Using Electronic Health Record Data to Evaluate the Epidemiology and Management of Inflammatory Arthritis

Samantha Sarah Rosemary Crossfield

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Leeds Institute for Data Analytics

&

The University of Leeds

School of Medicine

Leeds Institute of Rheumatic and Musculoskeletal Medicine

February 2021

Intellectual Property and Publication Statements

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

The publications from this thesis are listed on pages III-V.

Publications and Presentations Arising from this Thesis

Original articles

Crossfield SSR, Marzo-Ortega H, Kingsbury SR, Pujades-Rodriguez M, Conaghan PG (in review) Changes in ankylosing spondylitis incidence, prevalence and time to diagnosis over two decades

Crossfield SSR, Buch MH, Baxter P, Kingsbury SR, Pujades-Rodriguez M, Conaghan PG (2021) Changes in the pharmacological management of rheumatoid arthritis over two decades. *Rheumatology*

Crossfield SSR, Buch MH, Baxter P, Kingsbury SR, Pujades-Rodriguez M, Conaghan PG (2020) Has modern UK management of RA led to a reduction in use of steroids and NSAIDs? 20-year data from the clinical practice research datalink. *Rheumatology*. 59 (2)

Crossfield SSR, Hui Lai LY, Kingsbury SR, Baxter P, Pujades-Rodriguez M, Conaghan PG (2019) Variation in Methods, Results and Reporting in Electronic Health Record-based Studies Evaluating Routine Care in Gout: A Systematic Review. *PLOS ONE*

Co-author contribution

Changes in ankylosing spondylitis incidence, prevalence and time to diagnosis over two decades:

SSRC, HMO, SRK, MP-R and PGC conceived the study. SSRC designed the analyses with guidance from MP-R and PGC, and SSRC collected the data. SSRC had access to the data and conducted the analyses. SSRC generated the figures and tables and drafted the manuscript with review from the other authors.

Changes in the pharmacological management of rheumatoid arthritis over two decades:

SSRC, MHB, SRK, MP-R and PGC conceived the study. SSRC designed the analyses with guidance from MP-R and PB, and SSRC collected the data. SSRC had access to the data and conducted the analyses. SSRC generated the figures and tables and drafted the manuscript with review from the other authors.

Has modern UK management of RA led to a reduction in use of steroids and NSAIDs? 20-year data from the clinical practice research datalink:

SSRC, MHB, SRK, MP-R and PGC conceived the study. SSRC designed the analyses with guidance from MP-R and PB, and SSRC collected the data. SSRC had access to the data and conducted the analyses. SSRC generated the figures and tables and drafted the manuscript with review from the other authors.

Variation in Methods, Results and Reporting in Electronic Health Record-based Studies Evaluating Routine Care in Gout: A Systematic Review:

SSRC, MP-R and PGC conceived the study. SSRC wrote the proposal and registered the study with PROSPERO. SSRC designed the search strategy and data extraction protocol with guidance from PB and MP-R. SSRC performed the literature search. SSRC and LLYH screened the literature and extracted the data and assessed study quality; MP-R or SRK resolved discrepancies. SSRC generated the figures and tables and drafted the manuscript with review from the other authors.

Conference Presentation: Oral

Crossfield SSR, Buch MH, Baxter P, Kingsbury SR, Pujades-Rodriguez M, Conaghan PG (2020) Has modern UK management of RA led to a reduction in use of steroids and NSAIDs? British Society for Rheumatology Conference 2020; Glasgow, UK

Crossfield SSR, Hui Lai LY, Kingsbury SR, Conaghan PG, Pujades-Rodriguez M (2019) Methods, Results and Reporting in Electronic Health Record-based Studies of Routine Care in Gout. Leeds Annual Statistics Research Conference 2019; Leeds, UK

Crossfield SSR (2017) Using Routine Data to Inform Patient Management in Musculoskeletal Disease. Farr & EFMI Informatics for Health Conference 2017; Manchester, UK

Conference Presentation: Poster

Crossfield SSR, Hui Lai LY, Kingsbury SR, Conaghan PG, Pujades-Rodriguez M (2019) Methods, Results and Reporting in Electronic Health Record-based Studies of Routine Care in Gout: A Systematic Review. Leeds Annual Statistics Research Conference 2019; Leeds, UK

Crossfield SSR, Conaghan PG, Pujades-Rodriguez M, Johnson O, Baxter P, Kingsbury S (2018) The Burden and Management of Inflammatory Musculoskeletal Disease. University of Leeds Faculty of Medicine and Health Postgraduate Research Conference 2018; Leeds, UK

Crossfield SSR, Johnson O, Fleming T (2016) Large Scale Infrastructure for Health Data Analytics. IEEE International Conference for Health Informatics 2016; Chicago, USA

Acknowledgements

I would like to thank my supervisors Professor Philip Conaghan, Dr Mar Pujades-Rodriguez, Professor Paul Baxter, Associate Professor Sarah Kingsbury and Owen Johnson for their support and contribution to the work that has been undertaken. I thank the Medical Research Council Leeds Medical Bioinformatics Centre for the studentship grant that funded the research for this PhD, and the Leeds Institute of Rheumatic and Musculoskeletal Medicine that funded the Clinical Practice Research Datalink academic dataset licence. I would also like to thank the Data Analytics Team at the Leeds Institute for Data Analytics for their assistance in extracting the CPRD dataset. I would like to thank Dr Lana Lin Hui Lai for helping to extract the data for the systematic literature review and helping to assess the risk of bias in the literature. I would like to thank Professor Maya Buch for helping to design the study of prescribing patterns in rheumatoid arthritis and reviewing the associated manuscript. I would like to thank Dr Helena Marzo-Ortega for helping to design the study of the incidence, prevalence and time to diagnosis in ankylosing spondylitis and reviewing the manuscript. I would like to thank my family for their support and positivity.

Abstract

In healthcare, there are opportunities to utilise the growth of routine data capture in developing real-world evidence of chronic disease. Inflammatory arthritis encompasses a number of chronic diseases including gout, rheumatoid arthritis (RA) and ankylosing spondylitis (AS), for which timely treatment is necessary to limit joint damage. The hypothesis underlying this thesis is that the epidemiology and management of inflammatory arthritis can be evaluated using routine electronic health record (EHR) data. This was investigated through literature reviews and retrospective studies using a population-based primary care dataset.

Gout, AS and RA studies have used EHR data, and this thesis identified variation in methods that influenced reported trends in epidemiology and management. For future studies, considerations were raised for improving the reporting and assessment of EHR-pertinent biases.

Incidence and prevalence are uncertain in AS, and have not been investigated in RA in recent years following the incentivisation of diagnostic recording. Between 1998 and 2017, this thesis identified that AS incidence declined for ten years before it stabilised, while RA incidence trends were unclear, and prevalence rose in older patients. In an ageing population, managing these diseases is important and studies should consider changes in coding practice in the study period.

There have been efforts to reduce diagnostic delay in AS. This thesis found no improvement in time to diagnosis over two decades, largely driven by delay in rheumatology referral. This is concerning given the importance of treatment in early AS.

In RA, shifts in management principles have increased DMARD prescribing over time. This thesis identified that the prescribing of potentially toxic corticosteroids and non-steroidal anti-inflammatories nonetheless persisted across the last 20 years with suboptimal prophylactic therapy.

This thesis provides evidence of, and raises considerations for further improving, the use of EHR data in evaluating the epidemiology and management of inflammatory arthritis.

Table of Contents

List of Tables

List of Figures

Abbreviations

A&E	Accident and emergency
ACR	American College of Physicians
AHR	Adjusted hazard ratio
AIC	Akaike information criterion
Anti-CCP	Anti-cyclic citrullinated peptide
AOR	Adjusted odds ratio
APC	Annual percentage change
AS	Ankylosing spondylitis
ASAS	ASsessment in Ankylosing Spondylitis
BMI	Body mass index
BMJ	British Medical Journal
BNF	British National Formulary
BSR	British Society for Rheumatology
CI	Confidence interval
CINAHL	Cumulative Index of Nursing and Allied Health Literature
CKD	Chronic kidney disease
COPD	Chronic obstructive pulmonary disease
CoR	Comprehensiveness of reporting
CPRD	Clinical Practice Research Datalink
CRD	Centre for Reviews and Dissemination
CRP	C-reactive protein
CVD	Cardiovascular disease
DANBIO	Danish Registry for Biologic Therapies in Rheumatology
dL	Decilitre
dm+d	Dictionary of Medicines and Devices
DMARD	Disease-modifying anti-rheumatic drug
DMP	Data management plan
EAM	Extra-articular inflammatory manifestations
EHR	Electronic health record
ESR	Erythrocyte sedimentation rate
EULAR	European Alliance of Associations for Rheumatology
GDPR	General Data Protection Regulation

GI	Gastro-intestinal
GP	General practitioner
HES	Hospital Episode Statistics
HITSP IS	Healthcare Information Technology Standards Panel Interoperability Specification
HLA	Human leukocyte antigen
HR	Hazards ratio
HTA	Health Technology Assessment
IBD	Inflammatory bowel disease
IBP	Inflammatory back pain
ICD-10	International Classification of Diseases, tenth revision
ID	Identifier
IMD	Index of Multiple Deprivation
IQR	Interquartile range
IR	Incidence rate
IRC	Integrated Research Campus
IRR	Incidence rate ratio
ISAC	Independent Scientific Advisory Committee
ISO	International Organisation for Standardisation
LIDA	Leeds Institute for Data Analytics
MEDLINE	Medical Literature Analysis and Retrieval System Online
MeSH	Medical Subject Headings
mg	Milligram/s
MPR	Medication possession ratio
MRI	Magnetic resonance imaging
MSU	Monosodium urate
non-bDMARDs	Non-biologic DMARDs
NHS	National Health Service
NSAID	Nonsteroidal anti-inflammatory drug
ONS	Office for National Statistics
OR	Odds ratio
PDC	Proportion of days covered
PICOS	Population, Intervention, Comparison and Outcomes Study
PPI	Proton-pump inhibitor
PPV	Positive predictive value

PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	International Prospective Register of Systematic Reviews
QOF	Quality and Outcomes Framework
RA	Rheumatoid arthritis
RCV2	Read Code Version 2
RECORD	REporting of studies Conducted using Observational Routinely-collected Data
RoB	Risk of bias
ROBINS-I	Risk of Bias in Non-randomised Studies – of Interventions
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SD	Standard deviation
SQL	Structured Query Language
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUA	Serum urate
THIN	The Health Improvement Network
TNF	Tumour necrosis factor
UK	United Kingdom
ULT	Urate lowering therapy
UoL	University of Leeds
USA	United States of America
UTS	Up to standard
VRE	Virtual research environment

Chapter 1 Introduction

1.1. Background

In recent years the routine collection of data in healthcare services has grown considerably, bringing immense opportunity for research and the development of real-world evidence. Sources of routinely captured health data include electronic health records (EHRs), disease registries, medical imaging, and laboratory, health insurance, sensor, genetic and social media data. Estimates suggest that 30 petabytes of health data are generated annually and the volume of growth is expected to increase by ~40% each year (1, 2). In the United States (USA) alone, health data grew rapidly to 150 exabytes by 2011 and such "big data" is complex, variable and requires advanced technology for data handling (3). The growing uptake of clinical data systems in clinical practice can complement, but also disrupt, standard practices. For clinical system users, the training provided in using new systems may be insufficient, and data entry may be time-consuming (4). Nonetheless, the capacity for routine electronic data capture to support routine healthcare delivery while reducing costs has been lauded by a number of governments (5, 6). The huge array and amount of data bring opportunity for new approaches in data analytics and machine learning and for research to present more thorough "real-world evidence" that contributes to better outcomes (7, 8). Big data can help to overcome the limitations in health research brought by cohort selection biases and the practical restrictions in cohort size and follow-up duration that apply in traditional clinical trials and prospective cohort studies (9). Use of this pre-existing data in research can produce broader real-world evidence on healthcare and support the generalisation of findings from clinical trials and cohort studies to larger, more inclusive populations of patients (10). Despite challenges to research including lack of standardisation across health data, and issues of access, privacy, and quality assurance (11), real-world evidence may inform important aspects of health including the burden and management of disease (12).

Primary care plays an important role in the diagnosis and long-term management of many chronic diseases including inflammatory arthritides. Symptoms of many chronic diseases may first be presented to clinicians in primary care, which in some countries act as a 'gateway' in referring access to other health and care settings (13). The primary care population may be more representative of the general population, encompassing all stages of disease severity and the disease life-course. Primary care

data was therefore chosen in this study of the epidemiology and management of inflammatory arthritis.

The sources of data that are routinely collected in health settings, and inform research, include EHR databases, health insurance databases and registries. Electronic health records are routinely used to capture relevant aspects of care including symptoms, diagnoses, laboratory tests and prescriptions along with demographic information. Electronic health record uptake is often highest in primary care and in the United Kingdom (UK), most general practitioner (GP) practices use one of a small number of EHR clinical systems, which has facilitated data extraction for research on a large population (14). These data may, through secondary use in research, provide information on the epidemiology, timeliness of diagnosis, and real-world management of inflammatory arthritis in the general population, regardless of disease severity. Comparisons can also be made to data for a non-disease cohort derived from the general primary care population. Like EHRs, health insurance databases inform research as a secondary use, but may not be derived from a representative population. Further, the focus of data capture in insurance settings is claims-driven and data may be recorded retrospectively from clinical notes made during consultations. In contrast, data in registries, such as the nationwide Danish Registry for Biologic Therapies in Rheumatology (DANBIO), is recorded to facilitate research (15). However, registries focus on a specific disease, event or therapy, rather than reflecting healthcare across a general population. In addition, registries predominantly capture prescribing and management information from secondary care where it would not be as readily available from EHRs on a multi-site scale for research, given the difference in EHR uptake, and the preponderance of single-site EHR systems that create a ‘silo effect’, in secondary care (15, 16). For this study of disease epidemiology and management using primary care data, EHR data was selected as the most relevant source of routine health data on a representative, general population.

The UK has a long record of using EHRs in primary care, commencing in the 1970s and increasing to include 96% of GP practices by 1996 (17, 18). In addition, processes of payment for primary care service delivery, introduced in 2004, have improved aspects of the data quality in GP EHRs (19, 20). Therefore, UK-based EHR data are used in this thesis and there is a focus in subsequent chapters on UK EHR data sources and UK regulations.

Principles of disease management can differ between countries, related to national guidelines and historic practice. In the UK, many aspects of health and social care are

delivered ‘free at the point of delivery’ by the National Health Service (NHS), with pathways informed by national evidence-based guidelines published by the National Institute for Health and Care Excellence (21). The chapters in this thesis consequently have a particular focus on UK healthcare given that UK data are used and that principles of service delivery and pharmacologic management may differ across countries.

The burden of chronic disease represents a growing challenge for health care systems and health research globally (22, 23). Chronic musculoskeletal disease including inflammatory arthritis is a leading cause of disability and affects 2 billion people worldwide, and an estimated 14% of the UK adult population (24, 25). Inflammatory arthritis encompasses a group of diseases in which joint inflammation is caused by an overactive immune system. Inflammatory arthritides are characterised by painful joints with swelling and stiffness, while systemic manifestations include cardiovascular disease, weight loss and fatigue (26). In these diseases, sustained inflammation can lead to irreversible joint damage, extra-articular manifestations and further impaired quality of life. The inflammatory arthritides include gout, the spondyloarthropathies (including ankylosing spondylitis, AS), and rheumatoid arthritis (RA). The aetiology is often poorly understood, although genetic and environmental factors play an important role (27, 28). Timely diagnosis and early pharmacologic management of these conditions is vital in preventing symptoms, disease progression and disability (29).

The purpose of this thesis is to explore the opportunities afforded by routinely collected EHR data to investigate the epidemiology and management of inflammatory arthritis as an exemplar for chronic disease. The inflammatory arthritides gout, AS and RA were chosen in this study as conditions where early diagnosis and timely pharmacologic prescribing is important for quality of life as these are recorded in EHRs. Gout was selected as an exemplar of a common disease that is diagnosed and treated predominantly in primary care, having an established practice of pharmacologic-led management. In the UK, GP prescribing data is complete in EHRs and so EHR-based research in gout management can inform an understanding of the utility of comprehensive EHR data on a large cohort. In contrast, AS was selected as a less common disease in which the aetiology and prevalence is uncertain, where there has been an evolving disease concept and substantial delay in diagnosis, in order to investigate the constraints of primary care EHR-based research in diseases that may be under-recognised in primary care. Rheumatoid arthritis was selected as an exemplar of a common disease in which management guidelines have shifted in recent years, so that the utility of longitudinal EHR data in examining temporal trends could be

investigated. In these diseases, onset can be relatively early in life (e.g. around 25 years in AS (30)) and continual medication is commonly recommended (31-33), so the lifelong follow-up afforded by EHRs is pertinent. Over the life-course, these conditions can have major impact on quality of life, physical function and the ability to work, as well as increasing the risk of comorbidities and mortality (34-36). Additionally, they are a major and growing health service burden in ageing and increasingly obese populations (37). This thesis explores EHR-based research in these conditions and investigates real-world evidence in different aspects of epidemiology and management.

1.2. Thesis Outline

The hypothesis underlying this thesis is that EHR-based research can provide information on disease epidemiology and management, relevant to clinicians and decision-makers.

Chapter 2 presents the background for this thesis, presenting a knowledge base that forms the underlying rationale for this research. Electronic health records and EHR-based research are described, with detail provided on the common databases used in UK EHR-based research in inflammatory arthritis. A description is given on the three inflammatory arthritides that are investigated in this thesis – gout, AS and RA. This includes the epidemiology, diagnosis and management, and a summary of the breadth of EHR-based research in these diseases. Informed by this background, the aims and objectives of the thesis are described.

Chapter 3 gives a detailed description of the study design and data source used in this thesis. A feasibility study is also presented, which was conducted to explore the suitability of the EHR data source for the research described in subsequent chapters.

Chapters 4-6 present literature reviews of the EHR-based research in gout, AS and RA. In gout, the review aimed to describe the variation in methodological approaches and study findings, and to examine the comprehensiveness of reporting of EHR data utilisation and risk of bias. This review focused on studies of pharmacologic management given this is a focus of this thesis. In AS, the literature was reviewed to identify all common study themes, key findings, reporting on prescribed medication, the data source and study timeframe. These were also reviewed in RA for UK studies, along with methods and potential research gaps. The definitions of gout, AS and RA applied to EHR data sources were reviewed. These chapters identify the study themes,

the impact of methodology on the results, and the under-reporting of EHR data handling.

Chapters 7-9 present research into the epidemiology and management of inflammatory arthritis, conducted on GP EHR data: the incidence and prevalence of AS and RA (Chapter 7), the time to diagnosis in AS (Chapter 8), and prescribing and prophylaxis in RA (Chapter 9). The regional and demographic variation in incidence and prevalence in AS and RA, the diagnostic delay in AS, and changes in RA pharmacologic management are identified over the past two decades.

Chapter 10 gives a detailed and critical discussion of the thesis, including the aspects of methodology common across Chapters 7-9, highlighting the main contributions of this work and suggesting future directions in this field.

Chapter 2 Background Literature

2.1 Introduction

This chapter introduces the development of EHRs and research using EHRs, and provides a narrative review of the three inflammatory arthritides that will be investigated in this thesis. The rise in routine data recording and its increasing accessibility and utility in research is described. Then the clinical features, epidemiology, common comorbidities, diagnosis and management of gout, AS and RA are defined. Further, the EHR-based research conducted for each of these diseases is summarised to provide background, ahead of the literature reviews presented in Chapters 4-6. Finally, this chapter defines the thesis aims and objectives.

2.2 Electronic Health Records

Health records facilitate continuity in, and payment of, healthcare provision. The concept of EHRs developed in the 1970s with the aim of improving efficiency over paper versions (16). Electronic health records maintain a longitudinal record of a patient's health and interaction with the healthcare system, and may be reviewed, managed and recorded by clinicians, patients and healthcare organisations (38). Information can cover appointments, symptoms, diagnoses, referrals, laboratory tests, procedures, vaccinations and the prescription of medications and devices. EHR data are rich and varied; it may contain input from diverse specialties in forms including free text, coded fields, scans and images (39). Their role in the payment for service provision and in maintaining a legal record has driven EHR data quality improvements (19, 20). Despite barriers. including concerns over the 'de-personalisation' of healthcare (with the computer acting as a barrier in the consultation room), there is growing uptake of EHR systems across organisations globally, especially in primary care settings (14, 40-44).

Electronic health records are designed to structure electronic record-keeping in a clinical domain and meet data security, performance and usability criteria as well as local and national requirements (45, 46). Some organisations develop and own an 'in-house' EHR system, which is more common in hospital settings or across a health insurance system, while other EHR systems are vendor-owned and licenced use by health organisations is established through a procurement process (47). Even vendor-

owned EHR systems are often designed for one setting and encode specialty-specific requirements in patient management and differences in clinical focus and coding terminology (48). However, the range of EHR systems available and their specialisation can limit interoperability – data exchange between EHR systems and shared understanding across organisations of the content, meaning and context of information. This can hinder data sharing that could otherwise aid inter-organisational clinical workflow (49).

There have been increasing efforts to facilitate EHR utilisation and the use of routinely captured EHR data to generate novel insights and transform delivery of healthcare. In the USA, federal incentives for EHR adoption were included in the 2009 Recovery and Reinvestment Act (5). Correspondingly, the proportion of hospitals and physician offices in the USA with EHRs that met minimum standards was 9% and 17% respectively in 2008 and rose rapidly afterwards to 96% and 78% in 2017 (50, 51). Once the adoption was sufficient to fuel EHR-based research, in 2010, various USA federal agencies contributed to funding the development of a Healthcare Information Technology Standards Panel Interoperability Specification (HITSP IS #158) for the Use of EHRs for Clinical Research (52, 53). In the UK, EHR uptake progressed earlier, with Payne et al. in 2011 describing the achievement of “national-scale clinical information exchange”, with all GPs using EHR systems (17, 54). In 2010, the NHS Quality, Innovation, Productivity and Prevention programme encouraged the use of such existing resources, and in 2012 the Department of Health called for more efficient EHR research (55, 56). The increasing routine use of EHRs in recording clinical practice and government support for its secondary use in research has facilitated growing opportunities for EHR-based research. Accordingly, while studies of EHRs initially focused on adoption of EHRs, growing numbers are investigating their utilisation and the methods of EHR-based research (57).

2.2.1 EHR-based Research

Electronic health records hold prospectively collected information on healthcare delivery for a population that can span generations, which brings opportunities for research. EHR data can inform population-based understanding of epidemiology, real-world patterns of disease management and guideline compliance and the long-term effectiveness and safety of, and interactions between, medications (58-61). Using EHR data, regional variation, multi-morbidity and multidisciplinary clinical pathways, and routine adherence to treatment, can be studied at a population-level scale larger than that which traditional clinical trials could achieve. The lifelong span of EHRs facilitates

longitudinal research into associations with time-varying exposures and chronic conditions with features that take many years to develop, where multiple treatments may be used over time (62). EHR data from large populations can facilitate investigations of less common diseases and help to overcome low power and random error prone in smaller studies, and highlight both rare events that may be important risk factors and potential risk factors for rare events.

Such secondary data use facilitates research in a cost-efficient and timely manner that is inclusive of the general population. EHR-based research can mitigate some of the huge costs in large-scale recruitment that would be required in setting up a traditional cohort study or clinical trial (63). Using longitudinal EHR data, research can be conducted across the patient life-course and across generations, without, in comparison to cohort studies, employing high costs for lengthy follow-up or suffering the same degree of loss to follow-up (64). The comprehensive population coverage in EHRs can complement clinical trials that have narrow patient inclusion criteria, often excluding comorbid or elderly patients, or limited follow-up duration, and can highlight trends that require investigation in further trials (65). EHR data can also overcome the veracity challenges of patient reported data in survey or cohort studies.

Routinely collected health data also presents an opportunity for alternative, additional research practices. The richness of data in a variety of formats facilitates a range of studies applying tools such as natural language processing, process mining and machine learning, and exploiting super-computers. Electronic transfer readily enables simultaneous studies to be conducted, even across countries, on the same data. Research can be integrated in EHR systems, both enabling clinicians and patients to participate in research through less obtrusive means, e.g. by completing surveys, through pragmatic allocation to trial intervention, and disseminating EHR-based decision support (66-68). Where clinicians are tasked with completing case report forms in research trials, auto-population from EHRs can streamline data collection and ease the burden on clinicians of contributing time to research (69). In such ways, EHRs can facilitate alignment between research and healthcare provision in a “learning health system”, with knowledge generated from information captured during healthcare delivery supporting the continuous improvement of delivery processes (70).

As with other data sources used in developing real-world evidence, EHR data differs from that traditionally collected for research and there are limitations in the ‘repurposing’ of routine clinical EHR data for research. There are accessibility constraints and issues of data ownership, privacy and patient confidentiality. Data may

be of varying quality in aspects including completeness, correctness and currency (71). The quality of data recording may differ within and between organisations and vary over time, being affected by training, experience, changes in the incentives for data recording, the coding terminology used and the design of the EHR system. In the UK, some factors such as medication prescribing and GP laboratory test requests are fully captured electronically while aspects such as clinical reasoning and decision-making processes may not be. Strategies for handling missing data exist, though these rely upon assumptions regarding the reasons behind missing data, which may or may not be missing at random (71-73). Secondary use requires understanding of how EHRs are used in clinical practice and of any changes in recording practices that occur in the study timeframe in order to design and interpret studies appropriately. Further, patient healthcare utilisation may vary based on unknown variables and not all of the data required to address a study aim may be recorded in EHRs. A study may require the linkage of data collected from different sources with varied data structures, which brings technical, legal and privacy issues. While the volume, veracity and variety of data in EHRs can be utilised in research to develop complex algorithms, difficulty in interpreting and validating these is known as the 'black box' problem (74, 75). For such reasons, the full potential of EHR data has not yet been realised.

There are also ethical and legal requirements in data protection and privacy that apply to EHR-based research. These relate to all stages of data handling, including inter-organisational data linkage and storage. In the USA, The Health Insurance Portability and Accountability Act legislates requirements for the privacy and security of personal health information (76). In the European Union, the handling of personal data (that relating to an identified or identifiable individual) is governed by the General Data Protection Regulation (GDPR) (77) and many countries have similar legislative structures. In the UK, the Data Protection Act 2018 and common law duty of confidentiality apply to personal data (78, 79). Confidential information requires a legal gateway for disclosure, such as consent, the public interest, a legal obligation or approval to do so from the Secretary of State for Health (80). One common approach to meeting such requirements is to create a research dataset that is "anonymous in such a manner that the data subject is not or no longer identifiable" beyond 'reasonable effort' (77, 81, 82), so that the data are no longer considered to be confidential or personal. However, weaknesses in public engagement and understanding, and concerns around data protection and privacy, can form perceived barriers to the exploitation of EHR data in research (83, 84). Further, challenges remain in maintaining privacy while using EHRs to integrate research into routine patient management (85).

2.2.2 EHR Research Databases

In addition to national registries and institution-specific health insurance databases, in many countries, a small but growing number of organisations have developed EHR research databases to facilitate accessibility. In countries including the Netherlands, the UK, Spain and Italy, the majority of GP practices contribute to a research database (86, 87). Identifiable information is often excluded so that data may be made available either freely (e.g. the Medical Information Mart for Intensive Care database in the USA) or under a licence for research that meets academic criteria (e.g. the Clinical Practice Research Datalink [CPRD] in the UK) (88, 89). Many of these research databases contain data from a single organisation, or multiple organisations in the same health setting that use the same clinical EHR system.

Research must consider not only the EHR data provenance but also the processes of data extraction and transformation that are employed during creation of the research database. Data may be reformatted during extraction, undergo transformation or be excluded during automated processes of data quality review. In some cases, such as for the CPRD database, contributing health organisations are provided training in coding and the quality of extracted data are assessed in order to enhance data quality for epidemiologic research (89). Alternatively, as in ResearchOne, another UK database, no transformations or data quality assessment may be applied, giving researchers control over any data cleaning (48). Such data may be more representative of EHRs generally, which may facilitate the development of decision-making tools intended for incorporation into EHRs. Open standards and collaborative efforts such as the Integrating the Healthcare Enterprise initiative attempt to facilitate standardisation in EHR research databases (90).

