications other than glucocorticoids used for PMR. In PMR, entitlements for reimbursement are rarely applied because glucocorticoids are the most common treatment, they are inexpensive, and the most common, low-dose glucocorticoid prednisolone, is no longer reimbursable in Finland.

6.2.3 Laboratory data

Laboratory data may be available in some registry studies, as was the case in the studies of this dissertation, and we evaluated the impact of different laboratory values on the accuracy of the diagnoses.

6.2.3.1 Serological data

For RA, it was analysed whether serological data, specifically ACPA and RF, had an impact on the validity of the diagnosis. Previously in Denmark, Tenstad et al.

have investigated the use of ACPA or RF alone as a diagnostic test for RA and have shown the PPV associated with high positive ACPA to be higher with 43% than the PPV of 14% associated with high positive RF, when the high positivity was defined as values three times that of ULN (Tenstad et al. 2020). The results of this dissertation confirm these previous findings. The PPVs for seropositive RA diagnosis were a bit higher for both positive (96%) and high positive (98%) ACPA than for positive (92%) and high positive (94%) RF, respectively. The inclusion of ACPA status data in the study analyses clearly increased the validity of the diagnosis; the PPV of RA diagnosis in patients with recorded ACPA-positivity was excellent, 98%, when the PPV of RA diagnosis excluding serological data was 82%. It is important to note that ACPA and RF are an important component of the ACR/EULAR 2010 classification criteria for RA, and meeting the classification criteria had a major impact on whether the disease was deemed to be RA.

The impact of serological findings in patients diagnosed with SSc on the validity of SSc diagnosis was assessed using SSc-specific autoantibodies. Because nearly all the study patients were ANA positive, it was not relevant to separately analyse the effect of ANA positivity. The PPV for SSc diagnosis without data on specific autoantibodies was 78% and 80% for patients with lcSSc. For patients with specific autoantibodies of SSc, PPV of a diagnosis of SSc was 83,% and 81% for lcSSc. This means, that the effect of specific autoantibodies of SSc on the validity of SSc diagnoses was less pronounced than with RA. The reason for this is not known to us, but it can be speculated that seropositivity is perceived as more significant in diagnosing SSc than in diagnosing RA, meaning that it has been easier to diagnose seronegative RA than SSc if the patient lacks antibodies of the disease in question.

6.2.3.2 Inflammatory values

We analysed the effect of inflammatory laboratory values at diagnosis, CRP and ESR, on the persistence of PMR diagnosis. The results showed that these values were elevated on 94.3% of patients with persistent PMR diagnosis compared to 83.2% of patients with nonpersistent PMR diagnosis. This difference was statistically significant ($p<0.001$). During follow-up the diagnosis of PMR persisted in 69% of patients with elevated inflammatory values at diagnosis, compared to 40% with normal inflammatory values. In univariate logistic regression analysis, normal CRP and ESR values at diagnosis were a statistically significant predictor of a change in the diagnosis of PMR (OR 3.32, 95% CI 1.64–6.92). In the multivariate regression analysis normal CRP and ESR values at diagnosis did not quite reach statistical significance in predicting a change in PMR diagnosis (OR 2.27, 95% CI 0.98–5.24).

6.3 Challenges in diagnostics – Incorrectly input diagnoses and misdiagnoses

The meaning of validity of diagnoses can be considered from both the patient's perspective and an external perspective. For example, from the perspective of prevalence analysis, any false diagnosis is as bad as any other because they all distort the data. But from the patient's point of view, only the false diagnoses that worsen patient care either by delaying treatment or by causing the wrong treatment are detrimental. Based on this difference, invalid diagnoses can be divided into incorrectly input diagnoses and actual misdiagnoses.

6.3.1 Incorrectly input diagnoses

Incorrectly recorded diagnoses are often excluded from validation studies by some form of prescreening of patients. In a Swedish study of patients in the rheumatology department (Andréasson et al. 2014), 5.3% of all M34 diagnoses were incorrectly input, which is slightly lower than the 7.3% found in our study of SSc diagnoses.

In the studies of this dissertation, incorrectly input diagnoses accounted for 1.8% (8/455) of all diagnoses of PMR, 7.3% (28/385) of all diagnoses of SSc and 13.3% (31/233) of all diagnoses of RA. 73.8% (31/42) of the incorrect diagnoses of RA and 100% (29/29) of the incorrect diagnoses of seropositive RA were incorrectly input diagnoses. This means that incorrectly input diagnoses were a major problem compared to a change of diagnosis during follow-up or a clear misdiagnosis. Rheumatology departments were not immune to this problem; for example, 39.2% (20/51) of all incorrectly input RA diagnoses were recorded in rheumatology department. This number can be explained by the large number of visits in a department of rheumatology.

Incorrectly input diagnoses may be due to simple typing errors, or the actual recording of diagnoses may be done by administrative staff due to the requirements of administrative data systems. It can also be speculated that at least some of these errors are made when a patient is visiting a hospital with a completely unrelated condition without a good diagnosis, and the easiest diagnosis to input for a nonspecialist is “rheumatism”. This explanation is supported by the fact that a rare diagnosis such as SSc is seldom incorrectly input.

6.3.2 Misdiagnoses

According to our study, of the registry diagnoses of RA, 18% (42/233) of the final diagnoses were for conditions other than RA, and only 14% (6/42) of these incorrect diagnoses were considered misdiagnoses, for a total of 3% (6/233) of all diagnoses of RA. These six misdiagnoses were for patients that had a diagnosis of seronegative

RA in a registry. This means that 18% (6/33) of the incorrect diagnoses of seronegative RA were misdiagnoses. None of the recorded diagnoses of seropositive RA were misdiagnoses.

For the registry diagnoses of SSc, 34% (132/385) were considered not to be SSc. 79% (104/132) of these incorrect diagnoses and 27% (104/385) of all diagnoses of SSc were considered misdiagnoses. For lcSSc, 95% (44/46) of the incorrect diagnoses and 19% (44/226) of all lcSSc diagnoses were misdiagnoses.

6.3.2.1 How to determine limits of misdiagnoses?

It is sometimes difficult to determine the limits of misdiagnosis when we do not have actual diagnostic criteria, but instead only classification criteria for studies. For example, if a disease is at such an early stage of progression that it is only one point short of meeting the classification criteria, can the diagnosis be called a misdiagnosis? On the other hand, how mild does the disease or the risk of developing the disease have to be before it is uncertain whether it is a disease at all that it becomes a misdiagnosis?

With SSc as an example, a number of classification criteria have been developed (Van Den Hoogen et al. 2013; LeRoy et al. 2001; Masi 1980) and these criteria are often used in clinical practice as if they were diagnostic criteria, even though they are not intended to be used that way. The effectiveness of the 2013 revision of the ACR/EULAR classification criteria for SSc (Van Den Hoogen et al. 2013) has been reviewed, and SSc patients usually meet these criteria (Jordan et al. 2015; Araújo et al. 2017). But it is important to note that failure to meet the classification criteria does not necessarily exclude a diagnosis of SSc, especially when the disease is in early stages.