However, the capacity for EHR research databases containing data from across all health settings, e.g. primary and secondary care, at a population-level, are not fully realised. Often, such databases are derived from insurance settings using an integrated EHR system, such as Kaiser Permenante in the USA (91). An alternative source is organisations across health settings that use the same EHR system with a centralised, integrated EHR database e.g. SystmOne in the UK, which employs a modular approach to tailor the EHR view according to organisation type (66). However, there can be issues in representative coverage. Insurance system-derived research databases represent only a subset of the population demographic. Further, all organisations involved in a patient's care must contribute EHR data in order to form a comprehensive dataset at the patient-level. In addition, other issues in research, such as data validation, remain and research findings may require calibration before

translating to health organisations using different EHR systems. Country-wide, cross-organisational use of an interoperable shared care record, in a country with free healthcare at the point of delivery, as seen in Finland, is rare (92). The alternative for developing comprehensive and representative EHR research databases involves cumbersome data linkage. Given the complexity involved, some organisations maintain EHR-based research data derived from different EHR systems in separate research databases without cross-system linkage, or collect only aspects of data that are commonly available in all contributing EHR systems (93, 94).

Where data are collected from multiple organisations using different EHR systems, the different options and incentives embedded in those systems should be considered. Differences in system design can affect medication selection or the recording of diagnoses and test results using coded nomenclature rather than free text. For example, generic drugs are listed above branded drugs in some systems such as SystmOne. Coding conversion or transformations, and calibration and validation between data from different systems, may be necessary (94).

2.2.2.1 UK Databases

Electronic health record data from the UK is used in this thesis and so a background on the main sources of UK EHR data that are used in inflammatory arthritis research is presented. The sources commonly used are the CPRD, The Health Improvement Network (THIN) and Hospital Episode Statistics (HES). EHR data are commonly linked to Office for National Statistics (ONS) mortality statistics in mortality studies, and information on this source is also given.

The CPRD developed from the General Practice Research Database, which from 1987 had collected anonymous EHR data from consenting GP practices using the Vision EHR system (89). Over 674 GP practices (approximately 8% of UK GP practices) have contributed EHR data to the CPRD “GOLD” database, which contains data on over 11 million patients that have a comparable age, sex and ethnicity profile to the national census statistics and a body mass index distribution to the NHS Health Survey for England (89, 95, 96). Approximately 60% of these records have been linked with ONS death registry and HES data, although this does not cover hospital prescriptions, tests and accident and emergency (A&E) visits (89, 93, 97). Patient- and practice- level deprivation scores have also been linked to the data. The steps taken by the CPRD to improve data quality were described above (89). The database has flags for patient records that are deemed to be of “acceptable” quality for research and also contains

the date from which the quality of the data from each contributing GP practice was deemed to be “up to standard” (UTS) (89). Validation studies performed on the data have reported high validity of coding, although there is under-representation of acute events that present to secondary care, such as myocardial infarction (98-100).

Those GP practices using the Vision EHR system can also contribute anonymous EHR data to THIN, which holds longitudinal data from 532 GP practices, for over 10 million patients (101). As with the CPRD, contributing practices are provided with training (102). A study in 2011 reported comparable prevalence of major conditions (including diabetes, atrial fibrillation, obesity) and mortality rates to the UK population (101). The validity of AS diagnoses in THIN is also high (72% positive predictive value [PPV]) (62).

The HES database contains information on hospital admissions, outpatient visits and A&E attendances, which is collected monthly from NHS hospitals in England (93). Inpatient data has been collected from 1989 onwards, outpatient data from 2003 and A&E from 2007, with approximately 16 million, 60 million and 12 million episodes recorded annually (93). Patient information includes demographics, diagnoses and operations, and administrative information such as the methods of admission and discharge (93). Patient reported outcome measures also capture the patient perspective on specific procedures (93). However, data on prescribed medication or drugs issued through hospital pharmacies, is not available. HES data are based on aspects of data that are commonly captured across different hospital data systems and also fields that are derived by NHS Digital (93). HES data undergo quality checks and statistical disclosure control that suppresses small numbers and maintains patient confidentiality. The data are used in payment as well as in research and planning health services. HES data has been linked to ONS mortality data and a number of research databases including the CPRD and the UK Biobank, a large prospective cohort study dataset (89, 103).

ONS mortality statistics provide information collected weekly on deaths registered in England and Wales (104). The data include the clinically certified cause of death, patient age and sex, and geographic region (104). Death registration has been recorded using the International Classification of Diseases, tenth revision (ICD-10) coding system since 2001 and the data undergoes automated validation processes including checking for duplicates and completeness and comparison of the death date and registration date (105). Multiple causes of death can be recorded and a small proportion of deaths are labelled ‘uncertified’, e.g. where the medical certificate was completed by a doctor not fulfilling the legal requirements (105).

2.3 Inflammatory Arthritis

Electronic health records commonly hold comprehensive information on diagnoses and medications, two aspects that are important to the investigation of inflammatory arthritides, where early diagnosis and pharmacologic management play an important role. Therefore, this thesis utilises EHR data in investigating the epidemiology and routine management of such inflammatory arthritides.

Chronic diseases such as inflammatory arthritides are increasingly common and place a major burden on individuals and health economies. Inflammatory arthritides are generally conditions that result in inflammation of the joints and other connective tissues due to an overactive immune system. Inflammation is not appropriately self-regulated and timely pharmacologic treatment is required to limit or prevent irreversible damage. Environmental and genetic factors contribute to the development and progression of disease. For example, red meat and alcohol consumption can contribute to gout in susceptible individuals while AS and RA are associated with specific groups of human leukocyte antigen (HLA) alleles (27, 28). Diagnosis may be based on clinical history and symptoms, physical examination and, where needed, blood tests and scans. Gout and RA are the most common forms of inflammatory arthritis affecting peripheral joints while AS is a less common form of arthritis, predominantly affecting the spine. The following sub-sections review the characteristics, epidemiology and management of gout, AS and RA, and EHR-based research in these diseases.

2.3.1 Gout

2.3.1.1 Clinical Features

Gout results from monosodium urate (MSU) crystal deposition in articular and peri-articular tissues that leads to very painful and debilitating flares and consequent joint damage (106). Gout presents in acute episodic flares and may progress to chronic polyarticular gout (107). Acute gout presents with rapid onset of monoarthritis with intense pain, often within 24-48 hours. It affects the big toe (podagra) or foot in most initial cases, with the affected joint becoming intensely inflamed (108). Recurrent attacks are common, with approximately 80% of patients having further flares within three years (109). Over time, there may be chronic inflammatory and structural changes, as well as reduced quality of life and productivity loss (110). Between flares,

MSU deposition in extracellular fluid joints may continue asymptomatically if hyperuricaemia (high MSU serum level) is untreated. If untreated, MSU crystals can form under the skin, appearing as tophi (white nodules), and urolithiasis may develop, with severity usually being related to the duration of hyperuricaemia. The saturation point for MSU crystals is 6.8 milligrams (mg) per decilitre (dL) and hyperuricaemia, and risk of gout, is commonly defined at levels greater than 6 or 7 mg/dL (106, 111, 112).

2.3.1.2 Epidemiology

Gout is the commonest inflammatory arthritis, affecting approximately 2.5% of the UK population, particularly older patients and men (113). Prevalence varies between countries although is typically above 1% and highest in Taiwan, the Pacific and developed countries (114). In New Zealand, the estimated prevalence is 3.75% among adults, higher in regions with greater deprivation (115). In THIN and the CPRD, the incidence has been estimated as 2.68/1,000 person-years in 2000-2007 and 1.77/1,000 person-years in 2012 respectively, with risk of recurrent flare within one year increasing with age (116, 117). Studies suggest that prevalence is rising e.g. by 4.2% per year, and that incidence is rising most rapidly in post-menopausal women (113, 116, 118, 119).

Hyperuricaemia can lead to gout in susceptible individuals and the risk factors include diet, comorbidity and medication. The genetic contribution is poorly understood; a study of monozygotic twins reported greater influence from environmental than genetic factors, although genetic variants influence urate excretion and predict responsiveness to urate-lowering therapy (120, 121). Non-vegetable purine sources such as red meat, and dairy products, alcohol (beers and spirits), sugar-sweetened soft-drinks and shellfish can contribute to the risk of gout (27, 122). Dehydration can also be a trigger, with incident gout and recurrent flares being more common in summer (123). The risk of gout is increased by metabolic syndrome (e.g. hypertension), obesity, diabetes and cardiovascular disease (CVD), as well as early menopause (27). Kidney and liver disease can affect urate regulation and obstructive sleep apnoea may increase purine catabolism, increasing the risk of gout in affected patients (118, 124). Low-dose aspirin may increase urate retention and trigger gout, especially in women, as well as diuretics and immunosuppressants (125). Secondary gout can be common following chemotherapy or treatment to prevent transplant rejection. Factors contributing to the rising prevalence of gout include changes in diet, increasing obesity and comorbidity, and the rising use of loop diuretics for CVD (106).

2.3.1.3 Common Comorbidities

Gout is associated with increased risk of comorbidities and reduced overall survival (52). Gout shares common risk factors with a number of cardiovascular and metabolic diseases and patients with gout seem to accumulate comorbidities quicker (34). Compared with patients with osteoarthritis, patients with gout are more likely to have CVD, hypertension, chronic renal failure and diabetes (126). High serum urate (SUA) levels contribute to atrial remodelling and increased risk of atrial fibrillation (61). Gout may independently associate with increased risk of diabetes, especially in women (127). Risk of stroke, psoriasis, depression and anaemia is also higher in patients with gout (34). Further, pain from gout may make the management of comorbidities more complex (128). Urate lowering therapy (ULT) may also affect the risk of comorbidity, e.g. colchicine or nonsteroidal anti-inflammatory drug (NSAID) therapy may affect the risk of cardiovascular events, and there may be cardiovascular benefits of lowering SUA levels (110, 129). ULT may also increase the risk of cataracts in older patients and lessen the protective effect of gout in preventing Parkinson's (130, 131). Both gout and the attendant comorbidities contribute to increased mortality in patients with gout (34).

2.3.1.4 Diagnosis

Gout can be diagnosed by presence of urate crystals in synovial fluid or tophi, however microscopy is not always practical. In the UK, guidelines recommend diagnosis based on the assessment of flares, risk factors, and evidence of arthritis (132). Where patients present with red, hot swelling in a joint, diagnosis of gout can be aided by determining the rapidity of onset and frequency and duration of attacks, any previous attempted drug interventions and the impact on mobility and function. The cause of gout may be genetic where the age of onset is <30 years, and further investigation and aggressive treatment may be required. In older patients, presentation is often insidious and tophi may present early, especially in post-menopausal women. Diet and alcohol intake should be assessed as well as any use of urate-raising medication and the existence of associated comorbidities including metabolic disease, to determine the likelihood of gout. The examination of joints, especially of the toe and lower limbs, can reveal evidence of arthritis. Tophi may be present although commonly appear in untreated gout after ten years. Ultrasound and dual-energy computed tomography may assist in diagnosis and detecting tophi severity, but joint damage may not be visible until late in the disease course (106). Diagnostic challenges remain and late diagnosis is reasonably common, especially if septic arthritis occurs simultaneously or where pseudogout is suspected in patients with osteoarthritis (133). If the diagnosis is in

doubt, the aspiration of tophi or joint fluid can definitively determine the presence of urate crystals.

Table 1. Propositions to aid the diagnosis of gout

	Proposition	SOR (95% CI) VAS 100	A–B%*
1	In acute attacks the rapid development of severe pain, swelling, and tenderness that reaches its maximum within just 6–12 hours, especially with overlying erythema, is highly suggestive of crystal inflammation though not specific for gout	88 (80 to 96)	93
2	For typical presentations of gout (such as recurrent podagra with hyperuricaemia) a clinical diagnosis alone is reasonably accurate but not definitive without crystal confirmation	95 (91 to 98)	100
3	Demonstration of MSU crystals in synovial fluid or tophus aspirates permits a definitive diagnosis of gout	96 (93 to 100)	100
4	A routine search for MSU crystals is recommended in all synovial fluid samples obtained from undiagnosed inflamed joints	90 (83 to 97)	87
5	Identification of MSU crystals from asymptomatic joints may allow definite diagnosis in intercritical periods	84 (78 to 91)	93
6	Gout and sepsis may coexist, so when septic arthritis is suspected Gram stain and culture of synovial fluid should still be performed even if MSU crystals are identified	93 (87 to 99)	93
7	While being the most important risk factor for gout, serum uric acid levels do not confirm or exclude gout as many people with hyperuricaemia do not develop gout, and during acute attacks serum levels may be normal	95 (92 to 99)	93
8	Renal uric acid excretion should be determined in selected gout patients, especially those with a family history of young onset gout, onset of gout under age 25, or with renal calculi	72 (62 to 81)	60
9	Although radiographs may be useful for differential diagnosis and may show typical features in chronic gout, they are not useful in confirming the diagnosis of early or acute gout	86 (79 to 94)	93
10	Risk factors for gout and associated co-morbidity should be assessed, including features of the metabolic syndrome (obesity, hyperglycaemia, hyperlipidaemia, hypertension)	93 (88 to 98)	100

*A–B%: percentage of strongly to fully recommended, based on the EULAR ordinal scale (A = fully recommended, B = strongly recommended, C = moderately recommended, D = weakly recommended, E = not recommended).
CI, confidence interval; MSU, monosodium urate; SOR, strength of recommendation; VAS, visual analogue scale (0–100 mm, 0 = not recommended at all, 100 = fully recommended).

Reprinted with permission from British Medical Journal (BMJ) Publishing Group Ltd. (134)

2.3.1.5 Management

International guidelines on gout management, including publications by the British Society for Rheumatology (BSR), European Alliance of Associations for Rheumatology (EULAR), NICE and the American College of Physicians (ACR), agree on immediate treatment for acute gout, with rest and medication (33, 110, 135). Elevation of the affected joint and exposure in a cool environment, with application of topical ice, may ease symptoms (136). The guidelines all recommend prompt commencement of colchicine (e.g. 0.5-1 mg every two hours until toxicity sets in) or maximum-dose NSAIDs, or corticosteroids (intra-articular or oral) if these are not tolerated (Table 2, Figure 1, Figure 2) (129, 132). Pain reduction is the aim of acute management and if pain is not improved within 24 hours, combination therapy with NSAIDs and colchicine or interleukin-1 inhibitors may be beneficial (110). Intra-articular aspiration or opiate analgesic may also offer relief, though these are very infrequently used (122). Flare management should be followed by ongoing efforts to treat and prevent hyperuricaemia in order to reduce the risk of recurrent flares and comorbidity (137).

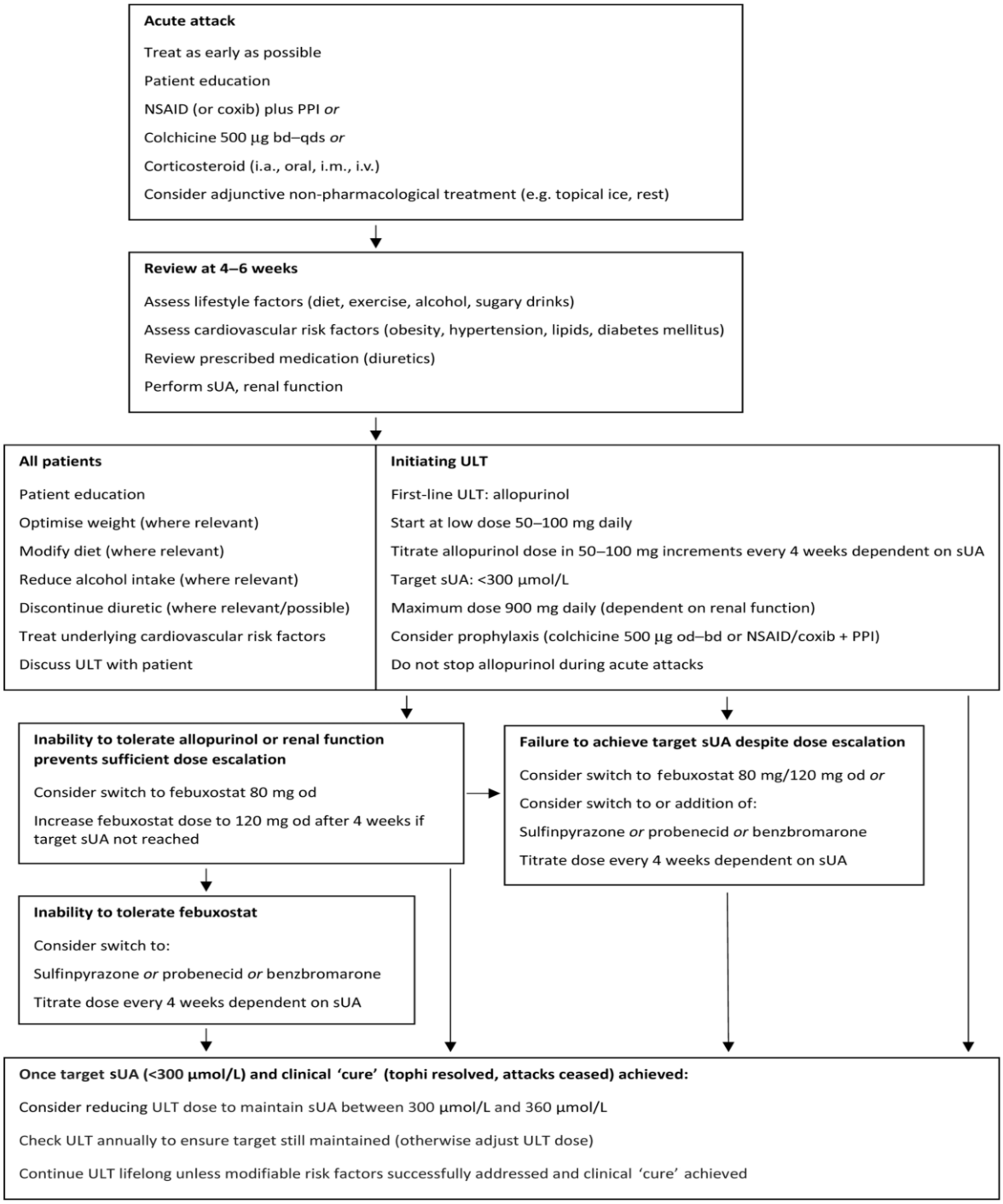
Table 2. Medication for acute and chronic gout

Agent	Class	Dosage (daily)	Route	Duration	Comments/Considerations[a]
Acute Flare Management					
Colchicine[62]	Anti-inflammatory	2 × 0.6 mg followed by 0.6 mg 1 h later	Oral	1 d*	Dose reductions required with P-gp and CYP3A4 inhibitors and with HIV protease inhibitors. *Wait 3 d before repeat dosing unless restarting colchicine for prophylaxis
NSAIDs					
Selective	Anti-inflammatory				Only celecoxib approved in the US; however, not FDA-approved for treatment of gout
Nonselective					
Naproxen sodium[3]		825 mg followed by 275 mg q8 h	Oral or suspension	5–8 d	Use with caution in elderly; increased risk of GI bleeds
Naproxen		750 mg followed by 250 mg q8 h	Oral or suspension	5–8 d	Use with caution in elderly; increased risk of GI bleeds
Indomethacin[3]		50–150 mg/d; consider 50 mg TID	IV, oral, or suspension	2–7 d	Immediate-release and rectal formulations are also FDA approved. 50 mg TID may provide more accurate dosing
Sulindac[3]		200 mg BID	Oral	7–10 d	
Corticosteroids					
Prednisone		40–60 mg once daily	Oral	3 d	After 3 d, taper dose by 10–15 mg q3d until discontinuation
Triamcinolone		60 mg/d	IM	Once	
Chronic Management					
Flare prophylaxis					
Colchicine[62]	Anti-inflammatory	0.6 mg once or twice daily; maximum dose 1.2 mg/d	Oral	6–12 m	Precede ULT initiation by 1–2 wks; dose reductions required with P-gp and CYP3A4 inhibitors and with HIV protease inhibitors
Urate-lowering therapy					
Allopurinol	Xanthine oxidase inhibitor	300 mg/d up to 800 mg/d, titrated to effect	Oral	Life-long	Dose reduction required for CKD based on creatinine clearance
Febuxostat[64]	Xanthine oxidase inhibitor	40–80 mg/d	Oral	Life-long	No renal or hepatic adjustments needed for mild to moderate hepatic/renal impairment
Probenecid	Uricosuric	250 mg BID	Oral	Life-long	Titrate dose to effect to maximum 2000 mg/d. Avoid use if CrCl < 50 mL/min
Pegloticase[63]	Uricase	8 mg q2 wk	IV	N/A[b]	Specific guidelines for renal and hepatic impairment dosage adjustments are not available

[a]Data on drug administration reviewed in Terkeltaub et al (2003),[38] except where noted.
[b]Optimal treatment duration not established.
Abbreviations: ACTH, adrenocorticotropic hormone; BID, twice daily; CKD, chronic kidney disease; CrCl, creatinine clearance; CYP, cytochrome P450; d, day; FDA, US Food and Drug Administration; GI, gastrointestinal; h, hour; HIV, human immunodeficiency virus; IM, intramuscular; IU, international units; IV, intravenous; m, month; NSAIDs, nonsteroidal anti-inflammatory drugs; P-gp, P-glycoprotein; q, every; TID, three times daily; ULT, urate-lowering therapy; wk, week.

Reprinted with permission from Taylor and Francis (138)

Figure 1. BSR algorithm for the management of gout



Note: coxib: cyclooxygenase-2 inhibitor; PPI: proton pump inhibitor; SUA: serum uric acid; ULT: urate-lowering therapy. Reprinted with permission from Oxford University Press (33)

Figure 2. EULAR recommendations for the management of gout

	Overarching principles
A	Every person with gout should be fully informed about the pathophysiology of the disease, the existence of effective treatments, associated comorbidities and the principles of managing acute attacks and eliminating urate crystals through lifelong lowering of SUA level below a target level.
B	Every person with gout should receive advice regarding lifestyle: weight loss if appropriate and avoidance of alcohol (especially beer and spirits) and sugar-sweetened drinks, heavy meals and excessive intake of meat and seafood. Low-fat dairy products should be encouraged. Regular exercise should be advised.
C	Every person with gout should be systematically screened for associated comorbidities and cardiovascular risk factors, including renal impairment, coronary heart disease, heart failure, stroke, peripheral arterial disease, obesity, hyperlipidaemia, hypertension, diabetes and smoking, which should be addressed as an integral part of the management of gout.
	Final set of 11 recommendations
1	Acute flares of gout should be treated as early as possible. Fully informed patients should be educated to self-medicate at the first warning symptoms. The choice of drug (s) should be based on the presence of contraindications, the patient's previous experience with treatments, time of initiation after flare onset and the number and type of joint(s) involved.
2	Recommended first-line options for acute flares are colchicine (within 12 hours of flare onset) at a loading dose of 1 mg followed 1 hour later by 0.5 mg on day 1 and/or an NSAID (plus proton pump inhibitors if appropriate), oral corticosteroid (30–35 mg/day of equivalent prednisolone for 3–5 days) or articular aspiration and injection of corticosteroids. Colchicine and NSAIDs should be avoided in patients with severe renal impairment. Colchicine should not be given to patients receiving strong P-glycoprotein and/or CYP3A4 inhibitors such as cyclosporin or clarithromycin.
3	In patients with frequent flares and contraindications to colchicine, NSAIDs and corticosteroid (oral and injectable), IL-1 blockers should be considered for treating flares. Current infection is a contraindication to the use of IL-1 blockers. ULT should be adjusted to achieve the uricaemia target following an IL-1 blocker treatment for flare.
4	Prophylaxis against flares should be fully explained and discussed with the patient. Prophylaxis is recommended during the first 6 months of ULT. Recommended prophylactic treatment is colchicine, 0.5–1 mg/day, a dose that should be reduced in patients with renal impairment. In cases of renal impairment or statin treatment, patients and physicians should be aware of potential neurotoxicity and/or muscular toxicity with prophylactic colchicine. Co-prescription of colchicine with strong P-glycoprotein and/or CYP3A4 inhibitors should be avoided. If colchicine is not tolerated or is contraindicated, prophylaxis with NSAIDs at low dosage, if not contraindicated, should be considered.
5	ULT should be considered and discussed with every patient with a definite diagnosis of gout from the first presentation. ULT is indicated in all patients with recurrent flares, tophi, urate arthropathy and/or renal stones. Initiation of ULT is recommended close to the time of first diagnosis in patients presenting at a young age (<40 years) or with a very high SUA level (>8.0 mg/dL; 480 μmol/L) and/or comorbidities (renal impairment, hypertension, ischaemic heart disease, heart failure). Patients with gout should receive full information and be fully involved in decision-making concerning the use of ULT.
6	For patients on ULT, SUA level should be monitored and maintained to <6 mg/dL (360 μmol/L). A lower SUA target (<5 mg/dL; 300 μmol/L) to facilitate faster dissolution of crystals is recommended for patients with severe gout (tophi, chronic arthropathy, frequent attacks) until total crystal dissolution and resolution of gout. SUA level <3 mg/dL is not recommended in the long term.
7	All ULTs should be started at a low dose and then titrated upwards until the SUA target is reached. SUA <6 mg/dL (360 μmol/L) should be maintained lifelong.
8	In patients with normal kidney function, allopurinol is recommended for first-line ULT, starting at a low dose (100 mg/day) and increasing by 100 mg increments every 2–4 weeks if required, to reach the uricaemia target. If the SUA target cannot be reached by an appropriate dose of allopurinol, allopurinol should be switched to febuxostat or a uricosuric or combined with a uricosuric. Febuxostat or a uricosuric are also indicated if allopurinol cannot be tolerated.
9	In patients with renal impairment, the allopurinol maximum dosage should be adjusted to creatinine clearance. If the SUA target cannot be achieved at this dose, the patient should be switched to febuxostat or given benzbromarone with or without allopurinol, except in patients with estimated glomerular filtration rate <30 mL/min.
10	In patients with crystal-proven, severe debilitating chronic tophaceous gout and poor quality of life, in whom the SUA target cannot be reached with any other available drug at the maximal dosage (including combinations), pegloticase is indicated.
11	When gout occurs in a patient receiving loop or thiazide diuretics, substitute the diuretic if possible; for hypertension consider losartan or calcium channel blockers; for hyperlipidaemia, consider a statin or fenofibrate.

IL, interleukin; NSAID, non-steroidal anti-inflammatory drug; SUA, serum uric acid; ULT, urate-lowering therapy.

Reprinted with permission from BMJ Publishing Group Ltd. (139)

Chronic gout can be considered 'curable' with ULT such as xanthine oxidase inhibitors, uricosurics or uricolytics (140). Allopurinol is a common first-line therapy and has a dose-dependent effect on SUA level (141). However, there is limited information on the optimum triggers for initiation and whether treatment can exacerbate an existing flare. Guidelines differ on whether to initiate therapy during, or 1-2 weeks following, a flare and whether early treatment is beneficial or whether to only commence treatment if there are ≥2 flares in a year, chronic kidney disease (CKD), tophi or history of urolithiasis (106, 122, 129, 142) (143). Initiation upon diagnosis, with co-prescription of uricosuric colchicine (0.5-1 mg daily for six months) or NSAIDs for six weeks for prophylaxis of flares, and continuation during flares, is increasingly recommended, in part to aid medication compliance (60, 110, 122, 129, 134, 135). Allopurinol is commonly initiated at 50-100 mg/day and titrated by 50-100 mg every 2-6 weeks until

the SUA level is maintained below <6 mg/dL (122, 139). Upon reaching target SUA level, continued SUA measurements and medication review (e.g. every six months) may limit non-adherence, maintain quality of life and prevent recurrent flare (110, 144).

The varied recommendations on dosage and the timing of initiation are partly due to the risks of toxicity and of flare that is brought on by sudden change in SUA level (129). In one rheumatology clinic, allopurinol intolerance (e.g. hypersensitivity, rash) was reported among 13.7% of patients (145). However, a recent study showed no correlation between allopurinol and mortality, and probable beneficial effects on hypertension have been noted (146). The HLA-B*5801 allele may affect treatment response and screening for this is recommended for Han Chinese and Thai patients prior to commencing therapy (110). Historically, allopurinol doses above 300 mg/day were avoided but <50% reach target SUA level on this and doses up to 900 mg/day are tolerated in patients with normal renal function (147, 148). Alternative or combination therapies include febuxostat, probenecid and benzbromarone (122).

However, the therapeutic management of gout is suboptimal despite the existence of international guidelines and quality indicators (33, 106, 122, 134, 139, 142). The potential of NSAID and colchicine toxicity may contribute to under-prescribing of flare prophylactics when allopurinol therapy is initiated (110, 140). In one study, over 70% had gaps in ULT, 75% of these within the first year of treatment (149). A GP study found that only 34% of patients had >1 ULT prescription, 20% initiating therapy on the date of diagnosis (44). The ULT dosage is the main modifiable factor in reaching target SUA level, yet only around a quarter of patients receive effective treatment and 80% experience further flares within three years of diagnosis (109, 116, 141, 150, 151). SUA level testing is key for therapy adjustment, yet studies report infrequent testing, e.g. one study found only 8% of patients on allopurinol had their SUA level tested, with under-monitoring being reported even among patients with flares (60, 116). Further, a GP-based study reported that only 58% of patients received ULT following a high SUA level result (152). In studies of patients managed in secondary rather than primary care, ULT was more frequently prescribed, and target SUA level more commonly reached, although management remained sub-optimal (145, 153). Alongside under-prescribing, 25-50% of patients may receive inappropriate prescribing (150). In a hospital study, gout was reported as one of the most common diseases associated with inappropriate medications among the elderly (142). Appropriate monitoring is the key to optimal prescribing, to achieve flare prophylaxis and to prevent ULT-related toxicity, but also given the converse risk of neurological issues (e.g. dementia) if SUA levels fall too low (34).

Patient education on the pathogenesis and management of gout may facilitate medication adherence and improve quality of life (129, 144). In a study involving a rheumatology-led education and nurse-led follow-up, 92% of patients achieved the target SUA level (154). Medication adherence may be improved through the provision of information on ULT around the time of first diagnosis with gout, and through patient access to clinical notes (152, 155). Importantly, mismanagement of gout was halved under a programme of patient self-management (129). Other patient factors including comorbidity may affect adherence (156). Organisational practice and clinician factors are also important: younger patients less commonly receive ULT and younger clinicians may be more likely to prescribe ULT (150, 157). Patient and GP practice demographics have been shown to associate with 7.82% and 13.49% of allopurinol prescription variance respectively (152).

There is also an important role for non-pharmacologic modalities of management (129). Reducing the intake of alcohol, fructose, shellfish and red meat, and avoidance of Atkins-type diets, can help to prevent hyperuricaemia (122). In obese patients, weight management can be beneficial in reducing flares and preventing life-threatening comorbidities (106). Medication reviews may highlight instances where urate-raising medication can be switched or is non-essential in the management of comorbidities (106). However, the long-term effect of lifestyle change on disease severity and burden is uncertain (106, 129).

2.3.1.6 EHR-based Research

EHR-based studies highlight a rising incidence and prevalence and persistent suboptimal gout management in terms of SUA level monitoring and the initiation, adherence and titration of ULT (116, 117, 137, 150, 151, 158, 159).

In EHR-based research, definitions of gout and flares have been based on diagnosis codes, symptoms including tophi, joint pain and erythema, ULT or NSAID prescription and joint aspiration (117). Gout diagnostic codes have been validated, with one study reporting that clinical review confirmed 83% of recorded diagnoses, and studies that performed sensitivity analyses that required ULT or NSAIDs in addition to diagnostic codes, have shown comparable results between primary and sensitivity analyses (140, 160, 161). A study showed that coded recording of podagra and tophi had high diagnostic specificity also (134), suggesting the suitability of these codes in defining gout. Patient exclusions in EHR-based studies of gout have been based on follow-up

duration or recorded diagnoses that are possible indications or contra-indications for ULT or for other rheumatologic conditions that may indicate misclassification (150, 159, 162).

The management of gout has been investigated in EHR-based literature, which report suboptimal prescribing, monitoring and comorbidity screening. Allopurinol has been reported as the most common prescribed ULT and the timing of treatment initiation in relation to diagnosis has been studied, along with the patient characteristics associating with the likelihood of receiving therapy (117, 152, 157). Importantly, contrary to some guideline recommendations, one study reported fewer subsequent flares and GP visits in patients prescribed allopurinol at diagnosis (117). Some studies have examined dosage, e.g. Cottrell et al. reported that 76% of patients remain on their starting dose of allopurinol despite 67% having high SUA levels, and that a third of patients on 300 mg/day had high SUA levels (60, 145). Care quality indicators have been readily assessed using EHR data, e.g. one study found that ULT was only prescribed for half of the patients eligible in the year post-diagnosis (150, 152). Comorbidity screening in patients with gout has also been assessed, with 5% and 26% having lipid and blood pressure monitoring respectively, for hyperlipidaemia and hypertension (157). Similarly, medication-adjustment has been investigated: diuretic prescribing continued after gout diagnosis in 64% of cases (140). Studies of trends in medication use have suggested declining NSAID prescribing, potentially following concerns of toxicity (126). The EHR-based research on gout pharmacologic management will be systematically reviewed in Chapter 4.

Limitations highlighted in EHR-based studies included the possibility of missing patients with unrecognised gout or misdiagnosis, including less severe cases that do not present to clinical care (126, 150). In the UK, NSAIDs have been available over-the-counter since 1983 and such consumption is missed from prescribing estimates; this may also introduce bias as may be more prevalent among younger, pre-retirement patients, who would pay prescription costs in England (126). Estimates of the timing and treatment of flares is hindered in EHR data, given the practice of prescribing treatment for the patient to use at a later date in flare self-medication as required (117). Similarly, patients prescribed allopurinol during a flare may be advised to commence therapy once the flare subsides. Such practice would hinder comparison of outcomes in patients following the contrasting guidance on the timing of allopurinol initiation. Some EHR research databases lack complete information on dosage and laboratory tests, which limits investigations of titration patterns and SUA level (150). There was also variation in the completeness of reporting of EHR-based studies, e.g. Kuo et al.

reported that 24.8% of patients had no recorded alcohol intake but this was not reported in all of the relevant studies (61).

2.3.2 Ankylosing Spondylitis

2.3.2.1 Clinical Features

Ankylosing spondylitis is an inflammatory autoimmune disorder that affects the axial skeleton and causes gradual fusing of vertebral and sacroiliac joints. It is a type of spondyloarthropathy (an arthritis involving both spine and peripheral joints) characterised by back pain with structural and functional impairment due to osteo-proliferation, with tissue and cartilage being replaced by bone (163). It is sometimes, especially in women, accompanied by peripheral arthritis in the lower limbs, enthesitis, restricted chest expansion and osteoporosis (164, 165). Extra-articular manifestations are common, including anterior uveitis, skin psoriasis and inflammatory bowel disease (IBD) (163).

However, the aetiology and pathogenesis of AS are poorly defined and the burden and management of AS is poorly understood. There is a genetic and immune-mediated basis, particularly involving HLA-B27, and an infectious pathogenesis related to enteric infection is posited, with a role played by gut bacteria in stimulating the immune system, (166, 167). The prevalence of AS correlates with the HLA-B27 positive rate in populations (168). Structural damage at presentation may be the best predictor of prognosis and much radiographic progression occurs in the first decade of disease (169, 170). The timely diagnosis of AS is therefore important.

Ankylosing spondylitis severely hinders quality of life and productivity, as the disease progresses in early adulthood (171, 172). Inflammation and fusion of the axial skeleton have far-reaching consequences both from spinal rigidity and systemic inflammation, with one study reporting disability among 94.2% of AS patients (173). Functional disability and pain are highly prevalent, with considerable effect on quality of life and impacting directly on work presenteeism and absenteeism (174). Disease onset occurs during early working life, around 25 years of age, and 21% of patients have been reported to leave the labour force within 10 years of AS diagnosis (30, 35). The chronic systemic inflammation in AS may contribute to the risk of serious comorbidities, including myocardial infarction and stroke (175). Mortality is higher in patients with AS (adjusted hazard ratio [AHR]: 1.60, 95% confidence interval [CI] 1.44-1.77), with

significant predictors being markers of high disease activity, such as prior hip replacement, work disability, and greater number of inflamed peripheral joints (176, 177).

Encouragingly, significant progress in AS treatment has been made in the last decade; effective management may improve and maintain spinal flexibility and reduce complications (e.g. adalimumab treatment may associate with improved work productivity) (168, 174). While delayed diagnosis associates with higher disease activity and worse quality of life, timely diagnosis and early treatment initiation substantially increases the likelihood of remission (178, 179).

2.3.2.2 Epidemiology

The epidemiology of AS is poorly studied, with few longitudinal studies. The prevalence of AS is mostly reported between 0.09 and 0.6%, generally seeming lower in more recent studies (180, 181). Varied incidence trends are reported. The unadjusted incidence rate rose between 1960 and 1993 in Northern Norway (182), while the age- and sex- adjusted incidence rate of AS per 10,000 person-years tended to decline in Minnesota, USA from 0.85 (95% CI 0.55-1.16) between 1935 and 1949 to 0.52 (95% CI 0.34-0.70) in 1980-1989, yet there was no significant change in a second study in Minnesota between 1980 and 2009 (183, 184). Incidence was recently recorded as 0.79 per 10,000 person years in South Korea in 2015 (185). Population and ethnic differences in HLA-B27 prevalence, and immigration, may contribute to variance across studies and over time. Prevalence of HLA-B27 and AS are lower in many black communities (186, 187). Men are more commonly diagnosed than women are (e.g. 77% of survey respondents with AS were men in a study in Wales by Cooksey et al.) though the reported difference is diminishing following greater recognition of symptoms and the publication of classification criteria in the mid-2000s (188, 189). A national review highlighted the absence of evidence regarding AS incidence and prevalence in the UK and made a key recommendation for such research (190).

The occurrence and severity of AS are influenced by genetics, with evidence for a role of intestinal inflammation in the pathogenesis of AS (191). HLA-B27 determines ~20% of the genetic risk and over 100 other contributing alleles have been identified (192, 193). Twin studies have shown that HLA-B27 affects the gut microbiome, which impacts on the intestinal microbiome composition (191). Factors including increased gut permeability in AS patients – up to 70% having clinical or subclinical gut disease – and potentially the role of HLA-B27 in the capability to present appropriate microbial

antigens, contribute to the hypothesis that gut disease and ‘leakiness’ of the gut drive the inflammation and development of AS (194-197). Evidence for a role of sex hormones in the aetiology of AS is conflicting, though there may be sex-specific differences in immune profiles (198, 199). The Th17 axis of the immune system is implicated in AS (200) and has been observed to undergo up-regulation in male but not female patients with AS compared with controls (199).

In addition, there may be sex differences in the expression of AS and in treatment response. While men are three times more likely to be diagnosed with AS (201), this may in part be due to AS diagnostic criteria traditionally focusing on the aspects of disease expressed more commonly among men, such as radiographic sacroiliitis and development in early adulthood (202). Women may have more peripheral arthritis and less radiographic spinal damage (203). While a later disease onset among women is generally reported, there are conflicting studies and the age at diagnosis may have changed following efforts to increase the recognition of AS and to prompt early diagnosis (203-205). The progression of AS may be worse in women than in men, as well as in patients who smoke or are obese (203, 206, 207). Further, the biologics developed for AS are less effective for women, who have less improvement and consequently carry a higher disease burden (201).

2.3.2.3 Common Comorbidities

Extra-articular inflammatory manifestations (EAMs) and peripheral musculoskeletal manifestations are common, systemic features of AS. Around half of AS patients at some point develop arthritis in peripheral entheses or joints, with the current prevalence of enthesitis and dactylitis being 13.6% (CI 1.8-31.8) and 5.6% (CI 0.0-16.2) at any given time (208). The organs affected by EAMs include the eyes, skin and bowels (163). Anterior uveitis causes blurred vision and pain and affects approximately one-third of AS patients, often repeatedly, and occasionally in both eyes simultaneously (209). Psoriasis and IBD may affect 10.2% (CI 7.5-13.2) and 4.1% (CI 2.3-6.5) of patients with AS respectively, compared with 2.8% (±0.02) and 0.4% (±0.04) of the general population (208, 210, 211). HLA-B27 positive patients may have more likelihood of developing anterior uveitis and lower occurrence of psoriasis and IBD (212). The prevalence of EAMs varies between ethnic populations and they are more common in patients with peripheral arthritis (which more commonly affects women) (168, 213, 214). In the management of AS, EAM treatment and prophylaxis is important, and guidelines recommend that EAMs inform the selection of anti-tumour necrosis factor (TNF) inhibitors for treating AS (31).

In addition to these inflammatory diseases, the rate of osteoporosis and CVD is also high, with serious implications. Osteoporosis affects 25% of AS patients and in combination with spinal rigidity this leads to high spinal fracture rates (10%) with associated risk of spinal cord injury (215). It is important to scan for fractures wherever new neck or back pain is presented in AS patients, to prevent neurological complications. Bone-protectant prophylaxis may be a suitable consideration alongside corticosteroid therapy in AS management (216). Hypertensive and ischaemic heart diseases are twice more common among AS populations than the general population (standard mortality rate ratio 1.98 [95% CI 1.72-2.28] and 2.20 [95% CI 1.77-2.70]. The systemic inflammation common in AS is a contributing factor, though therapy for AS may also affect CVD risk, as NSAIDs raise the risk while TNF inhibitors may be protective (217-220). Cardiovascular disease contributes to healthcare utilisation costs in AS patients and is the most frequent cause of death in AS patients. Overall mortality has been reported to be 1.5 times higher than for the general population although in one study this has not been found among women (176, 221). The early age of onset in AS compared with other joint diseases mean that early diagnosis and therapy is important for reducing the chronic burden of inflammation, alongside vigilance to osteoporosis and CVD.

2.3.2.4 Diagnosis

Symptoms of inflammatory back pain (IBP), including alternating buttock pain and back stiffness that is worse in the morning (lasting >30 minutes) and better for exercise but not rest, and AS-associated features such as anterior uveitis, are used in diagnosing AS (222). In the UK, national guidelines recommend that AS is suspected wherever patients present with “chronic or recurrent low back pain, fatigue, and stiffness” (223). A range of indicators may point to AS, including the patient being aged ≤45 with a history of peripheral arthritis, anterior uveitis, IBD or psoriasis and back stiffness symptoms that respond well to NSAIDs within 48 hours. Once AS is suspected, referral to a rheumatologist is recommended. Magnetic resonance imaging (MRI) is an important early diagnostic tool, detecting inflammatory lesions before sacroiliitis becomes radiographically detectable, and other indicators are HLA-B27 and elevated C-reactive protein (CRP) level (224). Both erythrocyte sedimentation rate (ESR) and CRP can indicate acute phase response and associate with poor radiographic progression (225). Classification criteria may aid in diagnosis (Table 3), however modern guidelines highlight that symptoms may be diverse and that AS cannot be discounted based on the presence or absence of a symptom or test result (223).

Table 3 Modified New York criteria for the classification of AS

1 Radiological criterion
Bilateral sacroiliitis grade ≥ II or unilateral sacroiliitis grade III to IV
2 Clinical criteria
(a) Low back pain and stiffness of at least 3 months duration improved by exercise and not relieved by rest
(b) Limitation of motion of the lumbar spine in both the sagittal and the frontal planes
(c) Limitation of chest expansion relative to values normal for age and sex
Definite AS is diagnosed if the radiological criterion plus 2 of the 3 clinical criteria are present.

Note: AS = ankylosing spondylitis. Reprinted with permission from BMJ Publishing Group Ltd. (226)

Despite the importance of early diagnosis for limiting the burden of AS (179), studies have highlighted a considerable delay between the presentation of non-specific back pain symptoms and AS diagnosis, particularly in women. The commonality of chronic back pain (227) and the complexity of AS, with diverse symptoms, may contribute to this. There may be a period of disease activity before patients present with symptoms; one study reported that almost half of patients may wait more than a year (228). In addition, the delay between symptom presentation and diagnosis ('diagnostic delay') may be substantial. Feldtkeller first highlighted such a diagnostic delay in the diagnosis of spondyloarthropies in 1999 in Germany: a median delay of 9.8 and 8.4 years in women and men respectively (229). An 8.5-year (interquartile range [IQR] = 3.0-16.0) median diagnostic delay was reported in the UK between 2011-2016 (230). Historic focus in diagnostic criteria on spinal damage may have contributed, given that such radiographic features may present late in the disease course and more commonly in men (231).

2.3.2.5 Management

The goal of modern management is to maximise the quality of life, and early diagnosis and treatment initiation is crucial for the control of inflammation and prevention of progressive structural damage. The recommendations of the ASsessment in Ankylosing Spondylitis working group (ASAS) and EULAR task force emphasise the tailoring of treatment according to the manifestations displayed, the severity of symptoms, and the wishes and general clinical status (e.g. age, comorbidity) of the patient (Table 4) (232). Ongoing disease monitoring should consider patient history,

clinical features, testing and imaging, with frequency determined based on symptoms, severity and medication.

Table 4. Summary of ASAS/EULAR recommendations for the management of AS

1. General treatment

The treatment of patients with AS should be tailored according to:

- The current manifestations of the disease (axial, peripheral, entheseal, extra-articular symptoms and signs).
- The level of current symptoms, clinical findings, and prognostic indicators.
- The general clinical status (age, gender, comorbidity, concomitant medications, psychosocial factors).

2. Disease monitoring

The disease monitoring of patients with AS should include:

- Patient history (eg, questionnaires)
- Clinical parameters
- Laboratory tests
- Imaging
  - All according to the clinical presentation as well as the ASAS core set

The frequency of monitoring should be decided on an individual basis depending on:

- Course of symptoms
- Severity
- Treatment

3. Non-pharmacological treatment

- The cornerstone of non-pharmacological treatment of patients with AS is patient education and regular exercise.
- Home exercises are effective. Physical therapy with supervised exercises, land or water based, individually or in a group, should be preferred as these are more effective than home exercises.
- Patient associations and self-help groups may be useful.

4. Extra-articular manifestations and comorbidities

- The frequently observed extra-articular manifestations, for example, psoriasis, uveitis and IBD, should be managed in collaboration with the respective specialists.
- Rheumatologists should be aware of the increased risk of cardiovascular disease and osteoporosis.

5. Non-steroidal anti-inflammatory drugs

- NSAID, including Coxibs, are recommended as first-line drug treatment for AS patients with pain and stiffness.
- Continuous treatment with NSAID is preferred for patients with persistently active, symptomatic disease.
- Cardiovascular, gastrointestinal and renal risks should be taken into account when prescribing NSAID.

6. Analgesics

- Analgesics, such as paracetamol and opioid (like) drugs, might be considered for residual pain after previously recommended treatments have failed, are contraindicated, and/or poorly tolerated.

7. Glucocorticoids

- Corticosteroid injections directed to the local site of musculoskeletal inflammation may be considered.
- The use of systemic glucocorticoids for axial disease is not supported by evidence.

8. Disease-modifying antirheumatic drugs

- There is no evidence for the efficacy of DMARD, including sulfasalazine and methotrexate, for the treatment of axial disease.
- Sulfasalazine may be considered in patients with peripheral arthritis.

9. Anti-TNF therapy

- Anti-TNF therapy should be given to patients with persistently high disease activity despite conventional treatments according to the ASAS recommendations.
- There is no evidence to support the obligatory use of DMARD before or concomitant with anti-TNF therapy in patients with axial disease.
- There is no evidence to support a difference in efficacy of the various TNF inhibitors on the axial and articular/entheseal disease manifestations; but in the presence of IBD a difference in gastrointestinal efficacy needs to be taken into account.
- Switching to a second TNF blocker might be beneficial especially in patients with loss of response.
- There is no evidence to support the use of biological agents other than TNF inhibitors in AS.

10. Surgery

- Total hip arthroplasty should be considered in patients with refractory pain or disability and radiographic evidence of structural damage, independent of age.
- Spinal corrective osteotomy may be considered in patients with severe disabling deformity.
- In patients with AS and an acute vertebral fracture a spinal surgeon should be consulted.

11. Changes in the disease course

- If a significant change in the course of the disease occurs, other causes than inflammation, such as a spinal fracture, should be considered and appropriate evaluation, including imaging, should be performed.

Reprinted with permission from BMJ Publishing Group Ltd. (232)

Both non-pharmacologic intervention and pharmacotherapy are employed in treating AS, with multi-disciplinary management often co-ordinated by a rheumatologist. Regular, supervised exercise on land and in water and involvement in self-help groups can help to maintain function and social participation. Guided exercise can have beneficial effects on outcomes, including balance, pain, disease activity and depression (233-235). A combined exercise programme that incorporates range of motion, strengthening and aerobic exercises, can bring particular benefit to this multi-modal disease (236).

In pharmacotherapy, NSAIDs are an important initial therapy for treating pain and stiffness, and may be prescribed long-term for persistently active, symptomatic cases (232). However, continuous therapy may prevent new bone formation, especially at large doses, and the cardiovascular, renal and gastrointestinal risks of NSAIDs should be considered given the potentially toxic nature of these drugs (217, 237-239). More recently, anti-TNF agents have offered improvement for cases of persistent high disease activity where NSAIDs are ineffective or not tolerated, and they may particularly be effective early in the disease course (178, 240). Guidelines for anti-TNF prescribing were published by the BSR in 2005 (240). Where this fails, newer medicines including interleukin-17 inhibitors and janus kinase inhibitors have shown good efficacy (31, 241). Other biologics are not generally considered effective at treating axial disease, although further research is required into the effectiveness of sulfasalazine, methotrexate and leflunomide for treating the earlier disease stages of peripheral arthritis in AS (242-247). If remission is reached and sustained, then the dosage of biologics may be tapered (31).

Corticosteroid injections may also help to reduce localised musculoskeletal inflammation, although systemic use should be avoided in patients with axial disease (232). Analgesics may provide symptomatic relief for residual pain or if recommended treatments fail (232). In cases of persistent pain and disability, or where there is evidence of advanced hip involvement or spinal instability, hip arthroplasty or spinal surgery may be necessary (248).

2.3.2.6 EHR-based Research

The EHR-based research in AS has shown a substantial associated health burden, which is contributed to by diagnostic delay. Compared with gout, AS has been less widely researched in EHR-based studies. Nonetheless, a high validity of AS diagnoses in GP EHR data has been reported and diagnostic coding and hospital prescription data have been used to investigate comorbidities, EAMs and the pharmacologic management of AS (62). Incidence and prevalence were relatively under-explored, with higher rates reported in primary care than in secondary care (182, 249). Prevalence estimates included 0.11% in California and 0.02% in Turkey (250, 251). Significant delays in diagnosis have been reported in hospital studies, contributing to poorer quality of life and higher mortality rates (176). One study defined a mean 13.4 year delay between disease onset and biologics initiation (252). The long-term effectiveness of biologics has been investigated using hospital and medical office data, with significant reduction in disease activity being reported (252). This EHR-based AS literature will be systematically reviewed in Chapter 5.

2.3.3 Rheumatoid Arthritis

2.3.3.1 Clinical Features

Rheumatoid arthritis is a chronic inflammatory autoimmune disorder that causes progressive symmetric polyarthritis of synovial-lined large and small joints, with irreversible joint damage (253). It is characterised by synovial inflammation and swelling, auto-antibody production and extra-articular features (254). RA commonly affects joints, tendons and bursa in the hands, wrists, knees and feet and causes painful joint swelling, extended morning stiffness (≥1 hour) and impaired function (26). The aetiology of RA is poorly defined, with evidence of environmental and genetic involvement and risk factors including cigarette smoking and female sex. The onset of RA may be acute or subacute, and may present initially as monoarticular disease or extra-articular tendinitis or bursitis. Symptoms often first present in the hands and feet, and joint symptoms are accompanied by systemic features such as fatigue and weight loss; there may also be extra-articular involvement (26). Extra-articular manifestations may involve the skin, cardiovascular, pulmonary, ocular and nervous systems, and include vasculitis, pleural effusions, scleritis and rheumatoid nodules. Prognosis is unpredictable and the clinical course of RA can fluctuate, although early treatment can be effective in altering the disease course (255).

Rheumatoid arthritis can lead to progressive disability, work loss, increased mortality and socioeconomic costs (256). Loss of hand function and the failure of larger weight-bearing joints, following cartilage damage and bone erosion, is a common burden with far-reaching consequences for quality of life. Over 50% of RA patients are of working age at symptom onset and in one study, from the pre-biologic era, ~80% of working age RA patients reported disabling pain and reduced functionality (257, 258). The consequences of systemic inflammation in untreated RA include increased rates of CVD, fatigue, sarcopenia and lymphoma (256). Mortality is 40-50% higher in patients with RA compared to the general population and mortality risk is higher in RA patients with high rather than low disease activity (AHR = 2.43, 95% CI 1.64-3.61) (259, 260). Encouragingly, Listing et al. reported lower mortality in RA patients that had effective treatment with the biologic disease-modifying anti-rheumatic drug (DMARD) rituximab (AHR = 0.57, 95% CI 0.39-0.84). Effective pharmacologic disease control is important in curbing the burden of RA.

2.3.3.2 Epidemiology

Rheumatoid arthritis affects ~0.5-1% of the UK population, with disease onset peaking at age 50-60; incidence is 50-75% higher in women than in men, although this disparity diminishes with age (261, 262). The age of disease onset is rising, although the reasons for this are unclear (263). There is high geographic variation reported in prevalence: 0.1-0.3% in Asian countries, 0.3-0.7% in Southern European countries and 0.5-1.1% in Northern European and North American countries (262). Some studies show comparatively different incidence rates (e.g. 0.15 and 0.39 per 1000 person years in two studies of incidence in the UK in 1996) and report a decline in incidence (e.g. at a rate of 1.6% per year across 1990-2014 in the UK) (264, 265). Silman et al. found no evidence of time or space clustering in RA incidence in East Anglia (266). However, methodological differences in epidemiologic studies of RA, including variation in diagnostic case definitions of RA, as well as regional or temporal differences in coding practice, health care access and diagnosis, may restrict the comparison and interpretation of geographic and temporal trends. In the UK, changes in the Quality and Outcomes Framework (QOF) indicators in 2012 introduced payment, annually from April 2013, for the recording of RA diagnosis using specific codes (20). This may have affected the practice of diagnostic coding in RA in the UK.

Rheumatoid arthritis is a multifactorial disease, with its occurrence, severity and expression being influenced by interactions between genetic and environmental factors. The nature and impact of risk factors are not fully understood, although

reported risk factors include genetic predisposition, sex and age, smoking, infectious agents, ethnicity, diet and hormonal factors. Twin and family studies have indicated an association of diagnosis, severity and outcome with HLA alleles encoding a "rheumatoid epitope", which may partly explain population- or ethnic- variations in prevalence (267, 268). In a UK GP EHR-based study, Rodriguez et al. found no association with alcohol use, obesity, or prescribing of hormone replacement therapy or low-dose aspirin (264). Rheumatoid arthritis is more common in women, and pregnancy may associate with remission, although the role of hormonal or reproductive factors is uncertain (262). A meta-analysis by Sugiyama et al. reported smoking, especially 20 or more pack-years among men, as a risk factor (269). There are conflicting associations reported with infections, such as Epstein-Barr virus and parvovirus, and while this is complicated by infections that mimic autoimmune disease, it may is suspected that infection can have a role in the triggering and severity of RA (270, 271). There are conflicting reports of a protective effect from a Mediterranean diet or fish oils, which may further indicate the complex interplay between genetic and environmental factors (262, 272).