In our study, rates of misdiagnoses were also calculated with early SSc considered as true SSc. Using this method, 25% (95/385) of diagnoses of SSc were other than SSc, and 71% (67/95) of the incorrect diagnoses were misdiagnoses, making a total of misdiagnoses 17% (67/385) of all SSc diagnoses. For lcSSc, 82% (9/11) of incorrect diagnoses and 4% (9/226) of all lcSSc diagnoses were misdiagnoses. These percentages are clearly smaller than when these early diseases are considered misdiagnoses.

In diseases such as SSc, early diagnosis is important to prevent organ damage, and advances in diagnostics give the possibility of earlier disease detection. For example, the increasing availability of NVC (Araújo et al. 2017) can improve identification and follow-up of patients with early SSc. On the other hand, overdiagnosis of SSc due to nonadherence to the classification criteria can lead to a lack of homogeneity of study populations and make it difficult to compare data from different countries.

6.3.2.2 Using follow-up to establish the diagnosis

Ultimately, follow-up will show whether the early-stage disease develops into a full-blown disease. It may also happen that the disease picture changes, either because new symptoms and status findings emerge, or for some other reason a new investigation is carried out, such as imaging or laboratory tests. These developments can lead to changing the diagnosis even if the original diagnosis seemed valid at the time.

Seronegative RA is one of the complex disease entities where another competing diagnosis can often be made during clinical follow-up. Paalanen et al. studied register data of patients with initially diagnosed seronegative RA, and during 15 years of follow-up, spondyloartrhropathy was diagnosed in 8.8% (Paalanen et al. 2021) of the patients. In another study by Paalanen et al., early arthritis patients diagnosed with seronegative RA had an in-depth clinical follow-up period of 10 years, and a more specific or competing diagnosis could be suggested for the majority of them (Paalanen et al. 2019). Only 2% developed erosions typical for RA. According to our study, of the diagnoses of seronegative RA that were deemed incorrect, 15% changed to some other more defined diagnosis based on data acquired during follow-up.

6.3.2.3 All misdiagnoses are not equally harmful

Different levels of misdiagnoses exist, and some are more harmful than others. For example, misdiagnosing psoriatic arthritis as another rheumatic disease is less harmful than misdiagnosing an infection or malignancy as a rheumatic disease. In our study of differential diagnostics of PMR, we examined the diagnoses that had been primarily misdiagnosed as PMR. Making the diagnosis of PMR may be difficult because there is no benchmark for the verification of the diagnosis, the diagnosis is essentially clinical, and many symptoms and findings of PMR also occur in other medical disorders. (González-Gay et al. 2017; Buttgereit et al. 2016; Lundberg et al. 2022; Nothnagl et al. 2006)

In our study, when the diagnosis of PMR was considered to be wrong, the conditions that best explained the patient's situation were predominantly the same as those reported earlier (González-Gay et al. 2000; Michet et al. 2008). The most common mimics of PMR were other inflammatory conditions of the joints, accounting for 35% of the revised diagnoses, followed by degenerative and stress-related musculoskeletal disorders at 13%. In our study, 9.3% of the patients with a subsequent change of diagnosis had a malignancy, and their delay in diagnosis was a median of two months. It is important to minimise a delay in diagnosis, especially for malignancies, because a delay can worsen the prognosis for the patient.

Those patients who had an atypical manifestation of PMR were more likely to be misdiagnosed, according to the results of our study. In a multivariable logistic regression analysis, patients who did not have characteristic pain in the shoulders and pelvic girdle, stiffness in the morning, or a full response to glucocorticoids were particularly at risk of a wrong diagnosis. The presence of atypical characteristics of PMR is a red flag that should prompt a search for other diseases that mimic PMR, such as malignant diseases and infections (González-Gay et al. 2000). The use of imaging, for example with PET-CT, can detect a possible underlying vasculitis of large blood vessels, particularly in patients who have marked pelvic girdle involvement, inflammatory pain in the lower back or diffuse pain in both lower limbs (Prieto-Peña et al. 2019).

The effectiveness of various PMR classification criteria in the identification of patients at a higher risk of a wrong diagnosis was also investigated in our study. When the patients met the PMR classification criteria of ACR/EULAR from 2012, diagnosis was deemed accurate in 81.3% of cases and only 45.5% when they did not meet these criteria. For a correct diagnosis of PMR in this study, patients had to fulfil two requirements: meet at least one set of classification criteria and have no other diagnosis that better explains the patient's condition during follow-up. Meeting the classification criteria suggests a more typical manifestation of PMR, and the PMR diagnosis changed less frequently in these patients than in those with a less typical manifestation. The classification criteria were frequently met even in those patients who were subsequently reclassified as having a disease other than PMR. This underlines the point that classification criteria are not the same as diagnostic criteria (Aggarwal et al. 2015), and they should only be used in patients who have already had alternative diagnoses ruled out with reasonable certainty.

6.4 Generalizability of the results of this dissertation

The results of this dissertation are based on large, centralised hospitals in Finland. The RA study (Study I) was conducted in four university hospitals and one large central hospital, the SSc study (Study II) in two university hospitals and the PMR study (Study III) in one university hospital. The findings are therefore most applicable to patients who receive diagnosis and treatment in large, centralised hospitals. The proportion of patients treated in hospitals at this level in Finland varies among the diseases studied in the thesis. For SSc, most patients are evaluated in central or university hospitals due to the rarity and severity of the disease, which enhances the overall validity of the study findings. In the case of PMR, most patients receive their diagnosis and treatment in primary care. Only those patients whose diagnosis is uncertain or whose response to treatment is inadequate are referred to a

rheumatologist. The patient population in a university hospital's department of rheumatology is therefore more complex than the patients in primary care. For instance, this could explain why a considerable number of patients in our study did not have elevated inflammatory values. RA falls between these diseases, as the majority of diagnoses are made by rheumatologists, but not quite as comprehensively as for SSc.

6.5 Limitations and strengths of the studies

A limiting factor in the studies of this dissertation is the retrospective study design. The data collection relied on medical records, which may not have contained all the necessary information for every patient. It is possible that some patients did not meet the classification criteria due to missing data. When we assessed the fulfilment of the classification criteria, any item that was not explicitly marked as positive in the medical records was considered negative. However, the PMR study required documentation on the absence of RF and/or ACPA through a negative result on the test in question. In some cases, the accuracy of the diagnosis could not be determined due to a considerable amount of missing information. The number of these patients was not very high, but 12 out of 250 RA patients, 27 out of 412 SSc patients and 73 out of 455 PMR patients had to be excluded from the final analysis. It was much more common for patients to have multiple visits and for medical records to be extensive and comprehensive, which in turn was a strength of the studies.

We did not have access to the medical records of health care facilities outside the hospitals participating in the study. This limitation affected Study I, as some patients initially defined as controls had to be excluded from the analysis altogether. This was due to the discovery that they had visited another Finnish hospital with a diagnosis of RA.

The complexity of diagnosing RA, SSc and PMR can also be seen as a limiting factor. In the absence of clear diagnostic criteria for these diseases, it was ultimately the decision of the reviewer whether the diagnosis recorded in the register was correct or incorrect, and these results might have been strengthened by a collaborative decision by multiple investigators. In the studies on SSc and PMR, this decision was based on the fulfilment of classification criteria as a prerequisite for a correct diagnosis. In the study on RA, meeting the classification criteria was not mandatory to confirm the diagnosis. However, in the study, of the patients with available data (about 55 percent of all) 88% of patients with seronegative RA and 97% of patients with seropositive RA met either or both of the ACR 1987 and ACR/EULAR 2010 classification criteria (Arnett et al. 1988; Aletaha et al. 2010). The multicentre design also strengthened the external validity of the RA study results.

In Study I, we analysed all patients with either seropositive or seronegative RA as a single group. However, we also analysed these patients separately, with one group diagnosed with seropositive RA and the other with seronegative RA. This approach allows for easier comparison of our study's results with those of previous studies (Waldenlind et al. 2014; Poulsen et al. 2017; Singh et al. 2004) that did not differentiate between serological subgroups of RA.

A clear strength of this dissertation are the long follow-up periods of the studies. This is particularly evident in Study III, which focused on the differential diagnosis of PMR. The follow-up period, with a median of 34 months, was significantly longer compared to several previous studies on the subject (Caporali et al. 2011; Falsetti et al. 2011; Hutchings et al. 2007), and of the changes in diagnosis, 13.7% occurred after the one-year follow-up limit of these shorter studies.

No patients were prescreened in any of the studies of this dissertation, and the recorded diagnosis of RA, SSc or PMR was the only criteria for inclusion. This method makes our results more generalizable than in many previous studies with prescreened patients.

In all studies of this dissertation, information was collected on the speciality in which each diagnosis was made, allowing for results to be calculated for both rheumatology department alone and also for all specialities. This approach also improves the generalizability of the study findings.

7 Summary/Conclusions

This dissertation investigated the validity of CRHC diagnoses of RA in biobank patients and diagnoses of SSc and PMR in hospital discharge registers, as well as how best to improve this validity by combining information from different sources. We also looked at how often a PMR diagnosis was changed to another disease, what these diseases misdiagnosed as PMR were, and which factors at the time of PMR diagnosis indicate a high risk of misdiagnosis.

The validity of the CRHC diagnosis of RA in the patients of the Finnish biobanks was moderate and could be considerably strengthened by the integration of information from the Drug Reimbursement Register. The validity was especially high in patients who were entitled to special reimbursement for the costs of medicine, had multiple visits with a diagnosis of RA, who were seropositive for autoantibodies of RA, and whose diagnosis was made in rheumatology department. Especially for seronegative RA, the validity of a diagnosis from one visit was only moderate, which is consistent with seronegative RA being a nonuniform disease entity.

In the two Finnish university hospitals studied, the hospital registry diagnoses of SSc were accurate only if diagnosed in the rheumatology department. In the case of lcSSc, the most common SSc type in Finland, the validity was high even if the diagnosis was registered in any department. These findings strengthen the reliability of registry studies based on SSc diagnoses, especially those of lcSSc.

Among patients diagnosed with PMR at a university hospital, a third of the original hospital registry diagnoses were found to be incorrect following a more comprehensive assessment and follow-up, with only a slightly better retention of diagnoses made at a department of rheumatology. This study emphasised the importance of considering differential diagnoses when making a PMR diagnosis and explored factors predictive of a change in diagnosis during follow-up. The risk of misdiagnosis was particularly high in patients with atypical disease patterns, inadequate response to medication, and in those who did not meet classification criteria. The most frequent mimics of PMR included other inflammatory arthritides, noninflammatory musculoskeletal conditions, infective diseases and malignancy.

It is essential to know the limitations of health care registers and the means to manage these limitations to best utilise the registry data and ensure the quality of future register-based studies.

Acknowledgements

This study was carried out at the Centre for Rheumatology and Clinical Immunology, Turku University Hospital. The study was financially supported by the FinnGen research project, the Finnish Foundation for Rheumatic Diseases, the Turunmaa Duodecim Society, the Finnish Governmental VTR-funding, and the Doctoral Programme in Clinical Research by University of Turku Graduate School.

First and foremost, my deepest gratitude goes to my supervisors, Docent Laura Pirilä and Docent Antti Palomäki, without whom this dissertation would not have been possible. Laura, you have acted as an inspiring mentor for me in several ways. Your enthusiasm for research was instrumental in persuading me to finally take up a PhD project that I had pondered for ages. I am grateful for the numerous ideas you have provided and for enabling me to complete this dissertation by allocating me time for research alongside clinical work. This allowed me to avoid using too many evenings and nights for this task. Antti, you have patiently guided me throughout the research and writing process. You didn't let me off the hook too easily but helped me to refine things until they were just right. I really appreciate how you managed to find an optimal balance between constructive criticism, emotional support and light-hearted banter, and I am grateful for your guidance.

I am extremely grateful to Professor h.c. Kari Puolakka and Docent Paula Vähäsalo for reviewing this dissertation and providing me with valuable comments and suggestions that have significantly enhanced the quality of this dissertation. I am grateful to Professor h.c. Dan Nordström for accepting the invitation to act as my opponent in the dissertation defense. I would also like to thank Professor Markus Juonala and Docent Johanna Huhtakangas for their contribution in the doctoral follow-up group of this dissertation.

I would like to express my gratitude to the co-authors of my articles, Kari Eklund, Aki Havulinna, Hanna-Kaisa Heikkilä, Johanna Huhtakangas, Pia Isomäki, Oili Kaipiainen-Seppänen, Saara Kortelainen, Kati Kristiansson, Markus Käyrä, Tuulikki Sokka-Isler and Saara Suuronen for all your important contributions in this study. I would also like to thank Eliisa Löyttyniemi for providing valuable advice on a tricky statistical problem.

My sincere thanks to my colleagues Saara Kortelainen, Ilpo Koskivirta, Taina Kotijärvi, Petri Kresanov, Laura Kuusalo, Reetta Laiho, Krista Lundelin, Markku Mali, Noora Mattsson, Maija Puurtinen-Vilkki, Kira Rintala, Laura Ryyppö, Milja Söderström and Kirsi Taimen as well as the entire staff of the Centre for Rheumatology and Clinical Immunology for those important moments around the coffee table. You have created a safe, fun and inspiring atmosphere where it was possible to exchange ideas and questions about all things important and not so important. It has always been a pleasure to come to work and it has meant a lot to me.

I wish to thank Louise Aaltonen, Markus Hakamäki, Tapio Hellman, Jonna Virtanen and Lassi Nelimarkka for the more or less scientific discussions and debates in our office at the Department of Internal Medicine. From you I received important peer support and a lot of practical advice for my dissertation.