2.3.3.3 Common Comorbidities

A number of comorbidities are common with RA, including osteoporosis, CVD and depression. The inflammation in RA can, by activating the cytokine pathway, trigger osteoporosis (273). Prolonged systemic inflammation from RA can contribute to arterial disease and increased cardiovascular risk. The prevalence of CVDs including hypertension, cerebrovascular disease and ischaemic heart disease are higher, and the rate of controlled hypertension is lower, than in non-RA patients (274, 275). In a study of patients without prior cardiovascular disease, RA patients were more likely to develop CVD than non-RA patients were (AHR = 2.06; 95% confidence interval 1.34-3.16) (276). Risk of cardiovascular mortality is also doubled in RA patients, with corticosteroid therapy contributing further to this risk (36, 277). Cardiovascular risk prediction tools can assist in determining instances where prophylactics such as statins, aspirin and platelet inhibitors are advised (278). Higher incidence of depression has been reported in RA than non-RA patients (15.69 vs. 8.95 per 1,000 person-years) (279). In the UK, national guidelines recommend annually assessing for the development of comorbidities in RA patients (280).

Comorbidities may also affect the likelihood of disease remission. For example, depression associates with elevated metrics of tender joint count and joint pain (281). Obesity worsens the disease course and impairs treatment, including through the pro-inflammatory action of white adipose tissue (282-284). The impact of comorbidities and

obesity is concerning, as the prevalence of RA is highest in the over 65s and in the UK over two-thirds of people in this age-group are multi-morbid (having ≥2 long-term conditions), and the prevalence of obesity has tripled globally between 1975 and 2016 (285-287). The management of comorbidities is therefore an important factor in treating RA.

2.3.3.4 Diagnosis

In the UK, guidelines recommend referral for specialist investigation if a patient presents with prolonged morning joint stiffness and persistent soft tissue joint swelling (synovitis) (280). Investigations for diagnosis can include clinical examination of synovitis, a blood test for rheumatoid factor, CRP, ESR, anti-nuclear antibody and anti-cyclic citrullinated peptide (anti-CCP) antibodies, and X-rays of the hands and feet to establish whether erosions are present (Table 5). Rheumatoid factors are proteins produced as an immune response that may attack tissue. The best indicators of an acute phase response are CRP and ESR (26). Anti-CCP antibodies and erosions may indicate increased risk of radiological progression. Ultrasonography can reveal any subclinical tendon and bursal involvement, for early clinical assessment of the disease extent. Following diagnosis, if these are not already recorded, a measure of anti-CCP antibodies, x-ray of the hands and feet, and a measure of functional ability should be taken to establish a baseline for assessing the subsequent response to treatment (280).

Table 5. ACR / EULAR classification criteria for RA

	Score
Target population (Who should be tested?): Patients who	
1) have at least 1 joint with definite clinical synovitis (swelling)*	
2) with the synovitis not better explained by another disease†	
Classification criteria for RA (score-based algorithm: add score of categories A–D; a score of ≥6/10 is needed for classification of a patient as having definite RA)‡	
A. Joint involvement§	
1 large joint¶	0
2–10 large joints	1
1–3 small joints (with or without involvement of large joints)#	2
4–10 small joints (with or without involvement of large joints)	3
>10 joints (at least 1 small joint)**	5
B. Serology (at least 1 test result is needed for classification)††	
Negative RF *and* negative ACPA	0
Low-positive RF *or* low-positive ACPA	2
High-positive RF *or* high-positive ACPA	3
C. Acute-phase reactants (at least 1 test result is needed for classification)‡‡	
Normal CRP *and* normal ESR	0
Abnormal CRP *or* abnormal ESR	1
D. Duration of symptoms§§	
<6 weeks	0
≥6 weeks	1

* The criteria are aimed at classification of newly presenting patients. In addition, patients with erosive disease typical of rheumatoid arthritis (RA) with a history compatible with prior fulfillment of the 2010 criteria should be classified as having RA. Patients with longstanding disease, including those whose disease is inactive (with or without treatment) who, based on retrospectively available data, have previously fulfilled the 2010 criteria should be classified as having RA.

† Differential diagnoses vary among patients with different presentations, but may include conditions such as systemic lupus erythematosus, psoriatic arthritis, and gout. If it is unclear about the relevant differential diagnoses to consider, an expert rheumatologist should be consulted.

‡ Although patients with a score of <6/10 are not classifiable as having RA, their status can be reassessed and the criteria might be fulfilled cumulatively over time.

§ Joint involvement refers to any *swollen* or *tender* joint on examination, which may be confirmed by imaging evidence of synovitis. Distal interphalangeal joints, first carpometacarpal joints, and first metatarsophalangeal joints are *excluded from assessment*. Categories of joint distribution are classified according to the location and number of involved joints, with placement into the highest category possible based on the pattern of joint involvement.

¶ "Large joints" refers to shoulders, elbows, hips, knees, and ankles.

"Small joints" refers to the metacarpophalangeal joints, proximal interphalangeal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints, and wrists.

** In this category, at least 1 of the involved joints must be a small joint; the other joints can include any combination of large and additional small joints, as well as other joints not specifically listed elsewhere (e.g., temporomandibular, acromioclavicular, sternoclavicular, etc.).

†† Negative refers to IU values that are less than or equal to the upper limit of normal (ULN) for the laboratory and assay; low-positive refers to IU values that are higher than the ULN but ≤3 times the ULN for the laboratory and assay; high-positive refers to IU values that are >3 times the ULN for the laboratory and assay. Where rheumatoid factor (RF) information is only available as positive or negative, a positive result should be scored as low-positive for RF. ACPA = anti–citrullinated protein antibody.

‡‡ Normal/abnormal is determined by local laboratory standards. CRP = C-reactive protein; ESR = erythrocyte sedimentation rate.

§§ Duration of symptoms refers to patient self-report of the duration of signs or symptoms of synovitis (e.g., pain, swelling, tenderness) of joints that are clinically involved at the time of assessment, regardless of treatment status.

Reprinted with permission from John Wiley and Sons (288)

2.3.3.5 Management

Early diagnosis and treatment initiation is important to limit or reduce joint damage and disability in RA (289, 290). Modern management employs DMARDs, with corticosteroids and NSAIDs offering short-term symptomatic relief (Table 6, Table 7). DMARDs disrupt the mediators of inflammation and provide disease modification. In the UK, national guidelines emphasise the importance of monitoring CRP and disease activity, maintaining tight control of inflammation, treating active cases to a target of remission (or low disease activity) and responding rapidly if symptoms worsen or flare (280). This is termed a “treat-to-target” strategy to reduce disease activity and even lead to remission (291). DMARDs can be “non-biologic” (“conventional”) or “biologic” and modify the disease by reducing the immune response. Non-biologic DMARD monotherapy is recommended upon diagnosis, transferring to combination therapy and biologic DMARDs where there is inadequate response (280). Encouragingly, treatment success has improved following the rising profile of DMARDs and then the development of biologic DMARDs from 1998 onwards (290, 292).

Table 6. EULAR recommendations for the management of RA

	Recommendations
1.	Therapy with DMARDs should be started as soon as the diagnosis of RA is made
2.	Treatment should be aimed at reaching a target of sustained remission or low disease activity in every patient
3.	Monitoring should be frequent in active disease (every 1–3 months); if there is no improvement by at most 3 months after the start of treatment or the target has not been reached by 6 months, therapy should be adjusted
4.	MTX should be part of the first treatment strategy
5.	In patients with a contraindication to MTX (or early intolerance), leflunomide or sulfasalazine should be considered as part of the (first) treatment strategy
6.	Short-term glucocorticoids should be considered when initiating or changing csDMARDs, in different dose regimens and routes of administration, but should be tapered as rapidly as clinically feasible
7.	If the treatment target is not achieved with the first csDMARD strategy, in the absence of poor prognostic factors, other csDMARDs should be considered
8.	If the treatment target is not achieved with the first csDMARD strategy, when poor prognostic factors are present, addition of a bDMARD*[1,2] or a tsDMARD*[3] should be considered; current practice would be to start a bDMARD[§]
9.	bDMARDs*[1,2] and tsDMARDs[#3] should be combined with a csDMARD; in patients who cannot use csDMARDs as comedication, IL-6 pathway inhibitors and tsDMARDs may have some advantages compared with other bDMARDs
10.	If a bDMARD* or tsDMARD[§] has failed, treatment with another bDMARD or a tsDMARD should be considered; if one TNF-inhibitor therapy has failed, patients may receive another TNF-inhibitor or an agent with another mode of action
11.	If a patient is in persistent remission after having tapered glucocorticoids, one can consider tapering bDMARDs, especially if this treatment is combined with a csDMARD
12.	If a patient is in persistent remission, tapering the csDMARD could be considered

The symbols (*, §, #) indicate different levels of evidence which are correspondingly provided together with voting results and levels of agreement in table 3.
[1]TNF-inhibitors: adalimumab, certolizumab pegol, etanercept, golimumb, infliximab boDMARDs or the respective EMA-approved/FDA-approved biosimilars.
[2]Abatacept, rituximab (as first bDMARD under special circumstances—see text), or tocilizumab or respective EMA-approved/FDA-approved biosimilars, as well as other IL-6 pathway inhibitors, sarilumab and/or sirukumab, once approved.
[3]Jak-inhibitors (where approved).
boDMARDs, biological originator DMARDs; bsDMARD, biosimilar DMARDs; csDMARDs, conventional synthetic DMARDs; DMARDs, disease-modifying antirheumatic drugs; EULAR, European League Against Rheumatism; Jak, Janus kinase; MTX, methotrexate; RA, rheumatoid arthritis; TNF, tumour necrosis factor; tsDMARDs, targeted synthetic DMARDs.

Table 7. Description of medication and combination therapy for RA

Drug category	Descriptions
Methotrexate	Used either oral or subcutaneous (a DMARD).
DMARDs§	Traditional/conventional DMARDs including HCQ, LEF, MTX, or SSZ (excludes azathioprine, cyclosporine, minocycline, and gold), it does not include tofacitinib, which is considered separately.¶
DMARD monotherapy	Most often defined as the use of MTX monotherapy, but may also be SSZ, HCQ, or LEF.
Double DMARD therapy	MTX+SSZ, MTX+HCQ, SSZ+HCQ, or combinations with LEF.
Triple DMARD therapy	MTX+SSZ+HCQ.
DMARD combination therapy	Double or triple traditional/conventional DMARD therapy.
Tofacitinib	Oral synthetic small molecule.
Biologics	TNFi biologic or non-TNF biologic (excludes anakinra).§
TNFi biologics	Adalimumab, certolizumab pegol, etanercept, golimumab, or infliximab.
Non-TNF biologics	Abatacept, rituximab, or tocilizumab (excludes anakinra).§
Low-dose glucocorticoid	≤10 mg/day of prednisone (or equivalent).
High-dose glucocorticoid	>10 mg/day of prednisone (or equivalent) and up to 60 mg/day with a rapid taper.#
Short-term glucocorticoid	<3 month treatment.

DMARD = disease-modifying antirheumatic drug; HCQ = hydroxychloroquine; LEF = leflunomide; MTX = methotrexate; SSZ = sulfasalazine; TNFi = tumor necrosis factor inhibitor; COBRA = Combinatietherapie Bij Reumatoide Artritis
§ Anakinra was considered but not included in these guidelines due to its infrequent use in RA and lack of new data since 2012.
¶ Azathioprine, cyclosporine, minocycline, and gold were considered but not included in these guidelines due to their infrequent use in RA and/or lack of new data since 2012.
# Regimen based on that described in the COBRA study (153).

Reprinted with permission from John Wiley and Sons (293)

The symptoms of swollen and tender joints in RA can also be treated with corticosteroids and NSAIDs. Corticosteroids may alleviate any symptoms and inflammation while DMARD initiation takes effect. As these are not disease-specific drugs, corticosteroids and NSAIDs would commonly have been initiated prior to RA diagnosis. One study reported a 9-fold increase in NSAID prescribing in the six-months pre-diagnosis (294). However, both drugs are potentially toxic, and the latter treat the symptoms of swollen and tender joints rather than preventing or reversing disease development (216, 295). Both have associated risk of mortality and hepatic, renal, cardiovascular, bone and gastro-intestinal (GI) disorders, particularly among the elderly and comorbid, which RA patients more commonly are (277, 296).

Following diagnosis, under modern management, targeted DMARD therapy should suppress inflammation and thereby limit the symptoms that underpin continued corticosteroid and NSAID prescribing. Therefore, UK national guidelines recommend only a short-term course (<3 months) of corticosteroids when initiating DMARD prescribing (280). In these guidelines, long-term corticosteroids are only advised where all other treatment options have failed and the associated long-term complications have been discussed with the patient. Similarly, in the UK, NSAIDs are recommended only at the lowest effective dose and for the shortest possible time (238). In the UK, the assessment of disease improvement or progression and medication toxicity, and resultant therapy changes, are generally performed by a rheumatologist, with follow-up managed in primary care. In other countries, there are varied approaches to corticosteroid and NSAID prescribing, e.g. recommendations for low-dose long-term corticosteroids with tapering attempted when remission is reached (297), or

approaches that taper the corticosteroid dosage over a specified duration, e.g. 10 months (298).

Where persistent long-term corticosteroid and NSAID prescribing is necessary, UK guidelines have evolved over recent years to recommend prophylaxis co-prescribing to mitigate their associated side effects. Long-term use of corticosteroids can trigger and exacerbate osteoporosis, with the adverse effects being dose-related (299-301). From 2010, UK national guidelines have recommended bone-protective treatment for patients taking high doses of oral prednisolone long-term, defined as ≥7.5 mg daily for ≥3 months, to prevent osteoporotic fractures (302). Prophylaxis is recommended in the form of co-prescribing of bisphosphonates (alendronate 10 mg once daily or 70 mg once weekly, or risedronate 5 mg once daily or 35 mg once weekly) and / or calcium and vitamin D (302). Bone-protectants are particularly important for women, who are more commonly affected by both osteoporotic fractures and RA (262, 299). NSAIDs are associated with risk of GI and cardiovascular complications (217, 303). From 2008, UK national guidelines for patients with RA have recommended co-prescribing proton-pump inhibitors (PPIs) alongside NSAIDs to prevent GI adverse effects (238).

Ongoing monitoring of RA in the UK is usually performed in primary care, and includes surveillance for osteoporosis, infections including tuberculosis, and malignancy (253). Given the immunosuppressant nature of DMARDs, patients receiving such therapy are advised to receive annual influenza, and five-yearly pneumococcal, vaccines, and to avoid live vaccines (304). The herpes zoster vaccination is also recommended for patients aged >50 years or lacking a history of varicella exposure. Patients should have ongoing drug monitoring, including an appointment 6 months after achieving treatment target, and should understand how and when to request rapid access to specialist care in the event of flares (280). Where treatment target is maintained for over a year without corticosteroids, a cautious step-down approach to therapy can be adopted. In all cases, an annual review should be offered to assess disease activity and functional ability, screen for comorbidities and symptoms related to complications such as vasculitis, and to assess any referral requirements. A multi-disciplinary team may monitor treatment and coordinate additional ongoing care including physiotherapy, occupational therapy, surgery and psychological intervention to support adjustments for living with RA (280).

2.3.3.6 EHR-based Research

In RA studies, EHR data have been used to derive information on comorbidities, management and screening practices, and treatment effectiveness. High validity of RA diagnoses in EHR data has been reported, e.g. 80-84% in studies in Canada and Sweden, and RA disease activity has been predicted from EHR data with high concordance to human expert agreement, suggesting the relevance of EHR data for RA research (305-307). The high prevalence of cardiovascular comorbidities in RA patients has been confirmed across a number of countries and especially in patients with CKD, suggesting the importance of tight control of comorbidities in RA patients (308, 309). EHR data has facilitated the assessment of RA management using quality indicators, and the suboptimal assessment and management of cardiovascular risk in RA patients has been highlighted (310, 311). Prescribing patterns have been examined and the underutilisation of methotrexate has been reported in the US, with only marginal improvement in dosage between 2009 and 2012 (312). EHR data has informed calculations of treatment cost, and the lower cost of biologic therapy with etanercept compared with infliximab has been suggested (313). The long-term safety of therapy, including the safety of surgery in patients with anti-TNF therapy, has been investigated using EHRs (314). UK-based studies will be systematically reviewed in Chapter 6, given the variation across countries in pharmacologic guidelines and the focus on UK prescribing later in this thesis.

2.4 Summary

The use of EHRs in clinical settings is becoming increasingly widespread and EHR research databases offer tremendous potential to accelerate health research. Despite the challenges in using data not initially intended for research, approaches are being developed to utilise and validate EHR data. Gout, AS and RA are inflammatory arthritides that start relatively early in life and contribute a significant individual disease burden, where early diagnosis and pharmacologic management is important for maintaining quality of life. EHR-based research has been conducted in these diseases. However, the research themes in these EHR-based studies and the methodological approaches applied to EHR data handling have not been systematically reviewed. Chapters 4-6 of this thesis aim to address this.

This chapter also established the guidance for the diagnosis and management of gout, AS and RA. The principles of management in gout include the prompt prescription of therapy for flares and long-term maintenance of SUA levels below 6 mg/dL. The

suboptimal management of gout, and resultant high recurrence of flares, has been reported using EHR data. The principles of management have shifted toward the early prescribing of biologics in AS, with studies reporting the effectiveness of early therapy (178, 240). However, the incidence and prevalence of AS in the UK are uncertain and studies using hospital EHR data have reported significant delays in diagnosis, though in the UK symptoms may first be presented to GPs. In RA, the principles of management have shifted toward early prescribing of biologic therapies. In RA management in the UK, early tapering of corticosteroids is recommended, and the use of prophylaxis where long-term corticosteroids or NSAIDs are prescribed. However, while biologic prescribing has increased, other trends in real-world management are uncertain. Chapters 7-9 of this thesis aim to address these uncertainties in the epidemiology and management of AS and RA in the UK.

2.5 Thesis Aims and Objectives

The hypothesis underlying this research is that novel understanding of disease epidemiology and management can be derived from EHR data and that this can provide important information required by clinicians and decision-makers to improve patient management. The aims of this thesis were therefore to describe the epidemiology and management of common inflammatory arthritides in routine general practice by appraising existing EHR-based research and doing novel exploration of EHRs.

The first objective of the thesis is to describe existing EHR-based studies (Chapters 4-6). Real-world trends in gout, AS and RA have been investigated in EHR-based research. Therefore literature reviews of the EHR-based studies in these diseases will be performed. The methodology applied to EHR data, the study themes (e.g. epidemiology, management), and the comprehensiveness of reporting will be investigated.

The second objective is to describe trends in disease epidemiology using EHR data (Chapter 7). In the UK, the incidence and prevalence of AS are uncertain and in RA these have not been examined since 2014 despite changes in payment for coding in 2013. Therefore, the incidence and prevalence of AS and RA in the UK will be investigated over the last two decades using population-based GP EHR data from the CPRD.

The third objective is to describe trends in the timeliness of diagnosis, given its importance for the success of early pharmacologic management (Chapter 8). Delays in diagnosis have been reported in AS using hospital EHR data, yet symptoms of back pain are often first reported in primary care and in the UK GPs play an important role in the timely identification of AS. Therefore, the diagnostic delay in AS in the UK in relation to first evidence of consultation for suggestive symptoms will be investigated using the same source of GP EHR data.

The fourth objective is to describe trends in real-world management using EHR data (Chapter 9). The principles of RA management have shifted following the development of biologic DMARDs from 1998 onwards and, in the UK, guideline recommendations for corticosteroid tapering and prophylaxis. Therefore, the same data source will be analysed to evaluate changes in the prescribing of RA medication and guideline-recommended prophylaxis in the UK over the last two decades.

Through meeting these objectives, this thesis will inform understanding of the contribution of EHR-based research and provide information regarding the epidemiology and management of disease. Some of the challenges and opportunities for future research will also be suggested.

Chapter 3 Materials and Methods

3.1 Introduction

This chapter describes the materials and methods that are common either to the following literature reviews (Chapters 4-6) or to the following analyses using EHR data (Chapters 7-9). Chapter-specific methodology is given in these subsequent chapters. For the literature review methodology used, details regarding the literature search, screening process and study selection are given. For the EHR analyses, the study design and data source are detailed; a feasibility study was performed when determining the data source, which is also presented in this chapter. This preliminary pilot study was undertaken to determine the suitability of the chosen EHR dataset for analyses of the epidemiology and management of inflammatory arthritis. The dataset extracted for the studies in Chapters 7-9 is described as well as the derived study population, study timeframe and criteria for the start and end of patient follow-up.

3.2 Literature Reviews

3.2.1 Literature Search

In order to identify EHR-based studies in gout, AS and RA, literature search terms were selected using the Population, Intervention, Comparison, Outcomes and Study design framework and synonyms were determined based on knowledge derived from the background literature review and by hand searching relevant manuscripts (EHR-based studies identified in Chapter 2, which were anticipated in advance to be eligible for inclusion) (315). The chosen search terms were applied in literature database searches and filters were applied where available to exclude citations published prior to 1970 (highlighted in Chapter 2 as the first decade of EHRs), of non-English language, or with non-human subjects. Google Scholar was also searched. Given the constraints regarding bulk download and complex search query strings in Google Scholar, key search terms were chosen and the publication timeframe was limited to more recent years, as is detailed in the relevant chapters. Harzing's Publish or Perish application (version 6) was used, which enabled bulk download of the first 1,000 citations returned by Google Scholar (316).

3.2.2 Literature Screen

Four tools were considered for the screening process: EndNote, Cochrane Covidence, Abstrackr and Rayyan (317-319). Endnote was selected for performing de-duplication and the screening was then performed in Rayyan (318, 320). Rayyan was selected for screening given its free online accessibility by all reviewers, capacity to set flags and highlight keywords, and tools for shared discussion between reviewers (318).

The citations returned by each literature search were uploaded into an Endnote library with automated de-duplication upon entry, followed by manual de-duplication ("Find Duplicates") to screen articles with matching authors, year, title or secondary title (320). The output was uploaded into Rayyan where a further automated duplication was performed and the remaining studies where screened.

3.2.3 Study Selection

The inclusion and exclusion criteria that were common across the studies are listed here.

Inclusion criteria:

1. Study using general population EHR data
2. Selection of patients with the relevant disease (gout, AS or RA)

Exclusion criteria:

1. Claims, insurance or such data with no reference to being derived from EHRs
2. EHR data were not used in cohort selection and results (e.g. enrolled cohorts and case studies) (as recording practices and management may be different from the general population)
3. Study using data from EHR and non-EHR sources, where separate methods and results are not provided for each source of data used
4. Non-gout / AS / RA specific disease registry or sub-population database, or rare event reporting database
5. Simulation studies (i.e. non-real-world data)
6. Clinical trial or intervention studies where medication management is non-routine
7. Non-human animal study

8. Unavailability of a full English manuscript despite communication with the main author
9. Publication prior to 1970, the first decade with established EHRs
10. Conference, meeting or protocol citation where there is an article already present in the screen list that was published subsequently with the same objectives and first, second or last author
11. Conference or meeting abstract, letter, research protocol or review. These were deemed to include insufficient information on methods and results

There were studies, particularly from the USA and Taiwan, which derived data from records or charts, health claims or insurance databases, with no reference to EHRs. It was unknown whether the data was sourced, for example, from paper charts or information not recorded during the routine course of care for the purpose of patient care. Such studies were excluded despite any methodological similarities to EHR-based studies: this is a previously published approach (321). Further, terms for paper records or charts, claims or insurance databases were not searched and so those that were identified through screening were likely to represent an incomplete sample. Studies referencing use of EHR-claims data were included on the assumption that the source data was EHR. Swedish studies were included, as Sweden is understood to implement an EHR system across the entire health setting.

The screening process was conducted using the citation review manager Rayyan and a pre-defined screening protocol based on the selection criteria (318). For the full-text screen, papers were located through a comprehensive online search and uploaded to Rayyan for review. Where an article could not be found or accessed, through open source, ResearchGate or University of Leeds Library subscriptions, then the corresponding author was contacted with an article request.

3.3 EHR-based Studies

To meet the aims of investigating the epidemiology and management of inflammatory arthritis, a retrospective longitudinal observational analysis was performed, using UK GP EHR data. As described in Chapters 1 and 2, primary care plays an important role in diagnosis and management of patients with inflammatory arthritis. EHRs are used across primary care in the UK, with payment processes contributing to data quality (17, 20). The studies were reported following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (322).

3.3.1 Data Source

Of the primary data sources described in Chapter 2, the CPRD was considered appropriate for conducting the work in Chapters 7-9. CPRD data undergoes quality assessment and patients have comparable demographics to the general UK population (89, 95, 96, 323). CPRD data includes information on data quality that can be used in defining the study population and criteria for follow-up. From information published in the database profiles of the CPRD and THIN, the latter had under-representation in men and people aged <25, and a greater number of GP practices contribute to the former (674 compared with 532) (89, 101). The under-representation of men and younger people in THIN had relevance to the AS analysis given the young age of onset in AS and its higher prevalence in men. Therefore, the suitability of CPRD data was investigated further: the following sub-sections describe the CPRD data in more detail and present a feasibility study that was conducted to confirm the suitability of the CPRD data for this thesis.

3.3.1.1 The Source Data

The selected source database was the CPRD GOLD, containing EHR data from contributing GP practices using the Vision EHR system. The database contains structured data that reflect the administration of routine care and are recorded by system users.

The GP data in CPRD GOLD includes clinical data coded in Read Code Version 2 (RCV2), (324). The RCV2 is a dictionary of legacy codes, last updated in April 2016 (325). Drug and appliance data which is coded in GP EHRs using the Gemscript Dictionary of Medicines and Devices (dm+d) product code system which is updated monthly, but prescriptions are mapped to the British National Formulary (BNF) in CPRD GOLD (324, 325). Prescriptions have a corresponding BNF identifier (ID), drug name and drug substance name. These BNF IDs are linked to one or more BNF chapters, each representing a group of drugs. The administration route and prescription date are also recorded. Most NHS laboratory tests are reported using the Pathology Bounded Code List (325). All code dictionaries were available via NHS Digital (325) and were searched by Samantha Crossfield (SSRC) to identify relevant codes during the study. Prescriptions were identified by SSRC through searches of the BNF chapters, drug names and drug substance names recorded in the dataset.

In addition, patient data in CPRD GOLD includes patient and current GP practice ID (both non-identifiable), sex, month-year of birth, current prescription exemption status and level of capitation supplement, first and current GP registration date, registration gap duration, transfer out date and reason and death date. Patient records include a flag indicating 'acceptable' quality for research, which can be used in designing study population selection criteria to exclude data of lower quality (89). The following are recorded per GP practice: non-identifiable practice ID, geographic region, last data collection date, UTS date. Coding quality is higher in data recorded while GP practices have UTS status (89).

3.3.2 Feasibility Study

A feasibility study was conducted to ensure that the CPRD GOLD database contained the necessary data to investigate the epidemiology and management of inflammatory arthritis.

3.3.2.1 Methods

From the July 2017 version of CPRD GOLD, patients with ≥2 years of data from a UTS GP practice were selected as the study population. Patients with gout, AS or RA were defined based on diagnostic codes recorded in the study period (01 January 1998 to 30 June 2017), excluding patients with <1 year of UTS registration prior to the first diagnosis code or <1 year of subsequent registration. Patients with a subsequent diagnostic code at least 7, 90 and 180 days following the index code were also identified. For patients with ≥3 years of GP data, the proportion with two or more of the three studied diseases was calculated.