I would like to acknowledge late Professor Matti Karp. My brief work in his research group during my student years led to the first articles I was involved in and Matti first demonstrated to me the practicalities of doing science. I thank Professor Timo Möttönen for awakening my interest in rheumatology with his charismatic presence and for inviting me into the world of rheumatology.

Special thanks to my parents, Raija and Kauko, for instilling in me an interest in nature, science and human being that has served as a solid foundation for my life and career. Thank you for always being there and making it possible for me to follow my dreams. I am also grateful for the privilege to grow up with my siblings, Anni and Juho, and subsequently with their spouses and children. I consider myself fortunate to have gained many friends who have been there for me over the years, and I thank all of you for that.

Finally, I would like to express my profound gratitude to my family. Thank you, Ville, for your support and companionship, I am immensely privileged to have your presence in my life. Veera and Roosa, my wonderful daughters. You are both still a little too young to understand this text, especially as it is in English. But if you read it later, I want you to know how important you are to me, now and always. Thanks to you, I was able to leave the work well behind when I came home, and you made me focus on more important things in life. Thank you for being you.

Turku, September 2024
Johanna Paltta

References

Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions between diagnostic and classification criteria? *Arthritis Care Res*. 2015;67(7):891–7.

Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, et al. 2010 Rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum*. 2010;62(9):2569–81.

Aletaha D, Smolen JS. Diagnosis and management of rheumatoid arthritis: A review. *JAMA*. 2018;320(13):1360–72.

Anderson BR, Gotlieb EG, Hill K, McHugh KE, Scheurer MA, Mery CM, et al. Registry-based trials: A potential model for cost savings? *Cardiol Young*. 2020;30(6):807–17.

Andréasson 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.

Araújo FC, Camargo CZ, Kayser C. Validation of the ACR/EULAR classification criteria for systemic sclerosis in patients with early scleroderma. *Rheumatol Int*. 2017;37(11):1825–33.

Arnett FC, Edworthy SM, Bloch DA, Mcshane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum*. 1988;31(3)315–24.

Aronowitz RA. When do symptoms become a disease? *Ann Intern Med*. 2001;134:803–8.

Baker JF, Conaghan PG, Gandjbakhch F. Update on magnetic resonance imaging and ultrasound in rheumatoid arthritis. *Clin Exp Rheumatol*. 2018;36 Suppl 114 (5):16–23.

Balogh EP, Miller BT, Ball JR, editors. *Improving diagnosis in health care*. Washington (DC): National Academies Press (US); 2015.

Benchimol EI, Manuel DG, To T, Griffiths AM, Rabeneck L, Guttmann A. Development and use of reporting guidelines for assessing the quality of validation studies of health administrative data. *J Clin Epidemiol*. 2011;64(8):821–9.

Bernatsky S, Linehan T, Hanly JG. The accuracy of administrative data diagnoses of systemic autoimmune rheumatic diseases. *J Rheumatol*. 2011;38(8):1612–6.

Bili A, Sartorius JA, Kirchner HL, Morris SJ, Ledwich LJ, Antohe JL, et al. Hydroxychloroquine use and decreased risk of diabetes in rheumatoid arthritis patients. *J Clin Rheumatol*. 2011;17(3):115–20.

Bird HA, Esselinckx W, J Dixon AS, Mowat AG, N Wood PH. An evaluation of criteria for polymyalgia rheumatica. *Ann Rheum Dis*. 1979;38(5):434–9.

Buttgereit F, Dejaco C, Matteson EL, Dasgupta B. Polymyalgia rheumatica and giant cell arteritis a systematic review. *JAMA*. 2016;315(22):2442–58.

Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203–9.

Böcking W, Ahrens U, Kirch W, Milakovic M. First results of the introduction of DRGs in Germany and overview of experience from other DRG countries. *J Public Health*. 2005;13:128–37.

Calderon LM, Pope JE. Scleroderma epidemiology update. *Curr Opin Rheumatol*. 2021;33(2):122–7.

Callhoff J, Albrecht K, Marschall U, Strangfeld A, Hoffmann F. Identification of rheumatoid arthritis in German claims data using different algorithms: Validation by cross-sectional patient-reported survey data. *Pharmacoepidemiol Drug Saf*. 2023;32(5):517–25.

Caporali R, Montecucco C, Epis O, Bobbio-Pallavicini F, Maio T, Cimmino MA. Presenting features of polymyalgia rheumatica (PMR) and rheumatoid arthritis with PMR-like onset: A prospective study. *Ann Rheum Dis*. 2001;60(11):1021–4.

Carrara G, Scirè CA, Zambon A, Cimmino MA, Cerra C, Caprioli M, et al. A validation study of a new classification algorithm to identify rheumatoid arthritis using administrative health databases: case-control and cohort diagnostic accuracy studies. Results from the RECord linkage On Rheumatic Diseases study of the Italian Society for Rheumatology. *BMJ Open*. 2015;5(1):e006029.

Carroll RJ, Thompson WK, Eyler AE, Mandelin AM, Cai T, Zink RM, et al. Portability of an algorithm to identify rheumatoid arthritis in electronic health records. *J Am Med Inform Assoc*. 2012;19(E1):e162–9.

Chaves SDA, Derumeaux H, Minh P Do, Lapeyre-Mestre M, Moulis G, Pugnet G. Assessment of the accuracy of using ICD-10 codes to identify systemic sclerosis. *Clin Epidemiol*. 2020;12:1355–9.

Chuang TY, Hunder GG, Ilstrup DM, Kurland LT. Polymyalgia rheumatica: A 10-year epidemiological and clinical study. *Ann Intern Med*. 1982;97(5):672–80.

Chung CP, Rohan P, Krishnaswami S, McPheeters ML. A systematic review of validated methods for identifying patients with rheumatoid arthritis using administrative or claims data. *Vaccine*. 2013;31(S10):K41–61.

Conrad P, Barker KK. The social construction of illness: Key insights and policy implications. *J Health Soc Behav*. 2010;51(1 suppl):S67–79.

Dasgupta B, Cimmino MA, Kremers HM, Schmidt WA, Schirmer M, Salvarani C, et al. 2012 Provisional classification criteria for polymyalgia rheumatica: A European League Against Rheumatism/American College of Rheumatology collaborative initiative. *Arthritis Rheum*. 2012;64(4):943–54.

de Angelis R, Giuggioli D, Bajocchi G, Dagna L, Zanframundo G, Foti R, et al. Sex-related differences in systemic sclerosis: A multicenter cross-sectional study from the national registry of the Italian Society for Rheumatology. *J Rheumatol*. 2022;49(2):176–85.

De Coster C, Quan H, Finlayson A, Gao M, Halfon P, Humphries KH, et al. Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: Report from an international consortium. *BMC Health Serv Res*. 2006;6:77.

Dejaco C, Singh YP, Perel P, Hutchings A, Camellino D, Mackie S, et al. 2015 Recommendations for the management of polymyalgia rheumatica: a European League Against Rheumatism/American College of Rheumatology collaborative initiative. *Arthritis Rheumatol*. 2015;67(10):2569–80.

Denton CP, Khanna D. Systemic sclerosis. *Lancet*. 2017;390(10103):1685–99.

Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet*. 2017;390(10092):415–23.

Dorland WAN. *Dorland's Illustrated Medical Dictionary*. 32nd ed. Philadelphia: Saunders, 2012.

Duodecim. Rheumatoid arthritis: Current care guidelines (In Finnish) [Internet]. 2022 [cited 2024 2 Aug]. Available from: <https://www.kaypahoito.fi/hoi21010#K1>.

Edwards CJ, Campbell J, van Staa T, Arden NK, Christopher D, Edwards J. Regional and temporal variation in the treatment of rheumatoid arthritis across the UK: a descriptive register-based cohort study. *BMJ Open*. 2012;2:e001603.

Espígol-Frigolé G, Dejaco C, Mackie SL, Salvarani C, Matteson EL, Cid MC. Polymyalgia rheumatica. *Lancet*. 2023;402(10411):1459–72.

EUR-Lex. General Data Protection Regulation [Internet]. 2016 [cited 2023 16 Jun]. Available from: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

Falsetti P, Acciai C, Volpe A, Lenzi L. Ultrasonography in early assessment of elderly patients with polymyalgic symptoms: A role in predicting diagnostic outcome? *Scand J Rheumatol.* 2011;40(1):57–63.

Felson DT, Anderson JJ. Methodological and statistical approaches to criteria development in rheumatic diseases. *Baillieres Clin Rheumatol*. 1995;9(2):253–266

Fimea. Biobanks [Internet]. 2023 [cited 2023 15 Jun]. Available from: <https://www.fimea.fi/web/en/supervision/biobanks>.

Finckh A, Gilbert B, Hodkinson B, Bae SC, Thomas R, Deane KD, et al. Global epidemiology of rheumatoid arthritis. *Nat Rev Rheumatol*. 2022;18(10)591–602.

Finlex. Act on electronic prescription (In Finnish) [Internet]. 2023 [cited 2023 15 Jun]. Available from: <https://www.finlex.fi/fi/laki/ajantasa/2007/20070061>.

Finlex. Act on the electronic processing of customer data in social welfare and health care (in Finnish) [Internet]. 2021 [cited 2023 15 Jun]. Available from: <https://www.finlex.fi/fi/laki/alkup/2021/20210784>.

Finlex. Act on the funding of welfare regions (in Finnish) [Internet]. 2023 [cited 2023 17 Nov]. Available from: <https://www.finlex.fi/fi/laki/ajantasa/2021/20210617>.

Finlex. Act on the National Institute for Health and Welfare (in Finnish) [Internet]. 2023 [cited 2023 15 Jun]. Available from: <https://www.finlex.fi/fi/laki/ajantasa/2008/20080668>.

Finlex. Act on the secondary use of social and health data (in Finnish) [Internet]. 2019 [cited 2023 15 Jun]. Available from: <https://www.finlex.fi/fi/laki/alkup/2019/20190552>.

Finlex. Biobank act [Internet]. 2012 [cited 2023 1 Jun]. Available from: <https://finlex.fi/en/laki/kaannokset/2012/en20120688>.

Finngen: an expedition into genomics and medicine [Internet]. 2023 [cited 2023 2 Jun]. Available from: <https://www.finngen.fi/en>.

Finnish Social and Health Data Permit Authority Findata [Internet]. 2023 [cited 2023 15 Jun]. Available from: <https://findata.fi/en/>.

Fors C, Bergström U, Willim M, Pilman E, Turesson C. Validity of polymyalgia rheumatica diagnoses and classification criteria in primary health care. *Rheumatol Adv Pract*. 2019;3(2):rkz033.

Funovits J, Aletaha D, Bykerk V, Combe B, Dougados M, Emery P, et al. The 2010 American College of Rheumatology/European league against rheumatism classification criteria for rheumatoid arthritis: Methodological report phase I. *Ann Rheum Dis*. 2010;69(9):1589–95.

Gissler M, Haukka J. Finnish health and social welfare registers in epidemiological research. *Nor Epidemiol*. 2004;14(1):113–20.

González-Gay MA, García-Porrúa C, Salvarani C, Olivieri I, Hunder GG. Polymyalgia manifestations in different conditions mimicking polymyalgia rheumatica. *Clin Exp Rheumatol*. 2000;18(6):755–9.

Gonzalez-Gay MA, Garcia-Porrúa C, Salvarani C, Olivieri I, Hunder GG. The spectrum of conditions mimicking polymyalgia rheumatica in Northwestern Spain. *J Rheumatol*. 2000;27(9):2179–84.

González-Gay MA, Matteson EL, Castañeda S. Polymyalgia rheumatica. *Lancet*. 2017;390(10103):1700–12.

Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med*. 2005;165:1493–9.

Graber ML. The incidence of diagnostic error in medicine. BMJ Qual Saf. 2013;22(Suppl 2):ii21–7.

Gravallese EM, Firestein GS. Rheumatoid arthritis — Common origins, divergent mechanisms. *N Engl J Med*. 2023;388(6):529–42.

Gunderson CG, Bilan VP, Holleck JL, Nickerson P, Cherry BM, Chui P, et al. Prevalence of harmful diagnostic errors in hospitalised adults: A systematic review and meta-analysis. *BMJ Qual Saf*. 2020;29(12):1008–18.

Guthridge JM, Wagner CA, James JA. The promise of precision medicine in rheumatology. *Nat Med*. 2022;28(7):1363–71.

Hafström I, Ajeganova S, Andersson MLE, Bala SV, Bergman S, Bremander A, et al. A Swedish register-based, long-term inception cohort study of patients with rheumatoid arthritis - Results of clinical relevance. *Open Access Rheumatol*. 2019;11:207–17.

Hakala M, Pöllänen R, Nieminen P. The ARA 1987 revised criteria select patients with clinical rheumatoid arthritis from a population based cohort of subjects with chronic rheumatic diseases registered for drug reimbursement. *J Rheumatol*. 1993;20(10):1674–8.

Hanly JG, Thompson K, Skedgel C. The use of administrative health care databases to identify patients with rheumatoid arthritis. *Open Access Rheumatol*. 2015;7:69–75.

Harrison JE, Weber S, Jakob R, Chute CG. ICD-11: an international classification of diseases for the twenty-first century. *BMC Med Inform Decis Mak*. 2021;21(Suppl 6):206.

Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208.

Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377–81.

Haverinen S, Vihervaara A, Löyttyniemi E, Peltonen S, Koulu L, Tasanen K, et al. Validation of psoriasis diagnoses recorded in Finnish biobanks. *Acta Derm Venereol*. 2020;100(17):adv00297.