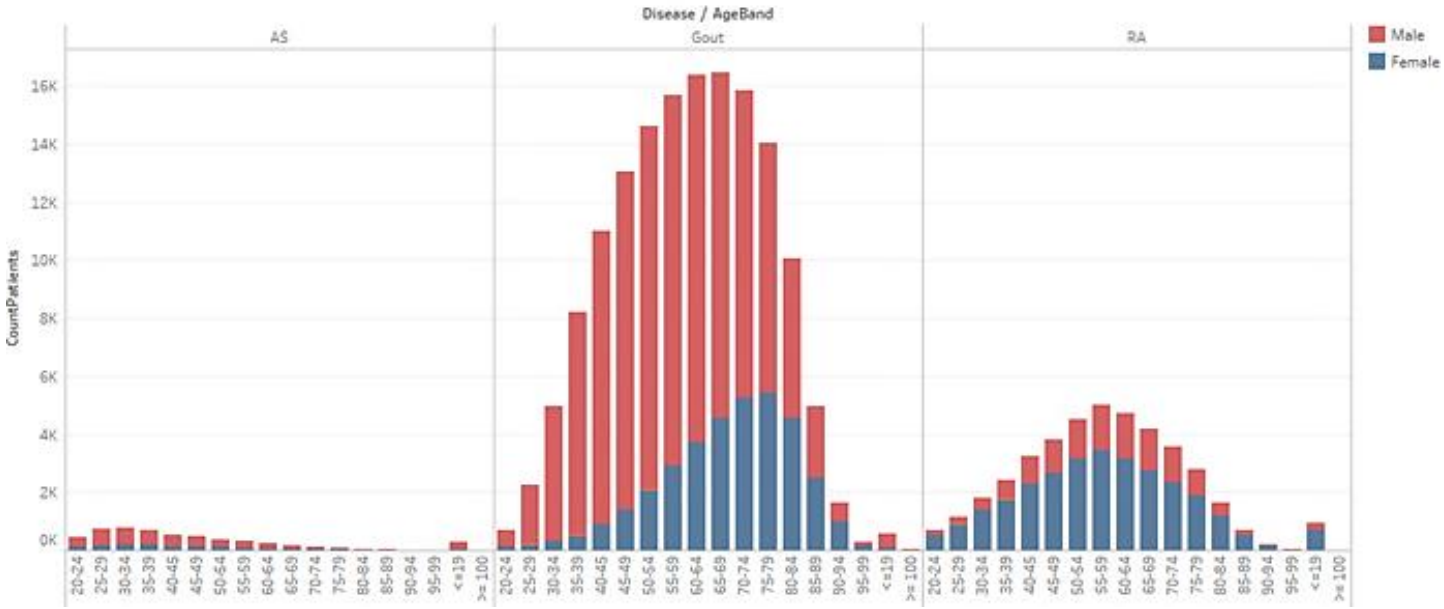
Summary descriptive statistics were used to explore the disease prevalence, follow-up duration, age at diagnosis, and availability of prescription and test data. The number and proportion of the study population with gout, AS or RA diagnosis recorded during the study period was described. The proportion of patients having ≥2, ≥3 and ≥5 years of GP data, and having active (current) GP registration was calculated. The mean age at diagnosis was calculated, for patients aged 18-100 (inclusive) at diagnosis. Patients were also grouped by age at diagnosis (≤19, 5 year bands from 20-24 to 95-99, ≥100) and the mode age group was calculated. The following proportions were calculated to assess prescriptions: patients with ≥1 and ≥5 NSAIDs, patients with gout and ≥1 and ≥5 allopurinol and patients with RA and ≥1 etanercept or adalimumab prescription. For

patients with gout and ≥3 years of GP data, the proportion with 0-10 and >10 SUA level tests was calculated.

3.3.2.2 Results

From a study population of 9,820,731 patients, 150,649, 5,412 and 41,456 patients had gout, AS or RA coded during the study period and one year of data pre-and post-diagnosis. The proportions of females were 23.7% (n = 35,746), 26.5% (n = 1,435) and 69.5% (n = 28,819) respectively (Figure 3). In patients with gout, AS or RA, 60.3% (n = 90,826), 60.2% (n = 3,260) and 66.8% (n = 27,711) had a subsequent diagnosis ≥7 days later; 54.0% (n = 81,293), 58.0% (n = 3,137) and 63.4% (n = 26,264) ≥90 days later; and 51.3% (n = 77,314), 56.0% (n = 3,030) and 61.2% (n = 25,371) ≥180 days later. In patients with ≥3 years of GP data, 519 had gout and AS, 3,238 had gout and RA, 319 had AS and RA, and 26 had gout, AS and RA.

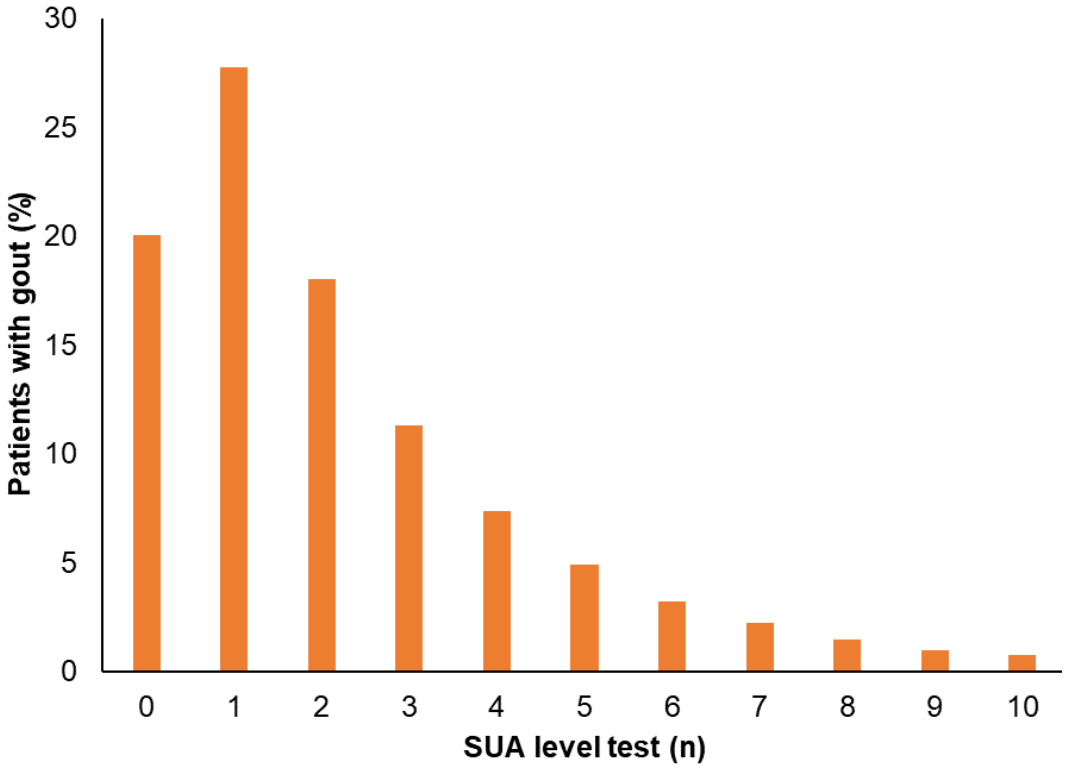
Figure 3. The number of patients with AS, gout or RA, by sex and age-group



Of the study population, 1.5% had a gout diagnosis recorded during the study period, 0.1% AS, and 0.42% RA. The proportions of gout, AS and RA patients with ≥2 years of GP data were 98.4% (n = 148,233), 97.2% (n = 5,259) and 98.3% (n = 40,762) respectively; 98.0% (n = 17,653), 96.3% (n = 5,212) and 97.8% (n = 40,528) with ≥3 years; and 96.9% (n = 145,939), 94.4% (n = 5,107) and 96.4% (n = 39,968) with ≥5 years. The proportions of gout, AS and RA patients that were alive and registered with a contributing GP practice at the time of data extract were 66.8% (n = 100,681), 70.1% (n = 3,792) and 62.9% (n = 26,095). The mean age at diagnosis was 59 (standard

deviation = 19.7; 56 in men, 68 in women), 38 (28; 39 in men, 38 in women) and 54 (26.4; 55 in men, 51 in women). At diagnosis, the mode age groups were 65-69, 30-34 and 55-59. The proportion with ≥1 and ≥5 NSAID prescriptions was high in each disease group, 93.1% (n = 137,496) and 62.7% (n = 92,639) in gout, 95.2% (n = 4,960) and 81.0% (n = 4,220) in AS, and 94.4% (n = 38,258) and 79.5% (n = 32,234) in RA. In patients with gout, 49.8% (n = 73,496) and 40.2% (n = 59,367) had ≥1 and ≥5 allopurinol prescriptions. In patients with RA, 0.6% (n = 251) and 0.3% (n = 128) had been prescribed etanercept and adalimumab. While 80.0% (n = 118,052) of patients with gout had ≥1 SUA level test, the majority of these (34.7%, n = 41,012) had only one test (Figure 4). The proportion with ≥1 SUA level test was higher in patients prescribed ≥1 allopurinol (91%) than in patients with no allopurinol prescription (68%).

Figure 4. Proportion of patients with gout having 0-10 SUA level tests



SUA = serum urate

3.3.2.3 Discussion

The CPRD GOLD database contained patients with inflammatory arthritis and information on their demographics, diagnosis, prescribed medication and laboratory tests. A large cohort of patients with gout, AS and RA was identified in the database and high numbers had >1 code, indicating that additional code requirements could be used in disease definitions for sensitivity analyses. The proportion of patients with each

disease (1.5%, 0.1%, 0.4%) seemed reasonable, given published reports of UK prevalence: 1.4% in gout, estimated 0.1-0.2% in AS, 0.5% in RA (137, 326, 327). Other studies also report comparable ratios of men to women: 4:1 to 3:1 in gout, 2:1 in AS, 1:3 in RA (30, 287, 328). The patterns of age at diagnosis was as expected, e.g. higher in women than men in gout, 30-34 in AS, and lower in women than men in RA (30, 328, 329). However, normal distribution of the data cannot be assumed and median age will be calculated instead of mean age in subsequent studies in this thesis. The feasibility study suggests that the large sample size and representative characteristics of the disease cohorts are suitable for the intended analysis.

The feasibility study highlighted the presence of prescription and test data necessary for investigating management over the disease life-course. For each disease, a large number of patients had ≥5 years of GP registration data in CPRD GOLD, suggesting that this database lends itself to assessing disease management long-term. Around 2/3 of patients had active registration at the date of data extract, suggesting that CPRD GOLD holds data of contemporary relevance. The assessed medication data suggests that GP-prescribed medication (NSAIDs, allopurinol) can be assessed using the database but not medication prescribed in secondary care (etanercept, adalimumab). Further, most patients with the studied diseases have been prescribed NSAIDs despite the potential for toxicity highlighted in Chapter 2. Most patients with gout, especially those prescribed allopurinol, had an SUA test as would be expected, suggesting that the use of this laboratory test can be studied using the database. Therefore, the database was deemed suitable for studies of the epidemiology and management of inflammatory arthritis, and was selected as the data source in this thesis.

3.3.3 Data Request and Data Security

A protocol was designed by SSRC for the studies in this thesis and was reviewed by an epidemiologist, statistician and clinician (Mar Pujades-Rodriguez, MP-R; Paul Baxter, PB; Philip Conaghan, PGC). In May 2018, the protocol was approved by the CPRD Independent Scientific Advisory Committee (ISAC; protocol reference 18_082). The ISAC is an expert advisory body established by the Secretary of State for Health to ensure that approved protocols are of an acceptable scientific standard, are viable and have no governance concerns (330). Extensions to the study timeframe in the protocol have been approved by ISAC annually thereafter.

Ethical considerations apply to the secondary use of patient data for research purposes. However, given the non-identifiable nature of the CPRD data, ethical

approval from the NHS Research Ethics Committee was not required for this study. The non-identifiable nature of CPRD meant that dissemination of the study results to study participants was not possible and there was no patient-public involvement in the study.

The non-identifiable data was extracted by SSRC under an academic licence with the CPRD. The data was stored in the Leeds Institute for Data Analytics (LIDA) Integrated Research Campus (IRC) at the University of Leeds (UoL). A project proposal and data management plan (DMP) were approved by LIDA and a satisfactory risk assessment was conducted. The DMP was developed using DMPOnline as recommended by the UoL Research Data Management Service and IRC (331, 332). All data handling was in accordance with the IRC Information Security Management System, which is self-assessed under the NHS Data Security and Protection Toolkit as providing suitable data security for NHS patient data, and has been externally certified as International Organisation for Standardisation (ISO) 27001 compliant. Remote access to the data in a virtual research environment (VRE) was provided to SSRC under an IRC User Agreement.

3.3.4 Study Dataset

The study dataset was extracted from the 2 April 2018 update of the CPRD GOLD database, which contained 17.6 million EHRs from 734 UK GP practices. The database included 15 million patients with continuous GP registration in CPRD, with 13.1 million having at least 1 year of registration and 10.9 million of these having registration at a single GP practice throughout their period of data contribution. In addition, linked data on patient and practice Index of Multiple Deprivation quintiles was also provided.

3.3.5 Study Population

The study population comprised patients in the CPRD population with the following:

1. An ‘acceptable’ patient flag (indicating record quality)
2. ≥1 day of continuous registration (not temporary status) during the study period (1 January 1998 to 1 April 2018)
3. Sex ‘male’ or ‘female’ (i.e. not missing/unknown/indeterminate)

Patients in the study population contributed data from the latest of:

1. Study start date (1 January 1998)

2. Becoming aged ≥18 years
3. Having ≥1 year of UTS registration at the GP practice

Patient follow-up in the study ended at the earliest of five events:

a. Study end date (1 April 2018)
b. Last data collection from the patient's GP practice
c. Patient de-registration (transfer out) from a GP practice contributing to CPRD GOLD
d. Death
e. Becoming aged ≥101 years

3.3.6 Statistical Analyses

The CPRD GOLD dataset was loaded into a structured, relational database in Microsoft Structured Query Language (SQL) 2017 using SQL Server Integration Services. R Version 3.6.2, Microsoft SQL 2017 and Microsoft Excel 2016 were used in all analyses.

The statistical analyses performed were chapter-specific and are defined in Chapters 7-9.

Chapter 4 Gout: A Systematic Literature Review of Electronic Health Record Research

4.1 Introduction

As described in Chapter 2, gout is an increasingly common inflammatory musculoskeletal disease, affecting approximately 2.5% of the UK population (113, 116). It results from MSU crystal deposition in articular and peri-articular tissues that leads to debilitating flares and joint damage (106). The role of medication in acute and chronic cases, and particularly ULT, was described in Chapter 2 (33, 106, 139). Adequate dosage and successful treatment can be assessed through quantification of SUA levels. There are comprehensive international guidelines and quality indicators for such management of gout in primary care (33, 106, 122, 139, 142). Despite this, suboptimal management is common, with only around a quarter of patients receiving effective treatment and 80% experiencing recurrent flares within three years of diagnosis (109, 116, 150, 151). This pattern of a rising burden with suboptimal management despite guideline recommendation for long-term medication is common to many chronic diseases (294, 333, 334).

Chapter 2 described the growing use of EHR systems in patient management and the issues in EHR-based research. Many EHR systems include electronic prescribing and these records enable evaluation of patient management and exploration of issues such as guideline-indicated treatment and medication adherence. These data are informative for diseases such as gout that are primarily managed with pharmacological interventions. For example, measures of compliance to treatment guidelines and medication adherence can inform understanding of routine care. EHR-based research in gout has examined temporal and demographic variations in treatment, quality of management and patient outcomes (61, 116, 152, 335). However, as noted in Chapter 2, the secondary use of EHR data requires understanding of how EHRs are used in clinical practice in order to design studies appropriately (e.g. clinicians may use different codes to record an event depending on their training, experience, the coding system used in the EHR and the design of the EHR system).

EHR-based studies may use a variety of approaches to define and validate events or cases, and to ascertain medication prescribing and outcomes. It is uncertain whether the estimations of medication management may vary depending on these approaches.

An understanding of the factors that determine heterogeneity in estimates is essential for interpreting study findings. Variation in the approaches and results, and in the comprehensiveness of reporting and study quality, in EHR-based studies of gout medication has not been assessed.

This chapter aims to inform the first objective of this thesis through a systematic literature review of the methods, results, reporting and risk of bias in EHR-based studies evaluating pharmacologic management. Gout, being a common inflammatory rheumatologic condition with predominantly pharmacologic management and commonly studied in the literature, was selected as the study focus. The review aims to describe the variation in methodological approaches across studies and assesses how this variation affects study findings, as well as changes in reporting and risk of bias over time. It identifies best practice and suggests necessary improvements for both reporting and for consideration of bias relevant to studies using such data. This systematic investigation of the quality of reporting on EHR studies may inform considerations for the standardisation of the definitions used, measurement, consideration of bias and reporting in EHR-based research, which would foster comparison of estimates across studies. This is the first review in this area and the findings should be relevant to the study design, reporting and interpretation of EHR-based studies both in gout and other diseases.

4.2 Methods

The systematic review was conducted in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Appendix A: Table A 1) (336).

4.2.1 Protocol

Having defined the study aim, a search was performed on Google Scholar, PubMed, International Prospective Register of Systematic Reviews (PROSPERO), NHS Economic Evaluation Database, Health Technology Assessment (HTA) and the Cochrane Database of Systematic Reviews, to check for existing systematic literature reviews, or any under development, on this topic. Having found none, a protocol was developed, informed by guidance from the Centre for Reviews and Dissemination, the Population, Intervention, Comparison and Outcomes Study (PICOS) design and the recommendations of Denison et al. (337, 338). The protocol was registered in

PROSPERO (number CRD42017065195) (339). Ethical approval was not required for this systematic review.

4.2.2 Literature Search

The protocol set the policy for database searches and for selecting studies using general population EHR data to report on gout medication exposure and outcomes. The search terms were ‘gout’, ‘medication’ and ‘EHR’. Synonyms and Medical Subject Headings (MeSH) for the search terms were identified, informed by clinical review (PGC) and a hand search of the terms used in the EHR-based studies of pharmacologic management identified in Chapter 2, which were anticipated in advance to be eligible for inclusion. A pilot search and full screen was conducted using Medline, and where the anticipated studies were missing they were again hand searched for additional terms until they were included. A final list of 75 synonyms and MeSH was derived (Table 8). The synonyms for a term were combined with ‘OR’ and then terms were combined with ‘AND’ so that the string searched was ((“EHR” OR “electronic health record”…) AND (“Gout” OR “podagra”….) AND (“medication” OR “treatment”…)).

Table 8. The search terms and synonyms used in the systematic literature review

EHR	Gout	Medication	
Electronic health record+	Podagra	Treatment	“drug monitoring”+
Medical records systems+	Gouty	Pharmacotherapy	Pharmacovigilance+
Record-linkage	Arthritis, gouty+	Drug*	“pharmaceutical preparations”+
Routin* ADJ5 data		Allopurinol+	Prescription drugs+
(Electronic OR link* OR compute* OR anonymi*ed) ADJ5 record		Benzbromarone+	Drugs, generic+
(Health OR patient OR clinic* OR medic* OR care) AND (record* OR data OR plan* OR chart*) AND (compute* OR system OR electronic OR warehouse OR link* OR dataset OR network)		Medication systems+	Prescription
		Drug therapy+	“antirheumatic agent”*+
“System”		Proben*cid+	Medication adherence+
EPR		Sulfinpyrazone+	Adheren*
EMR		Sulphinpyrazone	Anti-Inflammatory Agents, Non-Steroidal+
EHR		Colchicine+	Uricosuric
Database		Febuxostat+	Uricorsuric agent*+
Datalink		“Xanthine oxidase”+	Medication therapy management+
		“Urate lowering”	Drug therapy management+
		ULT	NSAID
		Prescribing	“gout suppressant”*+
		Therapy	“anti-gout agent”*
		“Anti-rheumatic drug”	Prescriptions+
		Drug prescriptions+	Non-steroidal anti-inflammator*

Note: MeSH terms are indicated by ‘+’ and a wildcard by ‘*’. EHR = Electronic health record

The literature search was conducted on 1 August 2017 and re-run twice (7 August 2018 and 20 February 2019) to update the review. An automated alert was set up on those databases that offered such functionality, which triggered email notifications of any subsequent additions that matched the search query. The notifications were checked during the screening process but all manuscripts were duplicates or deemed irrelevant based on their title and abstract and so were not included in the review.

The following were searched:

1. Scopus
2. Web of Science Core
3. Cumulative Index of Nursing and Allied Health Literature (CINAHL)
4. PubMed
5. Ovid Medical Literature Analysis and Retrieval System Online (MEDLINE) Epub Ahead of Print, In-Process and Other Non-Indexed Citations, Ovide MEDLINE Daily and Ovid MEDLINE
6. Embase Classic + Embase
7. Google Scholar

The MeSH terms were used on PubMed, CINAHL, MEDLINE and Embase. Filters were applied where available to exclude citations published prior to 1970, of non-English language, or with non-human subjects.

Given the constraints in Google Scholar, regarding bulk download and complex search query strings, Harzing's Publish or Perish application (version 6) was used to search the terms ("electronic" AND "record" AND "gout") AND ("medical" OR "health" OR "patient" OR "clinical" OR "medication" OR "prescription" OR "drug") and to download the first 1000 citations published between 01.01.2000 and 01.08.2017, 01.01.2017 and 07.08.2018, and 08.08.2018 and 20.02.2019 (316).

Table 9 provides the number of returns from each database, per search term and their combination, and the impact of filters, from the first search (August 2017). The returned citations were uploaded into Endnote for the purpose of de-duplication as described in Chapter 3. The de-duplicated manuscripts were then uploaded into Rayyan, where manual de-duplication was performed using the Rayyan matching probability algorithm, before screening commenced. Table 10 shows the number of manuscripts returned by each database, the number in EndNote after de-duplication, and the number in Rayyan after further de-duplication.

Table 9. Search term returns and the number from each database added to EndNote (August 2017)

Database	Search Term	Returns	Filter		
			1970+	English	Human
Scopus	1	14,219,620			
	2	3,053			
	3	13,517,062			
	1 AND 2	304			
	1, 2, AND 3	199	199	174	158
Web of Science	1	8,017,173			
	2	11,198			
	3	6,119,861			
	1 AND 2	907			
	1, 2, AND 3	523	523	504	504
CINAHL	1	276,874			
	2	2,123			
	3	1,283,690			
	1 AND 2	177			
	1, 2, AND 3	124	124	124	55
PubMed	1	885,874			
	2	16,300			
	3	11,870,379			
	1 AND 2	428			
	1, 2, AND 3	280	280	266	206
Medline	1	2,870,413			
	2	15,691			
	3	5,695,113			
	1 AND 2	935			
	1, 2, AND 3	476	468	427	321
Embase	1	5,454,939			

2	23,914			
3	10,210,395			
1 AND 2	3,184			
1, 2, AND 3	2,105	2,051	1,863	1,586

Note: Search term key: 1 = EHR (electronic health record); 2 = gout; 3 = medication

Table 10. Citation returns per database and the number added to Endnote and Rayyan after de-duplication

Database (in order)	Returns (Aug 2017)		Returns (Aug 2018)		Returns (Feb 2019)	
	Total	Added to Endnote	Total	Added to Endnote	Total	Added to Endnote
Scopus	158	102	203	139	1	1
Web of Science	504	391	248	140	200	86
CINAHL	55	43	14	1	12	3
PubMed	206	81	101	24	5	2
MEDLINE	321	305	83	14	111	23
Embase	1,586	1,201	191	121	11	6
Google Scholar	996	956	998	881	999	700
Total	3,826	3,172		1,316		821
Added to Rayyan after de-duplication		2,906		1,276		810

Further literature was identified through hand searches. During the process of full paper screening described below, any systematic reviews, conference and meeting abstracts and research protocols regarding EHR-based studies of gout medication management were identified and hand searched. These articles were themselves excluded for having limited information for systematic analysis. The references of any review articles were hand searched to find any missed EHR-based studies. For each conference and meeting abstract and protocol, searches were applied in Google Scholar using the first, second and last author names in turn for each manuscript, alongside the term ‘gout’, to select any papers published subsequently. Filters were set to exclude patents and citations and to search from the year of publication. Any manuscripts identified through these means underwent screening.

4.2.3 Study Selection

The selection criteria applied, in addition to those stated in Chapter 3, are as follows.

Inclusion criteria:

3. Reporting on routine gout medication exposure and / or outcomes among gout patients. This included aggregated reporting on medication types (e.g. ‘anti-gout medication’)

Exclusion criteria:

12. Gout is not defined but implied e.g. from a definition of hyperuricaemia or high SUA level, which may be asymptomatic, or anti-gout medication, which may be prescribed in other instances such as haematological malignancy (the exception being febuxostat, which is rarely prescribed except in gout treatment)

Using a study selection protocol, which was written to ensure reproducibility and standardisation, two reviewers (Lana Lai [LLYH] and SSRC) performed the blind screen. The study selection protocol was informed by the experience of the pilot. Taking a conservative approach, manuscripts with uncertain relevance were marked for inclusion during title-abstract screening and labelled ‘uncertain’ during full screening. The reviewers met following the first 50 title-abstract screens and again on completion of each screen, to resolve conflicts and uncertainties, with clinical and epidemiological guidance (MP-R, PGC, SRK). A full manuscript was found for over 99% of cases. Table 11 lists the pre-defined labels used to record exclusion reasoning during full-text screening, based on the exclusion criteria. The labels were ordered based on their relevance to the review aims - first by the exclusion reasons relevant to the three search terms (gout, medication and EHR). Each excluded article was assigned the first relevant exclusion reason.

Table 11. Ordered exclusion reasons with count of citations per exclusion reason

Order	Exclusion Label	Exclusion Reason	Count per Screen			
			Aug 2017	Aug 2018	Feb 2019	Total
1	1 Not EHR study	EHR data not used in cohort selection and results, and data collection is not routine (for example, a case study or trial)	88	19	16	123
2	2 EHR Combined	EHR data were only reported in combination with data from other sources	15	3	3	21
3	3 Not Gout	Gout is not a requirement in cohort or sub-group selection, or gout patients are not reported on separately from other cohorts	38	10	12	60
4	4 No Medication	No reporting on gout medication exposure or treatment outcome, among gout patients	102	25	6	133
5	5 Select Population	Non-gout disease registry or database, or rare event reporting database, rather than whole gout population	3	3	2	8
6	6 Non-human	Non-human animal study	0	0	0	0
7	7 Foreign language	English language version not available	0	0	0	0
8	8 No Full Text	Unavailability of a full manuscript	6	0	0	6
9	9 Replaced	A conference / meeting / protocol where a subsequent related article is already present in the screen list	18	1	1	20
11	11 Pre-1970	Publication prior to 1970, the first decade with established EHRs	0	0	0	0
12	12 Not Explicit EHR	Studies using claims, insurance or such databases that do not reference having an EHR source	45	20	0	65
13	13 Conference or meeting	Abstract from a conference, meeting or symposium	59	22	0	81
14	14 Review	Review, including systematic and meta-analysis	11	5	4	20

Note: EHR = Electronic health record

4.2.4 Data Extraction and Quality Assessment

A data extraction form was developed in Microsoft Access 2013 and was refined during the pilot. The form was based on the recommendations of the Centre for Reviews and Dissemination (CRD) for general information, study and participant characteristics, setting and results (337). The recommendations on page 30-31 were particularly instructive. Characteristics such as 'recruitment procedures', 'costs and resource use' were adapted to EHR-relevant details (e.g. clinical setting, data type, population coverage). Clinical and epidemiological guidance also informed the data elements extracted (PGC, PB and MP-R), along with the Critical Appraisal Skills Programme Cohort Study Checklist and published literature examining the methodology and reporting of EHR-based studies in other diseases (340-343). An accompanying extraction protocol facilitated standardised extraction.

During data extraction, the approach taken by each study to identify gout diagnoses was classified as "liberal" if there was a risk of over-classifying (high sensitivity) and "stringent" if there was risk of under-classifying. Liberal approaches required a single diagnostic code, free-text keyword or prescription unless a rheumatologist recorded it. Stringent approaches used further additional requirements concerning a specialist care setting, having further diagnostic codes or prescriptions, having tests, or meeting diagnostic criteria such as the 1977 ACR criteria (344).

The comprehensiveness of reporting (CoR) of EHR data use was assessed using the REporting of studies Conducted using Observational Routinely-collected Data (RECORD) statement (345). Study quality was assessed using the Cochrane Tool to Assess Risk of Bias (RoB) in Cohort Studies (346). Previously published systematic reviews have similarly used these tools (347, 348). Table 12 summarises the guidance used on CoR and RoB scoring, informed by the published literature and epidemiological reasoning. The scores were calculated applying the Care Quality Commission Survey Scoring Method; answers ranked from 0-10 for the least to most positive answer option (349). The sum of answers per study was divided by the number of questions evaluated to obtain an overall score for each study. Where a question was 'not applicable', it was excluded and did not affect the score (350).