Hazra A, Gogtay N. Biostatistics series module 7: The statistics of diagnostic tests. *Indian J Dermatol*. 2017;62(1):18–24.

Healey LA. Long-term follow-up of polymyalgia rheumatica: Evidence for synovitis. *Semin Arthritis Rheum*. 1984;13(4):322–8.

Helle N, Niinimäki M, Linnakaari R, But A, Gissler M, Heikinheimo O, et al. National register data are of value in studies on miscarriage—Validation of the healthcare register data in Finland. *Acta Obstet Gynecol Scand*. 2022;101(11):1245–52.

Hutchings A, Hollywood J, Lamping DL, Pease CT, Chakravarty K, Silverman B, et al. Clinical outcomes, quality of life, and diagnostic uncertainty in the first year of polymyalgia rheumatica. *Arthritis Rheum*. 2007;57(5):803–9.

Jaakkola E, Herzberg I, Laiho K, Barnardo MCNM, Pointon JJ, Kauppi M, et al. Finnish HLA studies confirm the increased risk conferred by HLA-B27 homozygosity in ankylosing spondylitis. *Ann Rheum Dis*. 2006;65(6):775–80.

Jessop ZM, Al-Sabah A, Francis WR, Whitaker IS. Transforming healthcare through regenerative medicine. *BMC Med*. 2016;14(1):115.

Johnson SR, Goek ON, Singh-Grewal D, Vlad SC, Feldman BM, Felson DT, et al. Classification criteria in rheumatic diseases: A review of methodologie properties. Vol. 57, *Arthritis Care and Research*. 2007. p. 1119–33.

Jones JG, Hazleman BL. Prognosis and management of polymyalgia rheumatica. *Ann Rheum Dis*. 1981;40(1):1–5.

Jordan S, Maurer B, Toniolo M, Michel B, Distler O. Performance of the new ACR/EULAR classification criteria for systemic sclerosis in clinical practice. *Rheumatology (Oxford)*. 2015;54(8):1454–8.

June RR, Aggarwal R. The use and abuse of diagnostic/classification criteria. *Best Pract Res Clin Rheumatol*. 2014;28(6):921–34.

Jutel A. Sociology of diagnosis: A preliminary review. Sociol Health Illn. 2009;31(2):278–99.

Kameda H, Kobayashi S, Tamura N, Kadono Y, Tada K, Yamamura M, et al. Non-radiographic axial spondyloarthritis. *Mod Rheumatol*. 2021;31(2):277–82.

Kanta Services [Internet]. 2023 [cited 2023 15 Jun]. Available from: https://www.kanta.fi/en/.

Katz JN, Barrett J, Liang MH, Bacon AM, Kaplan H, Kieval RI, et al. Sensitivity and positive predictive value of Medicare Part B physician claims for rheumatologic diagnoses and procedures. *Arthritis Rheum*. 1997;40(9):1594–600.

Kela. Kela's organisation [Internet]. 2023 [cited 2023 22 Nov]. Available from: <https://www.kela.fi/organizational-structure>.

Kela. Life situations [Internet]. 2023 [cited 2023 2 Nov]. Available from: <https://www.kela.fi/main-page>.

Kela. Reimbursements for medicine expenses [Internet]. 2023 [cited 2023 16 Jun]. Available from: <https://www.kela.fi/medicine-expenses>.

Kela. Research and statistics [Internet]. 2023 [cited 2023 9 Nov]. Available from: <https://www.kela.fi/kelas-research-and-statistics>.

Kela. Statistics on entitlements to reimbursement of medicines (in Finnish) [Internet]. 2023 [cited 2023 22 Nov]. Available from: <https://tietotarjotin.fi/tilasto/2855926/tilasto-laakkeiden-korvausoikeuksista>.

Kela. Statistics on purchased medicines reimbursed by health insurance (in Finnish) [Internet]. 2023 [cited 2023 22 Nov]. Available from: <https://tietotarjotin.fi/tilasto/2857522/tilasto-sairausvakuutuksesta-korvattavista-laaketoimituksista>.

Kim SY, Servi A, Polinski JM, Mogun H, Weinblatt ME, Katz JN, et al. Validation of rheumatoid arthritis diagnoses in health care utilization data. *Arthritis Res Ther*. 2011;13(1):R32.

Klein NP, Ray P, Carpenter D, Hansen J, Lewis E, Fireman B, et al. Rates of autoimmune diseases in Kaiser Permanente for use in vaccine adverse event safety studies. *Vaccine*. 2010;28(4):1062–8.

Kornelsen J, Atkins C, Brownell K, Woollard R. The meaning of patient experiences of medically unexplained physical symptoms. *Qual Health Res*. 2016;26(3):367–76.

Kortelainen S, Käyrä M, Hurme S, Paltta J, Pirilä L, Huhtakangas J. Systemic sclerosis: changes in the incidence rates in the Finnish population during the years 1999–2018. *Scand J Rheumatol*. 2024;53(1)29–35.

Koutsouris D. The evolution of medical care: from the beginnings to personalized medicine. *Health Technol*. 2017;7:3–4.

Kuhn GJ. Diagnostic errors. *Acad Emerg Med*. 2002;9(7):740–50.

Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613(7944):508–18.

König IR, Fuchs O, Hansen G, von Mutius E, Kopp MV. What is precision medicine? *Eur Respir J*. 2017;50(4):1700391.

LeRoy EC, Medsger Jr TA. Criteria for the classification of early systemic sclerosis. *J Rheumatol.* 2001;28(7):1573–6.

Li R, Chen Y, Ritchie MD, Moore JH. Electronic health records and polygenic risk scores for predicting disease risk. *Nat Rev Genet*. 2020;21(8):493–502.

Liao KP, Cai T, Gainer V, Goryachev S, Zeng-Treitler Q, Raychaudhuri S, et al. Electronic medical records for discovery research in rheumatoid arthritis. *Arthritis Care Res (Hoboken)*. 2010;62(8):1120–7.

Losina E, Barrett J, Baron JA, Katz JN. Accuracy of Medicare claims data for rheumatologic diagnoses in total hip replacement recipients. *J Clin Epidemiol*. 2003;56(6):515–9.

Ludvigsson JF, Håberg SE, Knudsen GP, Lafolie P, Zoega H, Sarkkola C, et al. Ethical aspects of registry-based research in the Nordic countries. *Clin Epidemiol*. 2015;7:491–508.

Lundberg IE, Sharma A, Turesson C, Mohammad AJ. An update on polymyalgia rheumatica. *J Intern Med*. 2022;292(5):717–32.

MacLean CH, Park GS, Traina SB, Liu HH, Hahn BH, Paulus HE, et al. Positive predictive value (PPV) of an administrative data-based algorithm for the identification of patient with rheumatoid arthritis (RA). *Arthritis Rheum*. 2001;44(S9):S106.

Maret-Ouda J, Tao W, Wahlin K, Lagergren J. Nordic registry-based cohort studies: Possibilities and pitfalls when combining Nordic registry data. *Scand J Publ Health*. 2017;45(17_suppl):14–9.

Masi AT. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum*. 1980;23(5):581–90.

Michet CJ, Matteson EL. Polymyalgia rheumatica. *BMJ*. 2008;336(7647):765–9.
Mishra A, Malik R, Hachiya T, Jürgenson T, Namba S, Posner DC, et al. Stroke genetics informs drug discovery and risk prediction across ancestries. *Nature*. 2022;611(7934):115-23.
Muilu P, Rantalaiho V, Kautiainen H, Virta LJ, Eriksson JG, Puolakka K. Increasing incidence and shifting profile of idiopathic inflammatory rheumatic diseases in adults during this millennium. *Clin Rheumatol.* 2019;38(2):555–62.
Ng B, Aslam F, Petersen NJ, Yu HJ, Suarez-Almazor ME. Identification of rheumatoid arthritis patients using an administrative database: A veterans affairs study. *Arthritis Care Res*. 2012;64(10):1490–6.
Nothnagl T, Leeb BF. Diagnosis, differential diagnosis and treatment of polymyalgia rheumatica. *Drugs Aging*. 2006;23(5):391–402.
Nuki G, Simkin PA. A concise history of gout and hyperuricemia and their treatment. *Arthritis Res Ther*. 2006;8(Suppl 1):S1. doi: 10.1186/ar1906.
Okkonen M, Havulinna AS, Ukkola O, Huikuri H, Ketonen M, Kesäniemi YA, et al. The validity of hospital discharge register data on non-ST-elevation and ST-elevation myocardial infarction in Finland. *Scand Cardiovasc J*. 2020;54(2):108–14.
Olliaro P, Torreele E. Managing the risks of making the wrong diagnosis: First, do no harm. *Int J Infect Dis*. 2021;(106):382–5.
Olsen J. Register-based research: Some methodological considerations. *Scand J Public Health*. 2011;39(3):225–9.
Paalanen K, Puolakka K, Nikiphorou E, Hannonen P, Sokka T. Is seronegative rheumatoid arthritis true rheumatoid arthritis? A nationwide cohort study. *Rheumatology (Oxford)*. 2021;60(5):2391–5.
Paalanen K, Rannio K, Rannio T, Asikainen J, Hannonen P, Sokka T. Does early seronegative arthritis develop into rheumatoid arthritis? A 10-year observational study. *Clin Exp Rheumatol*. 2019;37(1):37–43.
Paskal W, Paskal AM, Dębski T, Gryziak M, Jaworowski J. Aspects of modern biobank activity – Comprehensive review. *Pathol Oncol Res*. 2018;24(4):771–85.
Pease CT, Haugeberg G, Morgan AW, Montague B, Hensor EMA, Bhakta BB. Diagnosing late onset rheumatoid arthritis, polymyalgia rheumatica, and temporal arteritis in patients presenting with polymyalgic symptoms. A prospective longterm evaluation. *J Rheumatol*. 2005;32(6):1043–6.
Petersen I. Classification and treatment of diseases in the age of genome medicine based on pathway pathology. *Int J Mol Sci.* 2021;22(17):9418–42.
Pham A, Cummings M, Lindeman C, Drummond N, Williamson T. Recognizing misclassification bias in research and medical practice. *Fam Pract*. 2019;36(6):804–7.
Pietikäinen P, Turunen J. Diagnoses in and out of time: historical and medical perspectives on the diagnoses of distress. *Diagnosis*. 2017;4(1):3–11.
Pope JE, Denton CP, Johnson SR, Fernandez-Codina A, Hudson M, Nevskaya T. State-of-the-art evidence in the treatment of systemic sclerosis. *Nat Rev Rheumatol*. 2023;19(4):212–26.
Porter D, Basu N, Siebert S, Porter D. Classification criteria : time for a rethink? *Ann Rheum Dis*. 2020;79(10):1264–6.
Poulsen UE, Hetland ML, Pelck R, Nordin H, Jensen D V., Raun JL, et al. Validity and completeness of rheumatoid arthritis diagnoses in the nationwide DANBIO clinical register and the Danish National Patient Registry. *Clin Epidemiol.* 2017;9:627–32.
Prieto-Peña D, Martínez-Rodríguez I, Loricera J, Banzo I, Calderón-Goercke M, Calvo-Río V, et al. Predictors of positive 18 F-FDG PET/CT-scan for large vessel vasculitis in patients with persistent polymyalgia rheumatica. *Semin Arthritis Rheum*. 2019 Feb 1;48(4):720–7.
R. The R project for statistical computing [Internet]. 2024 [cited 2024 4 Jul]. Available from: <https://www.r-project.org>.
Renko M, Soini H, Rantala H, Tapiainen T, Pokka T, Uhari M. Systematic teaching of differential diagnostics and information retrieval in medical education – a controlled study. (in Finnish). *Duodecim*. 2010;126(5):549–56.

Rubinger L, Ekhtiari S, Gazendam A, Bhandari M. Registries: Big data, bigger problems? *Injury*. 2023;54:S39–42.

Rudwaleit M, Van Der Heijde D, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): Validation and final selection. *Ann Rheum Dis*. 2009;68(6):777–83.

Rutka K, Garkowski A, Karaszewska K, Łebkowska U. Imaging in diagnosis of systemic sclerosis. *J Clin Med*. 2021;10(2):248.

Rämö JT, Kiiskinen T, Seist R, Krebs K, Kanai M, Karjalainen J, et al. Genome-wide screen of otosclerosis in population biobanks: 27 loci and shared associations with skeletal structure. *Nat Commun*. 2023;14(1):157.

Schellekens GA, Visser H, De Jong BAW, Van Den Hoogen FHJ, Hazes JMW, Breedveld FC, et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. *Arthritis Rheum*. 2000;43(1):155–63.

Scherer HU, Häupl T, Burmester GR. The etiology of rheumatoid arthritis. *J Autoimm*. 2020;110:102400.

Shojania KG, Burton EC, Mcdonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: A systematic review. *JAMA*. 2003;289(21):2849–56.

Singh JA, Holmgren AR, Noorbaloochi S. Accuracy of veterans administration databases for a diagnosis of rheumatoid arthritis. *Arthritis Rheum*. 2004;51(6):952–7.

Smith V, Ickinger C, Hysa E, Snow M, Frech T, Sulli A, et al. Nailfold capillaroscopy. *Best Pract Res Clin Rheumatol*. 2023;37(1):101849.

Solomon A, Ngandu T, Soininen H, Hallikainen MM, Kivipelto M, Laatikainen T. Validity of dementia and Alzheimer's disease diagnoses in Finnish national registers. *Alzheimers Dement*. 2014;10(3):303–9.

Soussi B, Cordtz R, Kristensen S, Bork C, Christensen J, Schmidt E, et al. Incidence and prevalence of rheumatoid arthritis in Denmark from 1998 to 2018: a nationwide register-based study. *Scand J Rheumatol*. 2022;51(6):481–9.