Table 12. CoR and RoB scoring protocol

ID	Item	Guidance

RECORD 1.1	The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included.	Yes: health record / equivalent (no need for electronic as type not storage format) Partly: less clear; not all types named; database name only (345)
RECORD 1.2	If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract.	Yes: both region and timeframe Partly: one (or part of both)
RECORD 1.3	If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	No: using linked data and not stated. Integrated sources such as CPRD are one source, while CPRD and HES are linked
RECORD 6.1	The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.	
RECORD 6.2	Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.	Only refers to validation of codes used in cohort selection. No: no validation performed and no correct reference Partly: validation conducted but of poor quality or incomplete reporting; indirect study evidence e.g. if patients had events that confirm diagnosis; indirect referenced evidence e.g. reference to study selecting based on guideline definition, using the same outcome code/algorithm and finding similar results
RECORD 6.3	Consider use of a flow diagram or other graphical display to demonstrate the selection and linkage process, including the number of individuals with linked data at each stage.	Yes if in methods or results
RECORD 7.1	A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Partly: describe boundaries e.g. BMI or eGFR groups but no coded information
RECORD 12.1	Authors should describe the extent to which the investigators had access to the database population used to create the study population.	Yes: provide good detail about the database population and setting
RECORD 12.2	Authors should provide information on the data cleaning methods used in the study.	No if none stated Yes: 2 of the following (or similar) Partly: 1 of the following (or similar) • Handling of overlapping, misaligned or potentially duplicated prescriptions

		• Reference to estimation of duration of prescriptions • Detail about calculation of drug doses • Selection of 'up to standard' records and exclusion/deletion when this is not met • Reference to handling of records with missing values - imputation, missing category, exclusion etc. • Reference to use of range of impossible values (e.g. convert to missing if BMI<10) • Suitable detail about handling of missing values (e.g. count affected, analysis result pre- and post- imputation) • Excluded paper-only records • a posteriori quality control
RECORD 12.3	State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Partly: linked or pre-linked and little/no detail No: seem linked but no detail on linkage level or method is given NA: unlinked data
RECORD 13.1	Describe in detail the selection of the persons included in the study including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Yes: break down given by selection criteria, or full database count and final with much interim detail, in methods or results Partly: final cohort count given with little / no further detail, in methods or results
RECORD 19.1	Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Yes: discussion of all relevant aspects or ≥2
RECORD 22.1	Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Yes: information on ≥1 of these (e.g. contact for data request) Partly: other supplementary information or protocol ID and name of where submitted for review
Cochrane 1	Was selection of exposed and non-exposed cohorts drawn from the same population?	NA: descriptive study, no comparator Yes if match/control/group comparison, unless unequal exposure opportunity
Cochrane 2	Can we be confident in the assessment of exposure?	Definitely yes if rheumatology clinic, code plus prescription, sensitivity analysis with similar results or prescription with comprehensive coverage

Cochrane 3	Can we be confident that the outcome of interest was not present at start of study?	Definitely yes: outcome is measureable and steps taken to exclude extant cases. If gout incidence, 12 months prior data, if medication then 6 months prior data. NA: descriptive study where 'start' is not a concept
Cochrane 4	Did the study match exposed and unexposed for all variables that are associated with the outcome of interest or did the statistical analysis adjust for these prognostic variables?	Definitely yes: considered at least age, sex and organisation (if multiple), or variables selected following univariate analysis
Cochrane 5	Can we be confident in the assessment of the presence or absence of prognostic factors?	Definitely yes: laboratory or prescription (where there is comprehensive coverage) only and no unconsidered prognostic factors NA if descriptive study
Cochrane 6	Can we be confident in the assessment of outcome?	Definitely yes: prescription or laboratory (with comprehensive coverage) Probably yes: diagnosis without test, prescription without full coverage
Cochrane 7	Was the follow up of cohorts adequate?	Definitely yes: minimum follow-up is 1 year or suitable and exposure isn't likely to affect follow-up duration Probably yes: mean follow-up is >1 year but standard deviation is not No: follow-up duration is not balanced between groups, or too short NA: outcome is not about X following Y
Cochrane 8	Were co-Interventions similar between groups?	NA: observational. Largely not considering lifestyle, advice or other subjective / inconsistent / rare factors Probably yes: reason to suspect a difference in a relevant factor, e.g. if comparators are from diverse countries or have different demographic factors

Note: CoR = Comprehensiveness of Reporting; RoB = Risk of bias; RECORD = REporting of studies Conducted using Observational Routinely-collected health Data; CPRD = Clinical Practice Research Datalink; HES = Hospital Episode Statistics; NA = Not applicable

Two investigators (SSRC and LLYH) independently abstracted the data and met following abstraction of the first 1 and 5 manuscripts to discuss inconsistencies and refine the extraction protocol and data extraction form (as recommended by Whiting et al., 2011) (351). For example, for studies reporting findings for multiple cohorts, an amendment was made to the form, which enabled findings to be grouped by cohort.

For the objective aspects of study/participant characteristics and setting (cf. measures and estimates of medication / efficacy / safety, CoR and RoB), LLYH performed abstraction for a random 20% sample. A similar sampling approach is seen in published literature (347). Review discrepancies were resolved by consensus with a third reviewer (MP-R).

4.2.5 Study Outcomes

The outcomes were: indicators of gout diagnosis; medication types considered; methods/results relating to treatment utilisation (including their period of assessment in relation to the timing of gout diagnosis); efficacy and safety and association between these results and the gout definition used (liberal or stringent); CoR on EHR data use and RoB indicators, including analyses of time-trend and according to gout definition; and study size.

4.2.6 Statistical Analysis

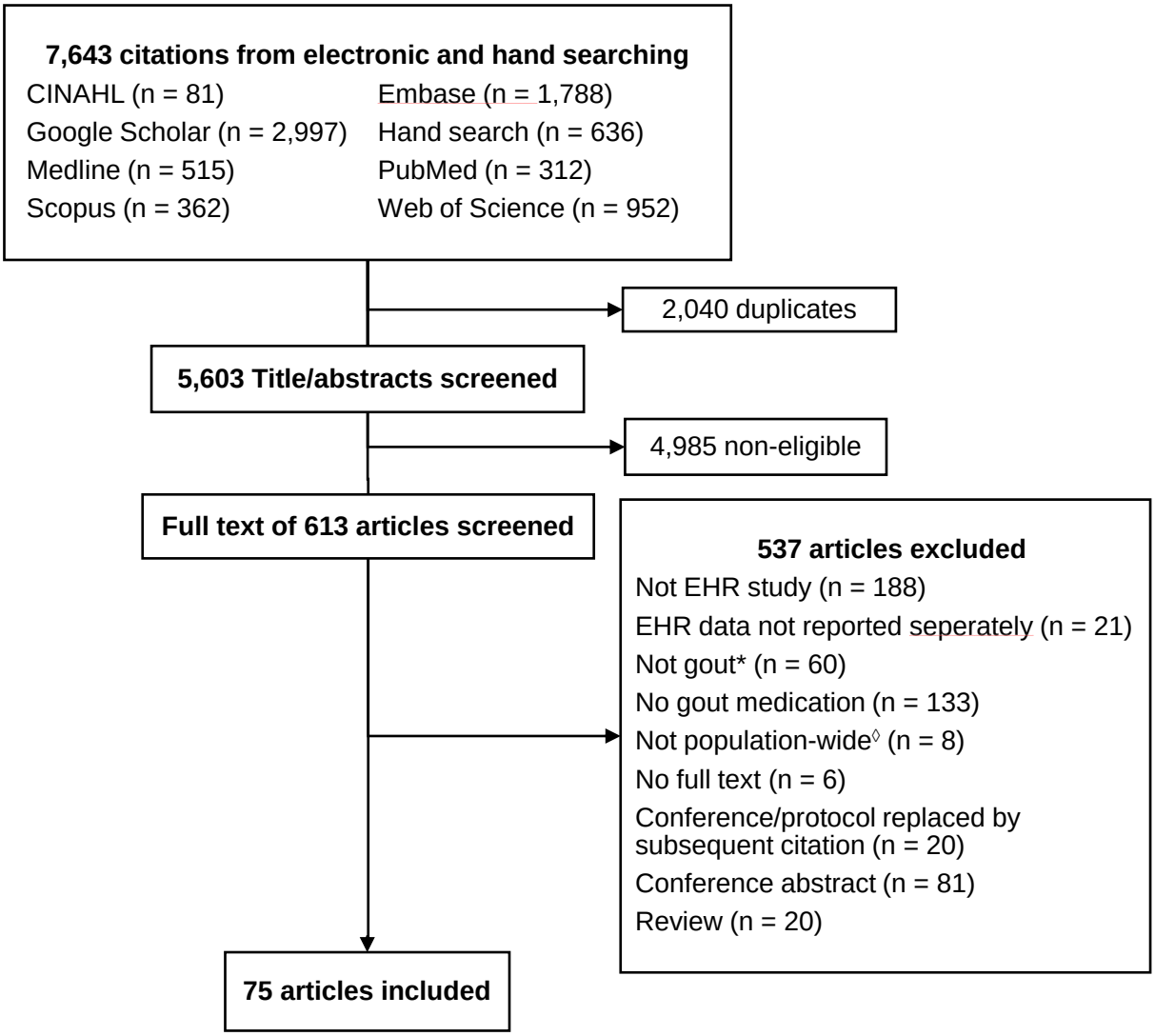
Descriptive statistics were performed by SSRC, using Microsoft SQL 2014 and Excel 2013, to describe the definitions of gout, methods used, treatment utilization, efficacy and safety, reporting quality and RoB. To account for study sample size, weighted proportions and weighted means or medians were calculated for cohort characteristics, overall estimates of medication use and treatment outcomes. Chi-square tests for trend were performed to assess time trends.

4.3 Results

The titles and abstracts of 5,603 articles were screened; 613 full-text articles were reviewed and 75 met the eligibility criteria (Figure 5, Appendix A: Table A 2). All of the selected studies were published between 2002 and 2019, with a rising publication rate ($R^2 = 0.86$) (Figure 6). The mean number of publications per year, in 5-year increments (2002-06, 2007-11, and 2012-16), was 0.6, 2.0 and 6.6. Most studies were conducted in the UK (n = 26) (Table 13, Figure 7). Amongst 67 studies that reported the study period covered, the median study duration was 8 years (IQR = 3.5-15) and 52 (77.6%) used data recorded since 2010 (Figure 8). The mean lag, from the year of study period end to the publication year, was 3.77 years (range 0-12). Fifty (66.7%) studies analysed data from primary care. Thirty-eight studies (50.7%) reported the number of sites considered and 19 (50.0%) of these studies used a single-centre site (range 1-

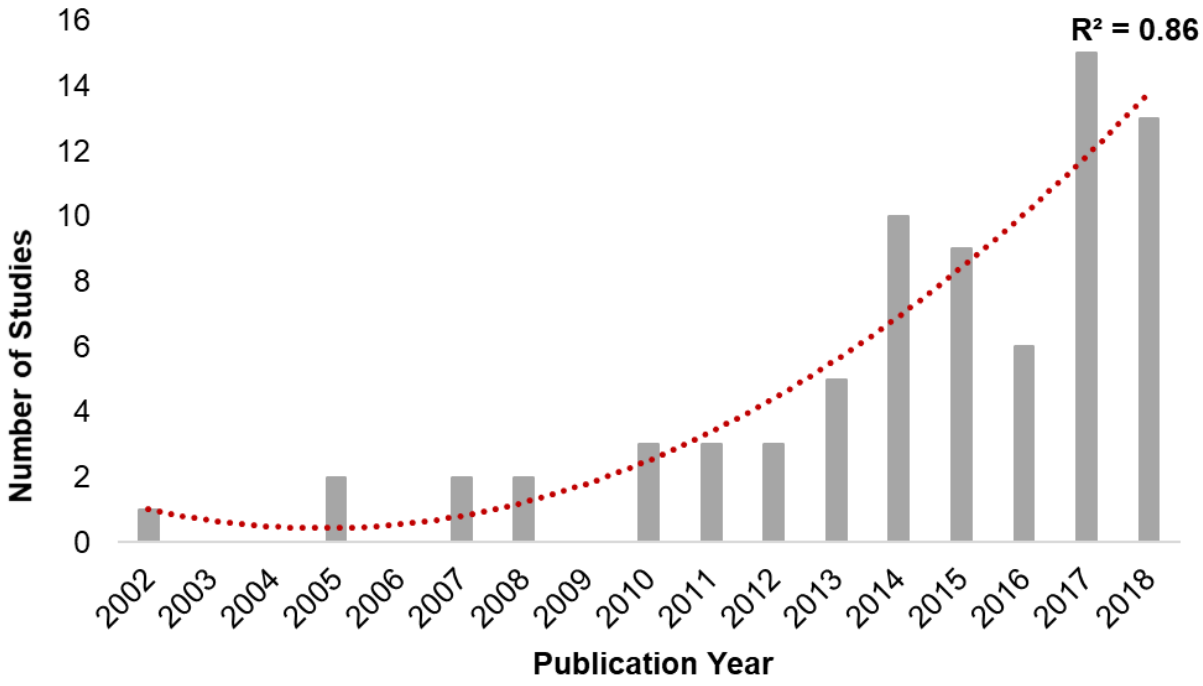
15,520). Only 31 (41.3%) reported the population size from which the cohort was drawn (total or study eligible), ranging from 8,686 to 35 million. Forty-six (61.3%) studies were cohort studies, 16 (21.3%) case-control, 6 (8.0%) matched, 6 (8.0%) cross-sectional and 1 (1.3%) site-randomised (352). The commonest themes were 'epidemiology of gout' and 'adherence to clinical guidelines'.

Figure 5. Flow chart of study identification and selection



Note: CINAHL = Cumulative Index of Nursing and Allied Health Literature; EHR = Electronic health record. * Studies including asymptomatic hyperuricaemia. ◊Studies using databases that are restricted to specific (non-gout) sub-populations (e.g. an adverse event database).

Figure 6. Frequency of articles by publication year, 2002-2018 (N = 74)



Note: The dotted line represents a polynomial regression line

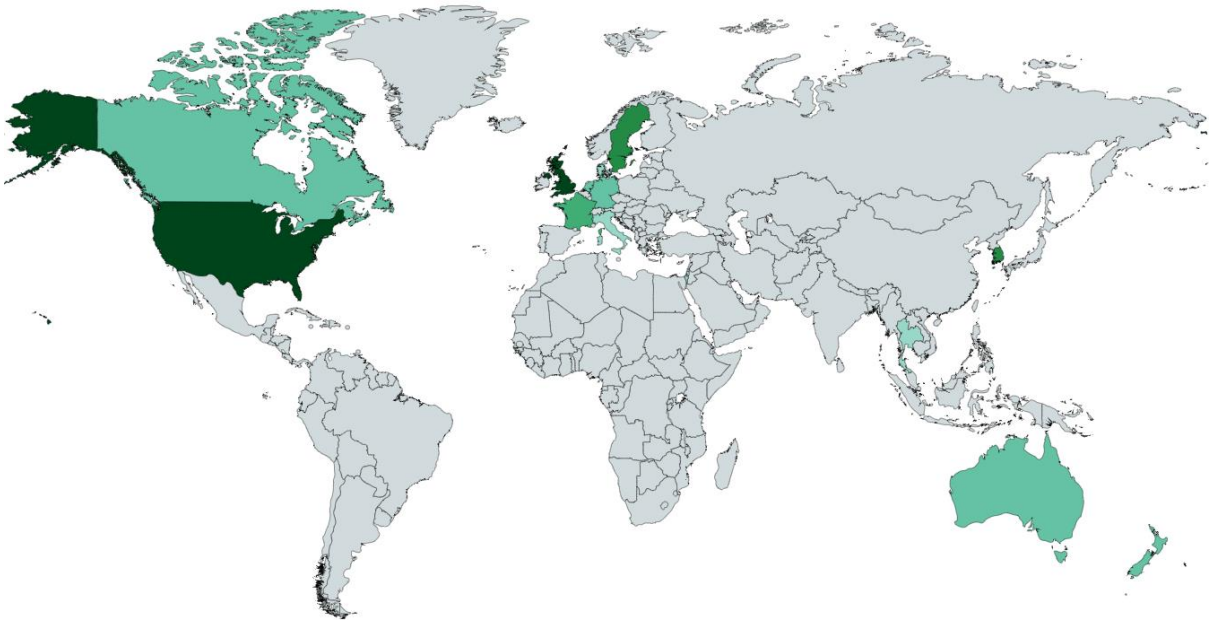
Table 13. Characteristics of the studies included (N = 75)

Characteristic	N (%)	Characteristic	N (%)
Geographic Setting*		Site Type	
Western Europe	41	Primary care	29 (39)
North America	25	Primary care and hospital	21 (28)
Asia	8	Hospital	13 (17)
Australia / New Zealand	4	Outpatient	4 (5)
Middle East	1	National dataset	7 (9)
Not specified	2	Nursing Facility	1 (1)
Study Design		Year of Publication	
Site-randomized trial (usual care cohort)	1 (1)	2000–2004	1 (1)
Matched cohort	6 (8)	2005–2009	6 (9)
Cohort	46 (61)	2010–2014	24 (35)
Case Control	16 (21)	2015–February 2019	44 (64)
Cross-sectional	6 (8)	Gout Cohort Size	
Study Aim		≤100	7 (9)
Epidemiology of gout	22 (29)	101 – 1,000	21 (28)
Patient management	6 (8)	1,001 – 10,000	15 (20)
Adherence to clinical guidelines	12 (16)	10,001 – 100,000	22 (29)
Adherence and gaps in therapy	5 (7)	>100,000	7 (9)

Treatment safety	10 (13)	Not specified	3 (4)
Treatment effectiveness	3 (4)	Drug Data Source	
Patient knowledge, beliefs & education	1 (1)	Prescription	45 (60)
Epidemiology; patient management	7 (9)	Dispensed prescriptions	20 (27)
Other combination	9 (12)	Prescription and dispensary	3 (4)
Time from Study End to Publication		Term search	1 (1)
0-2 years	20 (27)	Not specified	5 (7)
3-5 years	38 (51)		
6-12 years	12 (16)		
No end date specified	5 (7)		

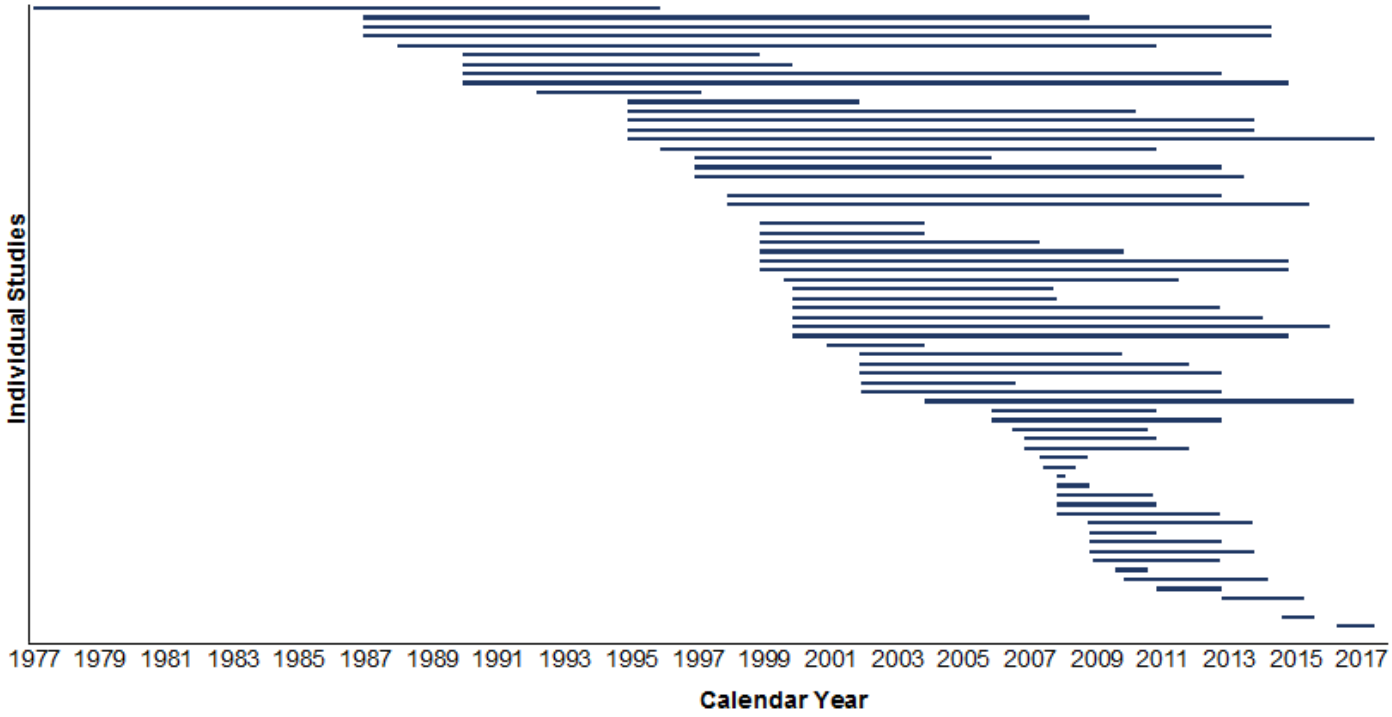
Note: *Some studies had multiple applicable settings

Figure 7. Chloropleth map of countries represented in the studies (N = 75)



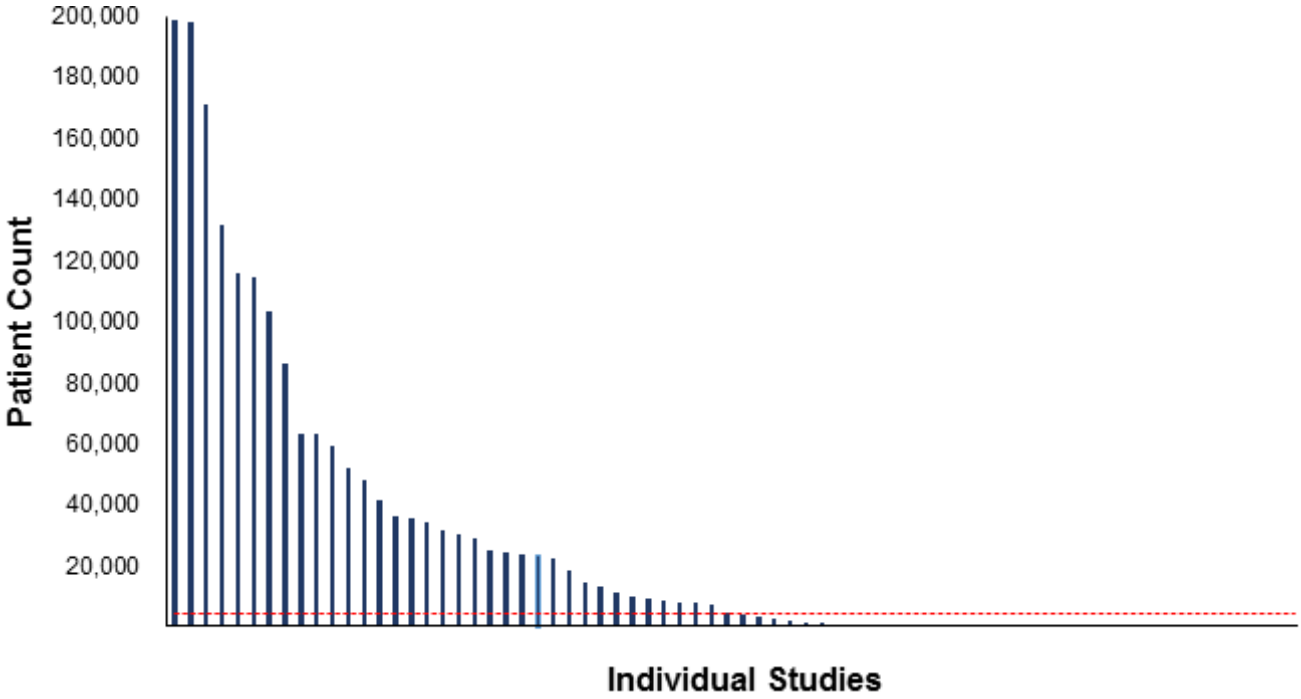
Note: Depth of green indicates the number of studies per country

Figure 8. Study observation period, per study (N = 67)



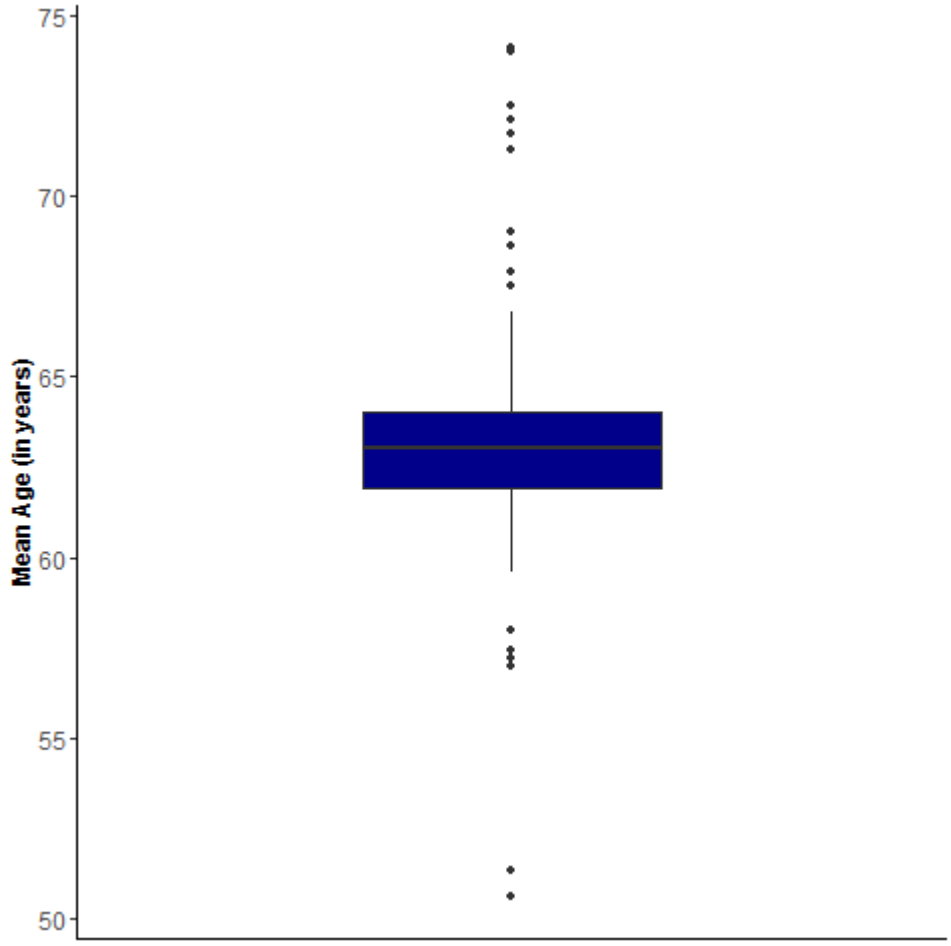
The study selection criteria, cohort size and baseline characteristics were assed. Besides inclusion criteria regarding gout diagnosis or medication, some studies used further criteria regarding: comorbidity (n = 17, 22.7%), age (n = 38, 50.7%, most commonly ≥18 or 20-89), minimum period of enrolment before follow-up start (n = 22, 29.3%), minimum follow-up duration (n = 14, 18.7%), and minimum number of visits during the study or current registration status (n = 13, 17.3%). Six studies (8.0%) excluded patients with missing demographic, prescription or laboratory values and 8 (10.7%) only included sites that met data entry quality standards. While 68 (90.7%) selected all eligible patients, 4 (5.3%) selected a random sample, and 3 (4.0%) selected consenting patients. The median gout cohort size was 4,368 patients (IQR = 435-30,767) (Figure 9) and increased over time (e.g. 19,564 (242-23,594) ≤2015 and 34,538 (612-34,505) >2015). Fifty-nine (78.7%) reported the sex distribution (weighted mean 74.1% male, range 47.8-100%), 50 (66.7%) the mean age (weighted mean 63.1, range 50.6-74.1) and 11 (14.7%) patient deprivation (Figure 10). Twenty-five (31.9%) reported mean/median follow-up duration, which was over 5 years for 12 studies.