Sparks JA. In the Clinic® rheumatoid arthritis. Vol. 170, *Annals of Internal Medicine*. American College of Physicians; 2019. P. ITC1–15.

Statistics Finland. About Statistics Finland [Internet]. 2023 [cited 2024 5 Aug]. Available from: <https://stat.fi/org/index_en.html>.

Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015;12(3):e1001779.

Summerton N. Making a diagnosis in primary care: symptoms and context. *Br J Gen Pract*. 2004;54(505):570–1.

Sund R. Quality of the Finnish Hospital Discharge Register: A systematic review. *Scand J Public Health*. 2012;40(6):505–15.

Tan YK, Conaghan PG. Imaging in rheumatoid arthritis. *Best Pract Res Clin Rheumatol.* 2011;25(4):569–84.

Tasca C, Rapetti M, Carta MG, Fadda B. Women and hysteria in the history of mental health. *Clin Pract Epidemiol Ment Health.* 2012;8:110–9.

Tenstad H, Nilsson A, Dellgren C, Lindegaard H, Rubin K, Lillevang S. Predictive values of anti-cyclic citrullinated peptide antibodies and rheumatoid factor in relation to serological aspects of the ACR/EULAR 2010 classification criteria for rheumatoid arthritis. *Scand J Rheumatol.* 2020;49(1):18–20.

Thammasitboon S, Thammasitboon S, Singhal G. Diagnosing diagnostic error. *Curr Probl Pediatr Adolesc Health Care*. 2013;43(9):227–31.

THL. About THL [Internet]. 2023 [cited 2023 15 Jun]. Available from: <https://thl.fi/en/web/thlfi-en/about-us/about-thl>.

THL. Care Register for Health Care [Internet]. 2023 [cited 2023 23 Nov]. Available from: <https://thl.fi/en/web/thlfi-en/statistics-and-data/data-and-services/register-descriptions/care-register-for-health-care>.

THL. Hospital benchmarking [Internet]. 2023 [cited 2023 17 Nov]. Available from: <https://thl.fi/en/web/thlfi-en/statistics-and-data/data-and-services/quality-and-statistical-principles/quality-descriptions/hospital-benchmarking>.

THL. Needs assessment of health and social services funding (in finnish) [Internet]. 2022 [cited 2023 17 Nov]. Available from: <https://www.julkari.fi/bitstream/handle/10024/144366/URN_ISBN_978-952-343-876-7.pdf?sequence=1&isAllowed=y>.

THL. Statistics by topic [Internet]. 2023 [cited 2023 9 Nov]. Available from: <https://thl.fi/en/web/thlfi-en/statistics-and-data/statistics-by-topic>.

THL. Statistics on health care services [Internet]. 2023 [cited 2023 17 Nov]. Available from: <https://thl.fi/en/web/thlfi-en/statistics-and-data/statistics-by-topic/health-care-services>.

THL. The Finnish Rheumatology Quality Register (in Finnish) [Internet]. 2023 [cited 2023 15 Jun]. Available from: <https://repo.thl.fi/sites/laaturekisterit/reumarekisteri/>.

Thomas SL, Edwards CJ, Smeeth L, Cooper C, Hall AJ. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? *Arthritis Rheum*. 2008;59(9):1314–21.

Thygesen LC, Ersbøll AK. When the entire population is the sample: Strengths and limitations in register-based epidemiology. *Eur J Epidemiol*. 2014;29(8):551–8.

Tomelleri A, van der Geest KSM, Khurshid MA, Sebastian A, Coath F, Robbins D et al. Disease stratification in GCA and PMR: state of the art and future perspectives. *Nat Rev Rheumatol*. 2023;19(7):446–59.

Valenzuela A, Yaqub A, Fiorentino D, Krishnan E, Chung L. Validation of the ICD-9-CM code for systemic sclerosis using updated ACR/EULAR classification criteria. *Scand J Rheumatol*. 2015;44(3):253–5.

Van Den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: An american college of rheumatology/European league against rheumatism collaborative initiative. *Arthritis Rheum*. 2013;65(11):2737–47.

Vesterinen T, Salmenkivi K, Mustonen H, Kuopio T, Lappi-Blanco E, Paavonen T, et al. Performance of Finnish biobanks in nationwide pulmonary carcinoid tumour research. *Virchows Archiv.* 2020;476(2):273–83.

Volkmann ER, Andréasson K, Smith V. Systemic sclerosis. Lancet. 2023;401(10373):304–18.

Vuori MA, Laukkanen JA, Pietilä A, Havulinna AS, Kähönen M, Salomaa V, et al. The validity of heart failure diagnoses in the Finnish Hospital Discharge Register. *Scand J Public Health*. 2019;48(1):20–28.

Vuorinen AL, Lehto M, Niemi M, Harno K, Pajula J, van Gils M, et al. Pharmacogenetics of anticoagulation and clinical events in warfarin-treated patients: A register-based cohort study with biobank data and national health registries in Finland. *Clin Epidemiol.* 2021;13:183–95.

Waldenlind K, Eriksson JK, Grewin B, Askling J. Validation of the rheumatoid arthritis diagnosis in the Swedish National patient register: A cohort study from Stockholm County. *BMC Musculoskelet Disord.* 2014;15:432.

Weman L, Salo H, Kuusalo L, Huhtakangas J, Kärki J, Vähäsalo P, et al. Initial presentation of early rheumatoid arthritis. *PLoS One*. 2023;18(7):e0287707.

WHO. ATC Structure and principles [Internet]. 2024 [cited 2024 4 Jul]. Available from: <https://atcddd.fhi.no/atc/structure_and_principles/>.

WHO. International Classification of Diseases (ICD) [Internet]. 2023 [cited 2023 10 Nov]. Available from: <https://www.who.int/standards/classifications/classification-of-diseases>.

Widdifield J, Bernatsky S, Paterson JM, Tu K, Ng R, Thorne JC, et al. Accuracy of Canadian health administrative databases in identifying patients with rheumatoid arthritis: A validation study using the medical records of rheumatologists. *Arthritis Care Res (Hoboken)*. 2013;65(10):1582–91.

Zhong L, Pope M, Shen Y, Hernandez JJ, Wu L. Prevalence and incidence of systemic sclerosis: A systematic review and meta-analysis. *Int J Rheum Dis*. 2019;22(12):2096–107.

Ziswiler HR, Urech R, Balmer J, Ostensen M, Mierau R, Villiger PM. Clinical diagnosis compared to classification criteria in a cohort of 54 patients with systemic sclerosis and associated disorders. *Swiss Med Wkly*. 2007;137(41–42):586–90.

List of Figures and Tables

Figures

Tables

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2024

ISBN 978-951-29-9875-3 (PRINT)
ISBN 978-951-29-9876-0 (PDF)
ISSN 0355-9483 (Print)
ISSN 2343-3213 (Online)