Figure 9. Number of patients with gout, per study (N = 72)



Note: Red line represents the median cohort size (n = 4,368)

Figure 10. Box plot of weighted mean cohort age across the studies (N = 49)



Note: Horizontal lines are median and interquartile ranges (25th and 75th percentiles); whiskers' ends indicate the maximum and minimum values at most 1.5 times the interquartile range from the hinge; dark individual dots are outlier values.

4.3.1 Gout Definition

Of 66 (88.0%) studies specifying the gout definition, 38 applied a liberal and 28 a stringent approach; 58 (87.9%) used diagnostic codes, 13 (19.7%) medication, 6 (9.1%) test results and 3 (4.7%) free text (Table 14, Appendix A: Table A 3). Diagnostic coding was optional in 8 studies and required by 50 (38 only required this). When using diagnostic codes, 51 (87.9%) referenced the coding system used but only 25 (43.1%) provided the code-list (Table 15). None of the 13 studies using medication in defining gout provided the medication code-list (although 7/75 (9.3%) provided these for medication variables studied). Of these studies using medication, 10 (76.9%) defined the window in which medication exposure was measured: in 3 (23.1%) studies this was before the study period, 5 (38.5%) during, 1 (7.7%) before or during (117), and 1 (7.7%) required "current" medication (353). Of the studies, eleven (14.7%) repeated their analyses using different gout definitions, in sensitivity analyses. While most studies did not distinguish between incident and prevalent gout, 31 (41.3%) defined incident gout: 29 (93.5%) of these by the first coded appearance and 16 (51.6%) of these required a prior 1-5 years with no diagnosis or medication.

Table 14. Definitions of gout and medication exposure used in the studies (N = 75)

Definitions	N (%)*
Gout Definition	
≥1 diagnosis	34 (45)
≥1 EHR reference (not specified)	2 (3)
≥1 gout medication prescription/dispense	2 (3)
≥1 diagnosis or gout medication	2 (3)
≥1 diagnosis or keyword	2 (3)
≥ keyword search of EHR	1 (1)
≥1 diagnosis; 1 diagnosis and medication (2 definitions)	4 (5)
1 liberal and ≥1 stringent definition (other than above)	6 (8)
≥2 diagnoses	3 (4)
Survey response and ≥1 diagnosis	2 (3)
≥1 diagnosis or medication and coded CKD, urolithiasis, tophus or >2 flares	2 (3)
≥ 1 test	3 (3)

Meet ACR criteria		3 (3)
Other stringent (not seen in >1 study)		9 (12)
No definition given		6 (8)
Incident Gout Definition		33 (44)
First code in the study or EHR	(% of 33)	31 (94)
No diagnosis in prior time period (1-3 y)	(% of 33)	13 (39)
Distinct codes for incident and prevalent	(% of 33)	1 (3)
No diagnosis and/or medication in prior time period	(% of 33)	5 (15)
No definition given	(% of 33)	2 (6)
Medication Minimum Duration/Dose Requirement		8 (11)
Minimum of 6 months		4 (5)
Minimum of 3 consecutive months		1 (1)
Minimum of 1 month		1 (1)
Minimum of 2 prescriptions		1 (1)
≥300 mg/day of allopurinol		1 (1)
Medication Exposure Measure		
Binary 'ever exposed' at any point in the study		23 (31)
Binary 'ever exposed' at a specific time point		14 (19)
Binary 'ever exposed' in a specific time window		9 (12)
Exposure within a window		26 (35)
Continuous exposure		4 (5)
Cumulative exposure		3 (4)
Reporting on Medication Exposure		
Use at baseline or prior to study		35 (47)
Dosage		33 (44)
% 'ever exposed' during the study		29 (39)
Use at or during follow-up periods		19 (25)
Temporal duration of medication use		9 (12)
Use in chronological periods		8 (11)

Note: ACR = American College of Rheumatology; CKD = Chronic kidney disease.
*Percentage is given as n out of 75 unless otherwise specified

Table 15. Distribution of studies according to elements considered in the definition of gout and medication exposure and their classification recording system (N = 75)

Indicator	Count (%)
Gout Diagnosis	
Diagnostic Code	58 (88)

Provision of code-list	(% of 58)	25 (43)
ICD	(% of 58)	33 (57)
Read Code / Oxmis	(% of 58)	18 (31)
Classification not specified	(% of 58)	7 (12)
Medication		13 (20)
Provision of code-list	(% of 13)	0 (0)
Multilex	(% of 13)	5 (39)
BNF	(% of 13)	1 (8)
Classification not specified	(% of 13)	7 (54)
Test Result*		6 (9)
UA crystals in synovial fluid		4
Radiologic evidence, e.g. DECT scan		2
Biopsy of tophus or synovial tissue		1
High SUA level		2
Free text		3 (5)
Medication Exposure		
Medication		75 (100)
Multilex		7 (9)
ATC		7 (9)
National ID		2 (3)
BNF		1 (1)
Classification not specified		58 (77)

Note: *Some studies used multiple tests in defining gout. ATC = Anatomical Therapeutic Chemical; BNF = British National Formulary; DECT = Dual-energy computed tomography; ICD = International Classification of Diseases; SUA = Serum uric acid; UA = Uric acid

4.3.2 Medication Assessment

Forty-five (60.0%) studies used prescriptions, 20 (26.7%) used dispensary data and 3 (4.0%) used both. Twenty-seven gout-related drugs or groups were reported, with ULT the commonest group (N = 72, 96%) and allopurinol, NSAIDs and colchicine the commonest drugs (Table 16). Interleukin-1 antagonist, oxypurinol and adrenocorticotropic hormone were each reported once. Eight (10.7%) studies specified a minimum prescription duration or dose. In 24 studies in the UK primary care setting having a cohort size of ≥100 patients, 16 (66.7%) referred to allopurinol, 15 (62.5%) to NSAIDs and 11 (45.8%) to colchicine.

Table 16. Gout medication types in the studies (N = 75)

Medication Type	Count (%)
Urate lowering therapy	72 (96)
Allopurinol	61 (81)
Febuxostat	16 (21)
Oxypurinol	1 (1)
ULT group	21 (28)
Prophylactic	53 (71)
NSAIDs	42 (56)
Colchicine	41 (55)
Corticosteroid	28 (37)
Other analgesic	9 (12)
Prophylactic group	1 (1)
Probenicid	19 (25)
Benzbromarone	5 (7)
Sulfinpyrazone	6 (8)
Uricosuric drugs	2 (3)
Uricosuric antihypertensive / diuretic	4 (5)
Prednisolone	2 (3)
Pegloticase	2 (3)
Other (medications only used in one study)	4 (5)

Note: NSAID = non-steroidal anti-inflammatory; ULT, urate lowering therapy

4.3.3 Measures of Medication Prescribing

Most studies reported the percentage with medication prescribed ‘ever’ (N = 44, 59%) or within specified windows (N = 26, 34.7%). Twenty-six (34.7%) studies limited the reporting of medication to exposure at baseline or during the study (Figure 11). Seven (9.3%) reported continuous or cumulative exposure, 8 (10.7%) temporal prescribing trends and 32 (42.7%) the prescribed dosage. Sixteen (21.3%) assessed the proportion initiating treatment at diagnosis or in periods of follow-up, 2 prescription gaps and 1 the percentage with ≥60 consecutive days of prescribing. ULT adherence was measured as a medication possession ratio (MPR) or the proportion of days covered (PDC) ≥0.80, in 1 and 8 (10.7%) studies respectively. A further 2 studies reported on gaps in ULT prescribing.

Figure 11. Venn Diagram of the reporting on patient-level medication exposure (N = 75)

4.3.4 Reported Estimates of Medication Prescribing

Gout medication initiation was low although higher in studies with stricter definitions, with little temporal change. Studies that selected incident patients using a liberal gout diagnosis (i.e. having ≥1 gout diagnostic code) reported 6.7% (range 0-9.4%) of patients initiating ULT at diagnosis, 22.9% (range 16.9-25.4%) by 1 year and 40.5% (range 40.3-40.53%) by 10 years (354-356). In studies of incident gout that required ≥1 gout diagnostic code but also prior registration without diagnosis, the estimates of ULT initiation were higher: 15% at diagnosis, 27.7% (range 23-31.9%) by 1 year and 43.1% by 9 years (117, 357-361). Kapetanovic et al. reported 47.8% and 60.6% of patients with incident gout receiving ULT in 2011 when using a liberal and stringent definition respectively (362).

Studies using a liberal gout definition found stable ULT prescribing (mean 28.3%), declining NSAID use (mean 36.3%) and rising colchicine use (mean 6.3%) across 1990

to 2014 among patients with incident and prevalent gout (117, 126, 357, 359, 363, 364). For example, Fisher et al. reported the proportion with medication during 1999-2006 and 2007-2014 respectively: 24.3% and 23.3% ULT; 11.4% and 23.8% colchicine; 34.4% and 30.7% NSAIDs (364). Rothenbacher et al. reported drug use in the first year following diagnosis, 2000-2001 and 2006-2007 respectively: 28.2% and 24.3% ULT, 13.8% and 19.7% colchicine (117). Rai et al. reported the proportion with medication across 2000-2012: stable around 22% ULT; 35.4% to 24.5% NSAIDs (363). By comparison, Arromdee et al. used a stringent definition involving adjudication by a rheumatologist and reported higher colchicine use: 19.8% in 1995-1996 (119).

The duration of ULT prescribing was short, although longer in studies with a stringent gout definition. It ranged from 0.33-0.8 years in studies requiring ≥1 diagnosis (liberal), 1 year in a study by Zandman-Goddard et al. requiring ≥1 rheumatologist-coded diagnosis and 2.5-4.0 years in studies including patients with ≥2 diagnoses (355, 365-369). A study by Mantarro et al., requiring ≥1 diagnosis, highlighted the steepness of drop-off with only 45.9% adherent across 0-29 days following initiation, 16.7% by 89 days and 3.2% 1 year on (370). Scheepers et al. reported mean PDC of 0.57 (standard deviation [SD] ±0.34) for patients with ≥1 diagnosis, while Coburn et al. selected patients with ≥2 diagnoses and reported median PDC 0.7 to 0.83 (355, 371). The proportion with PDC ≥80 in the first year was 38.6% among patients with ≥1 diagnosis (359, 370) and 59% among patients with ≥1 diagnosis validated through a survey (372). Hughes et al. reported 72% adherence (MPR ≥80) among cases with ≥1 diagnosis but 37% had MPR >100% (373). Other studies requiring ≥1 diagnosis reported 75% having a prescription gap of ≥25 or ≥60 days and only 13% of patients having ≥6 months of ULT-prescribing in the first year from diagnosis (358, 367, 374). Chronic prescribing of an acute medication, colchicine, following allopurinol initiation was also reported (median duration 0.72 years, IQR = 0. 26-1.43) (368). Only one study assessed temporal variation in ULT adherence, reporting stable non-adherence but increasing adherence (defined as PDC ≥80) over partial adherence from 1997-2012 (359).

ULT doses were low, with limited up-titration, although doses were greater in studies with a stringent gout definition. The mean proportion of patients with ULT up-titration was 5.4% (range 4-36%) and 29.0% (range 22.4-39.3%) when using liberal and stringent definitions respectively (60, 156, 366, 368, 373, 375). For example, Hughes et al. reported 28.3% up-titration within 90 days of an SUA test showing an above-goal level and that Rheumatologists more commonly up-titrated (373). The mean initial allopurinol dose was 148.1 mg/day and overall dose was 223.3 mg/day: 194.1 mg/day

and 231.4 mg/day in studies with liberal and stringent definitions respectively (352, 353, 365, 375-380). Inappropriate allopurinol dosing for renal disease patients was high (mean 24.8%, range 22-25.9%) and only reported in studies using liberal definitions (150, 381).

Some studies using liberal gout definitions reported demographic variation in medication prescribing, although this was not found in studies using stringent gout definitions. Allopurinol use was reported as more common in women than men in 2 studies using a liberal definition (AHR: 1.41), while comparable durations of allopurinol adherence were reported in men and women (0.98 and 1.04 years respectively) in a study using a stringent gout definition (117, 356, 369). Ethnic variation in the duration of allopurinol adherence was assessed using liberal and stringent gout definitions: Wahedduddin et al. reported variation (2.5 and 4 years among Hmong and Caucasians) using a liberal definition and no difference using a stringent definition (365).

4.3.5 Measures of Treatment Outcomes

Nine (12.0%) studies measured changes in SUA level with ULT: 7 measured mean change and 4 the percentage achieving SUA level goal. Eight (10.7%) examined the impact of ULT on disease control or SUA level, with 5 assessing associations with the starting dose, titration and drug adherence. Other evaluated measures were the percentage of patients reaching the SUA goal or switching treatment, the mean SUA change per treatment group, comparison of changes in repeated-measures and the time to treatment response. Fifteen (20%) evaluated treatment safety and the percentage with adverse outcomes, with 9 determining effect on the risk of adverse events: fracture, joint replacement, mortality, myopathy, CKD, hepatoxicity or CVD.

To model the treatment outcomes, 7 studies used logistic regression to assess the impact of ULT on disease control or SUA level and 2 used linear regression. Five used logistic regression to assess SUA level in association with the starting dose, titration and drug adherence. Nine utilized Cox regression to determine treatment effect on risk of adverse events or late effects. Models used in examining other outcomes were univariate analysis, logistic regression, competing risks regression and Kaplan-Meier.

4.3.6 Reported Estimates of Treatment Outcomes

Studies reported on the effectiveness of gout therapy. Higher SUA level goal attainment was reported in patients with higher starting allopurinol doses, adherence and up-titration, more so in studies with stringent gout definitions (366, 370-372, 378). For example, Mantarro et al. used a liberal definition and reported lower odds of hyperuricaemia in the first 90-149 days for adherent patients (adjusted odds ratio [AOR] = 0.40; 95%CI 0.24-0.67) (370). Rashid et al reported higher starting doses (AOR = 1.92; 95% CI 1.86-2.22 for 100-300 mg compared with ≤100 mg) and adherence (AOR = 2.52; 95% CI 2.41-3.01) associating with goal achievement in patients with ≥2 diagnoses (366). ULT had a positive dosage- and duration-dependent effect on SUA level (7/7 studies), and combination therapy was more effective than monotherapy (2/2 studies), regardless of gout definition (156, 366, 370-372, 377, 378, 380, 382). For example, Rashid et al. reported that 73.5% of patients with 0 flares in the year following ULT initiation were adherent, compared to 28.2% of patients with ≥3 flares (156). Hatoum et al. reported greater SUA level goal attainment with febuxostat than allopurinol (AOR=1.73; 95% CI, 1.48-2.01) but 5.4 times more treatment switching (380). In patients hospitalised with gout, Theuringer et al. reported resolution of the gout episode, or marked improvement, in all patients (N = 13) prescribed with anakinra: 50% obtaining this within 24 hours (383).

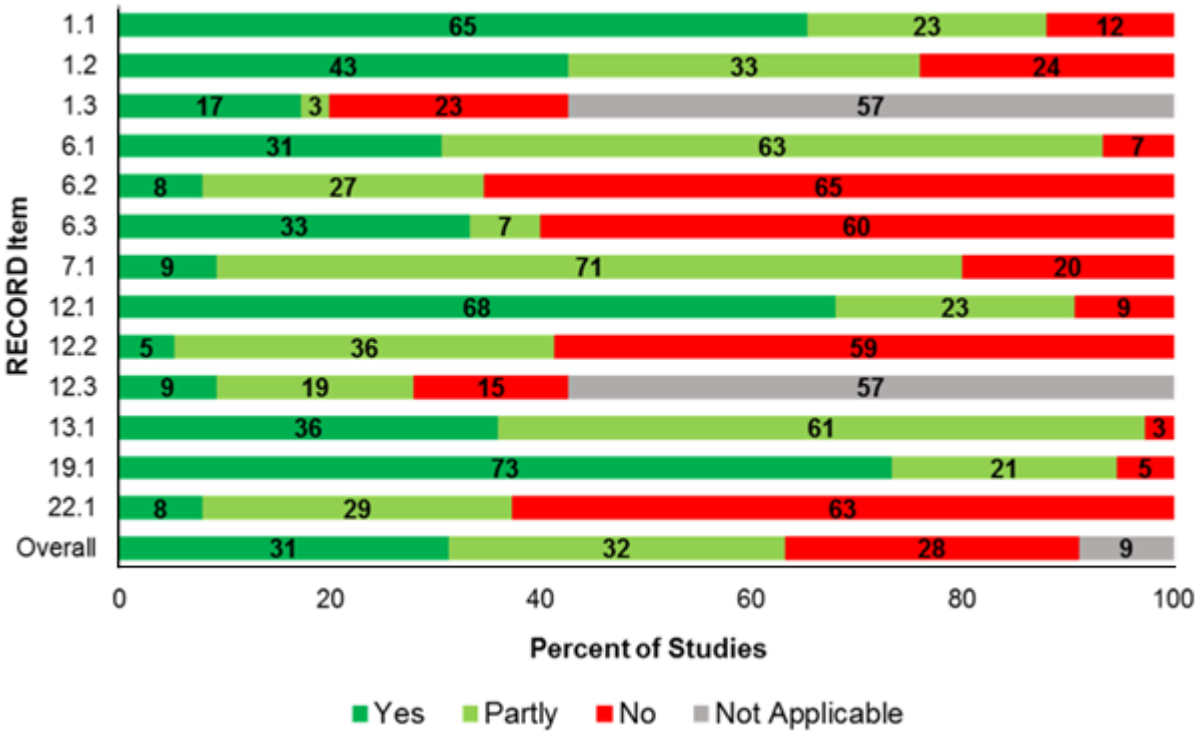
Most studies reported that gout treatment was safe and well tolerated, with few switches, regardless of gout definition. For example, with a liberal definition, colchicine associated with lower cardiovascular risk and mortality compared with no medication; and with a stringent definition it was unrelated to myopathy risk when prescribed with vs without statin (384-386). Spaetgens et al., using a liberal gout definition, reported that current colchicine use associated with no increased pneumonia risk and slightly raised risk of urinary tract infection (387). Using a liberal definition, studies reported raised risk of pneumonia with current allopurinol use and osteoporotic fracture with high doses, no or minor fracture risk and no joint replacement risk associated with ULT (387-390). Only the association between allopurinol and all-cause mortality was examined using varied gout definitions (in studies that both used propensity-matched cohorts): Kuo et al. used ≥1 diagnosis and reported no improvement compared with non-exposed gout patients (hazards ratio [HR] = 1.01; 95% CI, 0.92-1.09); Coburn et al used ≥2 diagnoses and noted no improvement with titration (HR = 1.08; 95% CI, 1.01-1.17) (371, 391). Both studies were hindered in investigating dosage-dependent effects by pervasive low dosage prescribing.

4.3.7 Comprehensiveness of Reporting

The overall mean score for CoR of EHR data use was 5.2/10 (SD ±1.5, range 1.8-8.5): 5.6/10 (SD ±1.3), 5.3/10 (SD ±1.4) and 3.0/10 (SD ±0.9) for studies with a liberal, stringent and no stated approach to defining gout respectively. Table A 3 in Appendix A shows the definition and mean score per study. Those with a liberal definition reported less comprehensively on validation, database population and linkage methodology but more frequently provided full code-lists and the patient count at each selection stage.

In the title or abstract, 66 (88.0%) mentioned ("yes"/"partly") the data (although 11 only named the database) while 15 (20.0%) reported the geographic region, 9 (12.0%) the study timeframe and 32 (42.7%) both (Figure 12, Table 17). Of 32 (42.7%) data-linkage studies, 15 (46.9%) mentioned linkage in the title/abstract and 21 (65.6%) mentioned the method, quality or level of linkage in the methods. Study titles commonly referenced gout, medication, setting, country and study design. The commonest words in the study titles included 'cohort', 'risk', 'GP', 'allopurinol' and 'population-based' (Figure 13).

Figure 12. Percentage of studies with comprehensive reporting on RECORD items (N = 75)



Note: RECORD = REporting of studies Conducted using Observational Routinely-collected Data

Table 17. Frequency of studies with comprehensive reporting on RECORD items and additional relevant items (N = 75) (345)

RECORD	Question	Count (%)			
		Yes	Partly	No	NA
1.1	Is the type of data used specified in the title or abstract (with database name if applicable)?	49 (65)	17 (23)	9 (12)	
1.2	Is the geographic region and timeframe within which the study took place reported in the title or abstract?	32 (43)	25 (33)	18 (24)	
1.3	If databases were linked as part of the study, is this stated in the title or abstract?	13 (17)	2 (3)	17 (23)	43 (57)
6.1	In the methods, is the method of study cohort selection (e.g. codes or algorithm used to identify subjects) listed in detail, or an explanation as to why this is not shared?	23 (31)	47 (63)	5 (7)	
6.2	In the methods, was any validation conducted during study published?	6 (8)	20 (27)	49 (65)	
6.3	Is there a flow diagram / graph with the number of individuals in the data at each stage (e.g. of linkage or cohort selection)?	25 (33)	5 (7)	45 (60)	
7.1	Are codes or algorithms provided for all exposures, outcomes, confounders and effect modifiers, or an explanation as to why this is not shared?	7 (9)	53 (71)	15 (20)	
12.1	Do authors describe the extent to which they had access to the database population used to create the study creation?	51 (68)	17 (23)	7 (9)	
12.2	Is information provided on the data cleaning methods?	4 (5)	27 (36)	44 (59)	
12.3	If databases were linked as part of the study, are methods of linkage and linkage quality evaluation provided, and is linkage at the person, organization or other level?	7 (9)	14 (19)	11 (15)	43 (57)
13.1	In the results, is the cohort selection described in detail, including filtering based on data quality, data availability and linkage? (text or diagram)	27 (36)	46 (61)	2 (3)	
19.1	In the discussion, is there discussion about the implications of using data not primarily collected for the study? For example, discussion of misclassification bias, unmeasured confounding, changing eligibility over time	55 (73)	16 (21)	4 (5)	
22.1	Authors provide information on how to access any supplemental information such as the study protocol, raw data, or programming code	6 (8)	22 (29)	47 (63)	
NA	Study observation period, start and end dates (month, year)	47 (63)	25 (33)	3 (4)	
NA	Count of sites contributing to the source database	39 (52)		36 (48)	
NA	Count of patients in the source database	31 (41)		44 (59)	

Note: RECORD = REporting of studies Conducted using Observational Routinely-collected Data; NA = Not applicable

Figure 13. Word cloud of the top 90 words* in study titles (N = 75)

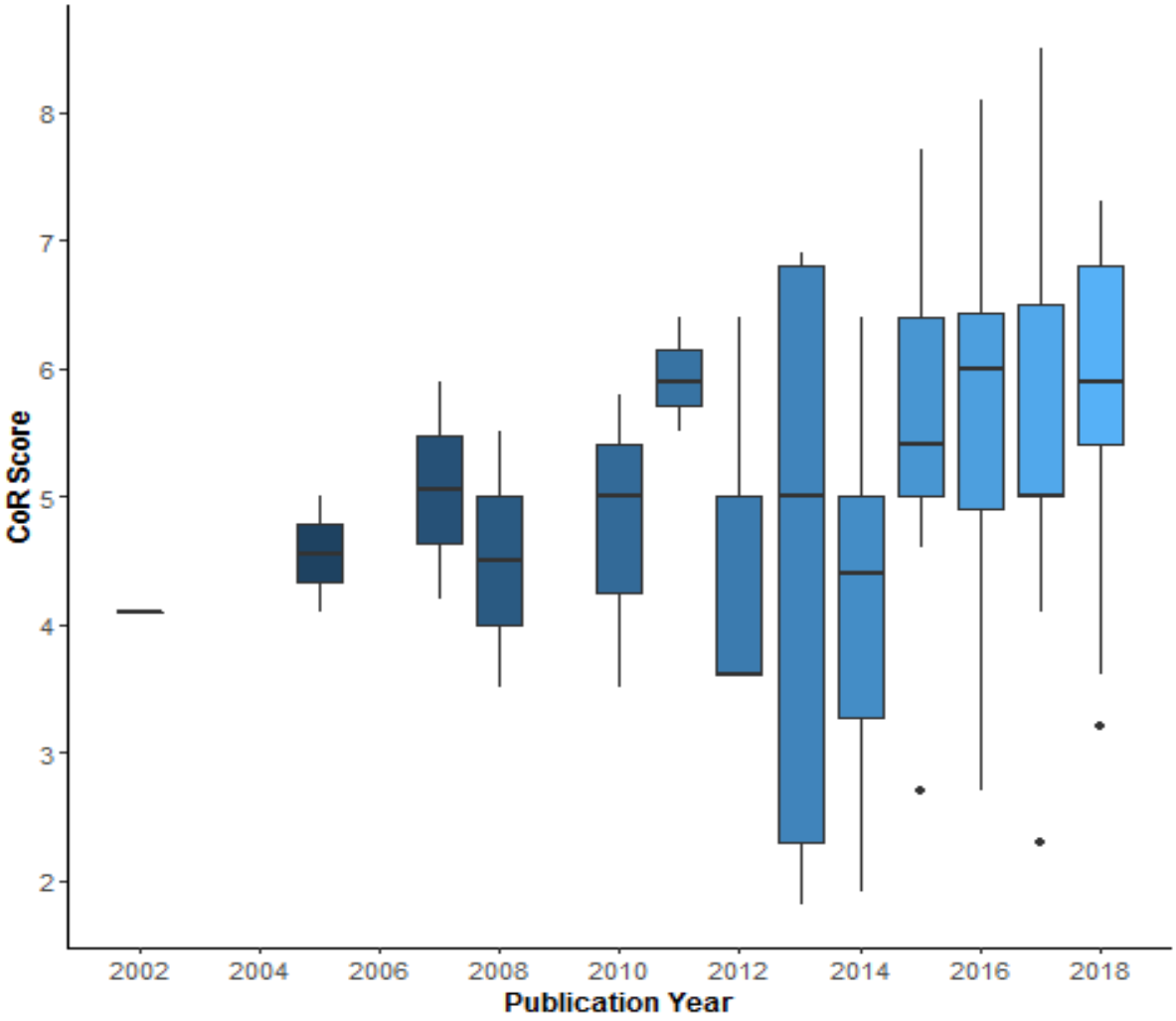
Note: *excluding basic words e.g. 'the'

In the methods, cohort selection was the most comprehensively recorded item ("yes"/"partly" = 70, 93.3%), while only 7 (9.3%) provided the codes or algorithms for all variables, 6 (8.0%) provided a data availability statement and only 4 (5.3%) made ≥2 references to data cleaning or preparation. Seventeen (22.7%) provided incorrect references to a gout diagnosis validation study: 7 referenced a study by Meier et al. (392) that used more stringent selection criteria than that used by the citing study, 5 referenced studies using a different coding system to that used in the citing study and 3 referenced studies that validated a disease other than gout. The study by Meier et al used ≥1 diagnosis plus a high SUA level test result or subsequent anti-gout medication, yet the reported diagnostic validity was cited by subsequent studies using only ≥1 diagnosis. Fifteen studies used data from the CPRD, which trains clinicians in data-entry (89), but only 8 (53.3%) noted this. Fifty-five (73%) studies discussed ≥2 implications or limitations of secondary data use.

Overall, CoR scores rose over time (Figure 14). The median score was 5.0 (IQR = 3.6-5.9) for 40 papers published by 2015 (date of RECORD guideline publication), and 5.8 (IQR = 5.0-6.5) from 2016 (n = 35). This rise following RECORD publication was not significant ($x^2 = 8.2$, $p = 0.3$). The proportion scoring "yes" or "partly" was greater for all items from 2016 except 13.1, with a mean change of +15.5% (SD ±16.1). This ranged from 0.4% decrease in studies describing the selection of included persons, to 56.3%

increase in those describing the linkage methodology. However, from 2016, “yes” scores decreased for items 1.1, 6.1, 6.3, 19.1; notably 15.4% fewer mentioning the data type in the abstract and 14.3% fewer discussing the limitations of secondary data use. The proportion with “yes” scores remained ≤25% for 6/13 items (6.2-7.1, 12.2, 12.3, 22.1).

Figure 14. Boxplot of overall CoR scores for studies by publication year (N = 74)



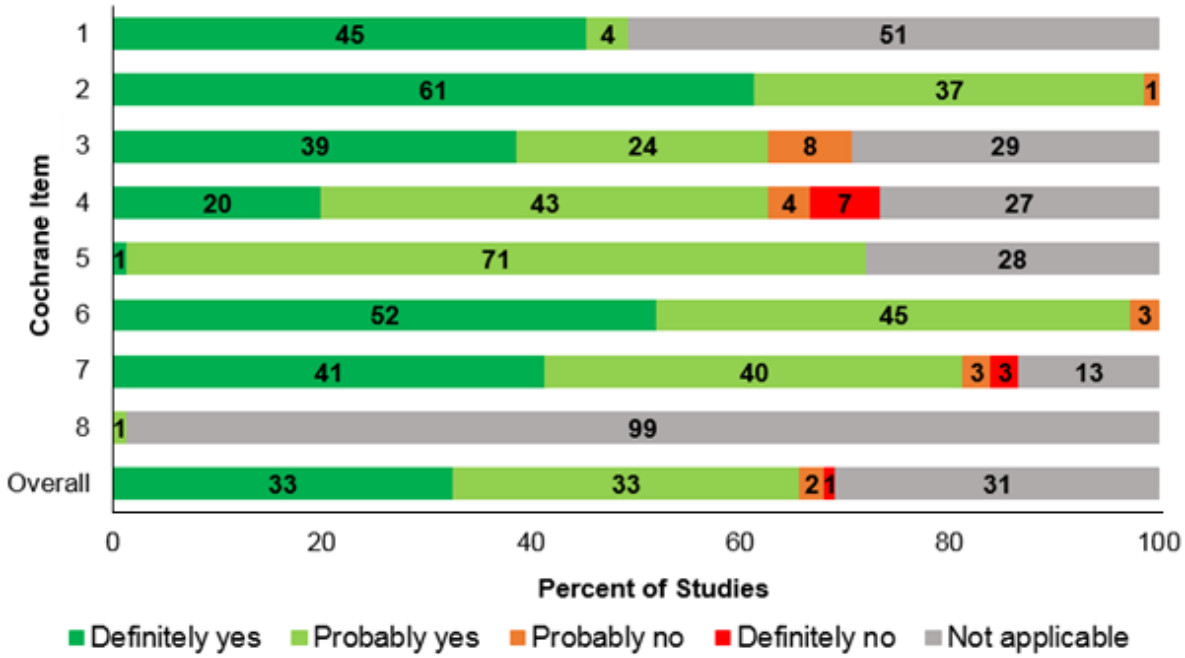
Note: Horizontal lines are medians and interquartile ranges (25th and 75th percentiles); whiskers’ ends indicate the maximum and minimum values at most 1.5 times the interquartile range from the hinge; dark individual dots are outlier values.

4.3.8 Risk of Bias

The mean RoB score was 12.0/15 (SD ±1.6, range 7.5-15) and, where items were applicable to a study, over 85% of studies were scored with low or probably low RoB per item. Table A 3 in Appendix A shows the definition and mean score per study. All studies scored “yes”, “probably yes” or “not applicable” for having exposed and non-

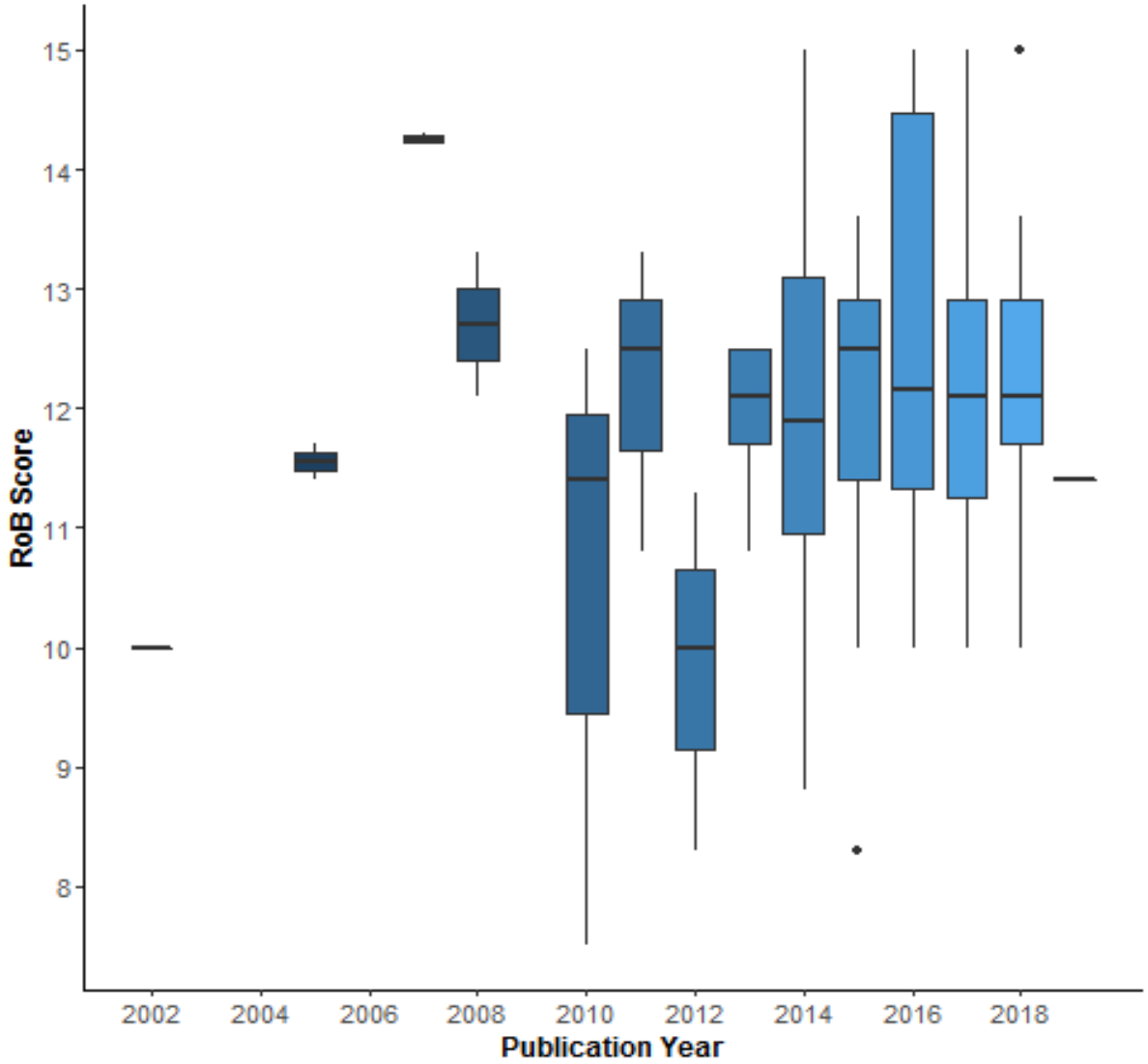
exposed cohorts drawn from the same population and confidence in the assessment of the presence or absence of prognostic factors (Figure 15). In 8 (10.7%) studies, patients were inappropriately matched or estimates were incorrectly adjusted and 6 (8.0%) inadequately assessed the outcome at follow-up start. However, the RoB measures were not applicable to, on average, 30.8% of studies per item (IQR = 10.0-34.7%). The similarity of co-interventions between groups compared, and whether cohorts were drawn from the same population, were non-applicable for 74 (98.7%) and 38 (50.7%) studies respectively.

Figure 15. Percentage of studies with low risk of bias, as assessed with the Cochrane Tool for Cohort Studies (N = 75)



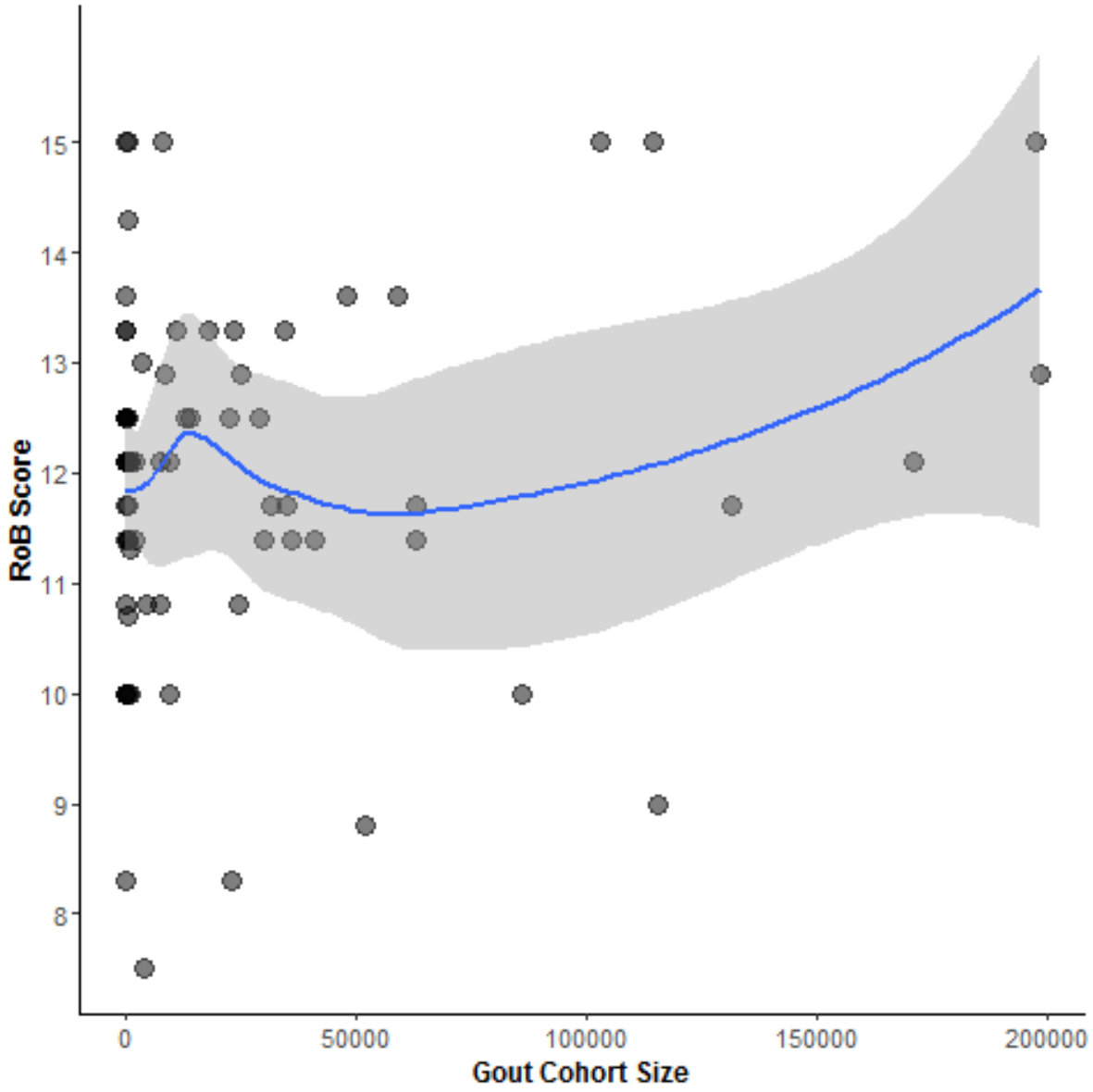
Mean RoB scores were 12.0 (SD ±1.3), 11.8 (SD ±1.8) and 12.6 (SD ±1.9) for studies with a liberal, stringent and unspecified gout definition. The overall RoB score did not change over time (e.g. mean 11.7 (SD ±1.7) for studies published ≤2015 and 12.3 (SD ±1.5) for studies published >2015) (Figure 16). There was a non-significant association with cohort size (e.g. mean 11.7 (SD ±1.9) for studies with a cohort size <1,000, compared with 12.0 (SD ±1.2) for larger studies) (Figure 17).

Figure 16. Boxplot of overall RoB scores for studies by publication year (n = 74)



Note: Horizontal lines are medians and interquartile ranges (25th and 75th percentiles); whiskers' ends indicate the maximum and minimum values at most 1.5 times the interquartile range from the hinge; dark individual dots are outlier values.

Figure 17. Scatterplot of overall RoB scores for studies by cohort size (n = 75)



Note: The blue line is the smooth local weighted regression line (LOESS curve). The shaded area indicates the 95% confidence interval.

4.4 Discussion

This is the first systematic literature review of variation in methods, results and reporting in EHR-based studies of gout medication. The studies demonstrated wide variation in gout diagnostic definitions and medication-related methods. This did not widely affect efficacy and safety estimates, though reporting quality was variable. The risk of bias in the studies was acceptable, as assessed by the Cochrane tool.

4.4.1 Definition and Methods

Variation in gout diagnostic definition and medication-related methods may explain the differences in medication utilisation and treatment efficacy estimates across the studies. Most studies employed more liberal diagnostic approaches that may lead to more misdiagnosis (false positives) when suspected cases are recorded in EHRs during pre-diagnostic evaluation. Stringent approaches are likely to have higher specificity but may lack sensitivity for identifying mild or recently diagnosed cases (i.e. severe or long-standing cases benefit from additional clinical contact, increasing opportunities for prescribing, tests and diagnostic coding). There were varied approaches to measuring adherence; MPR seemed less suitable than PDC because scoring >100% MPR was common, potentially due to dose changes and early fills (373).

Most studies measured variables that are well recorded in EHRs: SUA level, medication, comorbidities and procedures. No studies considered the time since diagnosis, and the assessment of severity or flare frequency is difficult in EHR data (367). The medications studied were varied, but studies largely reported on the proportion 'ever exposed' to medication, with less consideration of detail such as prescription duration, cumulative dosage and timing in relation to clinical events such as testing.

4.4.2 Reported Estimates

Based on EHR prescription and dispensary data, reported ULT initiation, adherence and titration was sub-optimal and doses rarely reached above 300 mg/day (note: the reported low initiating doses were appropriate). Despite the publication of guidelines, ULT prescribing remained stable with poor adherence and gaps or discontinuation in prescribing were common after only a short duration of prescribing. Only one study showed temporal improvement, and only in full over partial ULT adherence, rather than in a reduction of non-adherence (359). The limited improvement is of concern given that the studies comprehensively reported on prescriptions and trial evidence suggests that monotherapy doses ≤300 mg/day fail to reduce SUA levels (33). Prescribing for patients with renal insufficiency, and chronic prescribing of prophylactics alongside ULT, also require particular attention. Encouragingly, interventions that showed promise for improved pharmacologic management and reduced urate levels included nurse- or pharmacist-led follow-up and home testing of ULT levels (393-395). If

prescribing were to improve, EHR-based research could evaluate prescribing patterns, dosage-dependent safety and efficacy, and safety of concomitant prescribing.

Estimates of medication prescribing and effectiveness were higher and less heterogeneous in studies using stringent gout definitions. These studies may select more severe and long-standing cases with greater opportunity for prescribing, tests and diagnostic coding, so gout management may be more critical in these instances. ULT was generally reported as safe and effective regardless of definition or methodology, which indicates the opportunity for optimizing gout control though ULT.

Estimates of gout incidence were not evaluated but it was noted that 48.4% of studies did not specify a prior disease-free period in determining incidence. Lewis et al. reported that this leads to over-estimation of incidence of similar chronic diseases such as rheumatoid arthritis in the first 12 months of a study (396). This is due to inclusion of prevalent cases diagnosed during previous registrations at other GP practices or incorrectly recorded as diagnosis instead of as medical history of the disease during registration.

4.4.3 Comprehensiveness of Reporting and Risk of Bias

There is significant scope for further improving the reporting of EHR-based research to facilitate reproducibility and understanding of bias and representativeness. Improvement could be made in reporting of the timeframe and data linkages in the abstract; sharing code-lists (e.g. in supplemental material or publically available repositories); adequately describing definitions, validation and linkage; and reporting the cohort size during each selection stage. Indeed, more studies failed to define their gout definition than provided information on data availability.

There were reporting differences between studies using liberal and stringent gout definitions. The former more commonly provided code-lists for definitions and the cohort count at each selection stage, though this may be an artefact of the fewer selection stages. The latter reported more on validation, database population and linkage, which may be explained by a greater understanding of how healthcare provision determines data collection and the use of this knowledge to develop a more systematic approach to cohort design and study definitions.

The publication of RECORD guidelines in 2015 may account for some of the observed temporal improvements in CoR, though no studies referenced RECORD. The recent shifts, from full to partial reporting of the data type in the abstract (with increasing reference to the database source instead of data type), and toward discussion of the implications of secondary data use, may reflect the increasing familiarity of EHR data among researchers. Falling levels of reporting of the cohort size during patient selection, together with the reported rising mean cohort size, may reflect increasing use of centralised databases and modalities of data sharing.

There were difficulties in assessing certain RECORD items. For example, researchers rarely have full access to the EHR database population, yet no studies reported whether they had full data access or received data via a third party (RECORD item 12.1). We therefore scored for this item instead by whether studies described and appropriately referenced the database (profile, coverage or validation studies), which enabled consideration of selection bias. Another issue was that the RECORD items were listed by the area of the manuscript (e.g. abstract, methods) in which they should appear. For items listed in the methods and results, we awarded a score if they were reported in either section. RECORD item 6.3 considers a graphical display of the count of individuals in each linkage stage; we adapted this to apply to all (i.e. also non-linkage) studies by considering the display of individuals in each selection stage. RECORD did not capture some aspects relevant to CoR, which were found to be inadequately reported, including the study observation period (start and end dates) and the number of contributing sites and patients in the source database.

Scores of RoB were generally low, especially in recent studies. However, the Cochrane items assessed were “non-applicable” in a third of instances (e.g. those related to co-interventions or estimate adjustment), because many EHR-based studies were descriptive.

4.4.4 Assessment Tools

During the development of the data extraction form used in this study, published guidance was tailored to suit data extraction on EHR-based studies. The recommendations from the CRD that related to aspects such as recruitment and resource use, were relevant to more traditional study designs (337). These were adapted to ensure that the details relevant to EHR-based studies were captured, such as the clinical setting, data type and database population coverage. The adaptations

performed suggest the need for further guidance to be published on extracting information from EHR-based studies.

Some factors relevant to CoR and RoB were difficult for researchers to account for. These include consideration of temporal changes in code classifications, completeness and accuracy of recording, or local / regional / national guidelines or policymaking that alter clinical practice or EHR utilisation. Data providers, particularly those providing 'research-ready datasets', should publically detail, with regular updates, the steps undertaken to create a dataset, database profiles and results of data quality assessments (particularly data completeness, correctness, concordance, plausibility and currency) (71). This would facilitate reporting, calibration and the capacity to relate findings back to EHRs for personalised interventions.

Other items not considered by commonly used CoR and RoB assessment tools are pertinent to EHR-based research. Researchers should consider the adequacy of the dataset to answer the research question (e.g. whether to use hospital or primary care data in studying diabetes). Where possible, temporal changes in clinical practice or EHR use during the study period should be accounted for in study design, analysis and interpretation. The use of multi-site data is a strength of EHR-based research but may bring site-level bias if appropriate statistical methods (e.g. random-effects models or use of a site indicator) are not used and this is not specifically covered by common RoB tools for cohort studies. In this review, no multicentre studies reported using these methods. We looked for discussion of unmeasured confounding, selection bias and changing eligibility over time when assessing reporting of the limitations of secondary data use (RECORD 19.1). However these are not assessed by Cochrane and only the former two are assessed by the Risk of Bias in Non-randomised Studies – of Interventions (ROBINS-I) and Newcastle-Ottawa tools (346, 397, 398). Acquisition bias, where events occurring outside the study window (e.g. diagnoses or prescriptions) affect estimates, or variation in patient engagement with clinical care, were not discussed by any study, nor assessed by any of these tools. Unequal follow-up duration (data window length), inappropriate conditioning, competing risks or loss-to-follow-up due to patient- or EHR system- migration can introduce bias but these were not specifically considered in the tools or in all studies. For example, one study reported on prescribing for incident gout cases in 2014, where diagnoses made later in 2014 had less follow-up (375). Even studies with large datasets should consider sparse data bias where modelling multiple variable combinations (399). Given the rising number of EHR-based studies, such considerations should inform future CoR and RoB tool development (Table 18).

Table 18. Commonly missed factors that affect EHR-based research, with considerations for further improving CoR and RoB tools

Factor	CoR Consideration	RoB Consideration
Temporal changes in code classification, EHR system, clinical practice, guidelines or policy	Are these reported on in longitudinal studies?	Are these temporal changes appropriately taken into account (e.g. through adjustment) and/or their impact examined through sensitivity analyses in longitudinal studies?
EHR data accuracy, adequacy (e.g. detail) and completeness (including missingness)	Are these reported and previous validation studies referenced correctly?	Is the research question and analysis appropriate, given these?
Steps applied and assumptions made during data extraction, processing and cleaning	Are these reported or referenced correctly?	Is the research question and analysis appropriate, given these?
Site-level bias		Is this appropriately addressed in multicenter studies? E.g. include site-level in the model
Unmeasured confounding, misclassification bias, selection bias, changing eligibility over time		Are these appropriately addressed or acknowledged? E.g. replication of analysis with different definitions
Bias from unequal follow-up duration		Are longitudinal studies accounting for follow-up duration? E.g. standardization or minimum follow-up requirement, use of survival methods, use of time-variant variables
Bias from competing risks		Are these appropriately addressed in survival analysis?
Bias from change in the population structure, e.g. changes in sites providing data in open cohort studies of long duration	Is description of the population structure (size, demographics) reported over time in longitudinal studies?	Are these appropriately addressed in longitudinal studies?

4.4.5 Strengths and Limitations

A strength of the systematic review was the attempt made to enhance the sensitivity of the literature selection process. A wide breadth of synonyms and MeSH were used for each search term, and a conservative screening approach was adopted whereby abstracts were included for full-text review if their eligibility was uncertain. This review

comprehensively evaluated all aspects of EHR data-use in gout management research (methodology, use and outcomes, reporting and study quality) and the use of common tools for bias risk and reporting assessment. Examples exist in the literature that assessed only CoR, RoB or neither. Aspects of CoR beyond RECORD were also assessed, and a few RECORD items were adapted to enhance applicability. Other commonly used assessment tools were considered and referenced in the Discussion. Protocols and data extraction forms were used by both reviewers, with review of differences, to reduce variation. The study followed the PRISMA statement (336), which strengthened the comprehensiveness of reporting in this study.

The limitations included restriction to publications in English, the risk of publication bias in academic research, and the lack of a standardised term for EHR-based research (e.g. studies may only name a source database or allude to “records”). Due to the EHR focus of the review, we adopted a previously published approach to exclude studies not referencing an EHR-based source even though insurance and claims data could have been EHR-derived (321). No comparison of medication use estimates was made between studies that had medication use as a selection criterion, though such studies may have greater estimates due to excluding non-users. However, studies generally reported on more medication types than those selected on and up-titration measures would not be affected. The minor adaptation of RECORD may affect comparison with other reviews using RECORD. The Cochrane tool for RoB in cohort studies was appropriate for reviewing the predominantly cohort studies although unmeasured confounding and selection bias would have been assessed by ROBINS-I and Newcastle-Ottawa tools.

4.4.6 Conclusion

The number of EHR-based gout medication studies conducted has risen over time. The studies used varied case-definitions and medication-related methodology, which affected the ability to evaluate and compare treatment outcomes. Nevertheless, they consistently reported that ULT is effective, safe and sub-optimally prescribed. There was some evidence of temporal improvement in CoR but not RoB, which was generally low. There is scope for researchers to further improve reporting of methods for reproducibility, particularly through provision of code-lists, data preparation steps and coding validation. Adapted CoR and RoB tools are required for improved evaluation of EHR-based research.

Chapter 5 Ankylosing Spondylitis: An Initial Thematic Scoping Literature Review of Electronic Health Record-based Research

5.1 Introduction

This chapter informs the first thesis objective of describing the existing EHR-based studies. As described in Chapter 2, AS is an inflammatory disease that begins in early adulthood and may eventually cause fusing of vertebral and sacroiliac joints, which can severely hinder quality of life and productivity (163, 172). In more recent years, a diagnostic delay has been highlighted, with implications for disease activity, and prompt treatment with biologics in early AS has been found to be particularly effective (178, 240). As also noted in Chapter 2, EHR data has been used to investigate comorbidities and EAMs in AS, and the pharmacologic management of AS. However, the themes and key findings in the existing literature has not been reviewed. This chapter therefore aims to describe the study themes and findings of EHR-based studies of AS.

5.2 Methods

5.2.1 Literature Search

A protocol was defined for database searches and for selecting EHR-based studies of AS. The search terms were 'AS' and 'EHR' and the synonyms and MeSH were informed by the knowledge base in Chapter 2 and the terms used in Chapter 4 (Table 19). As described in Chapter 4, in database searches the synonyms and MeSH for a term were combined with 'OR' and then the terms combined with 'AND'.

Table 19. Search terms and synonyms used in PubMed for the AS EHR search, with count of returned citations (individually and in combination)

Search Term	Return
Ankylosing spondylitis (AS)	17,200
Ankylosing spondylitis+	16,812
Ankylosing spondylarthritis	16,846
Ankylosing spondyloarthropy	17,127
Ankylosing spondyloarthropic	1

Ankylosing spondyloarthritis+	16,833
Bekhterev's disease	16,823
Marie–Strümpell arthritis	22
Ankylosing Spondylarthritides	16,813
Marie-Struempell Disease	16,813
Bechterews Disease	16,813
Bechterew Disease	16,824
Ankylosing Spondyloarthritides	16,815
Bechterew's Disease	16,853
Marie Struempell Disease	16,813
Marie Struempell arthritis	11
Electronic health record (EHR)	1,024,016
(Health OR patient OR clinic* OR medic* OR care) AND (compute* OR system OR electronic OR warehouse OR link* OR dataset OR network OR database)	994,547
Medical records systems+	36,083
Electronic health record+	33,861
eHealth	26,738
EPR [electronic patient record]	19,826
Record-linkage	6,518
EMR [electronic medical record]	5,855
EHR	4,705
CPRD	324
e-health	2,625
Datalink	977
Routin* n5 data	7
(Electronic OR link* OR compute* OR anonymi*) n5 record	2
EHR and AS	674
Limit to humans	589
Since 1970	586
English language	586

Note: MeSH terms are indicated by '+' and a wildcard by '*'

The literature search was performed on 30 October 2017, using PubMed and Google Scholar. In PubMed, MeSH terms were used and filters were applied to exclude citations published prior to 1970, of non-English language, or with non-human subjects. Given the constraints in Google Scholar noted in Chapter 4, Harzing's Publish or

Perish was used to search the terms “ankylosing spondylitis” and “electronic health record” and download the first 1000 citations published between 01 January 2000 and 30 October 2017 (316). The returned citations were uploaded into Endnote and then Rayyan, with de-duplication performed in each, as described in Chapter 3.

5.2.2 Study Selection

In addition to the selection criteria stated in Chapter 3, an inclusion criteria was that studies must report on AS in adults (aged >17). Registry and health insurance studies were included owing to a small number of AS studies specifically referencing EHR data and the similarity in methodology between these studies. One reviewer (SSRC) performed the study selection in Rayyan using these selection criteria (318). Taking a conservative approach, manuscripts with uncertain relevance were marked for inclusion during title-abstract screening. A full English manuscript was found for 96% of cases. Excluded full-text manuscripts were assigned a pre-defined exclusion label based on the exclusion criteria.

5.2.3 Data Extraction

A data extraction form was created using Microsoft Excel 2013 and information was extracted for each study regarding the title, authors, year of publication, country, study timeframe, cohort size, data source, study objective/s, methods / measures, AS medications studied, key AS-related findings. Study limitations and opportunities for further work were also noted. The form was based on the recommendations of the CRD regarding general information, study and participant characteristics, setting and results (337). The data was abstracted by SSRC and descriptive statistics were performed using Microsoft Excel 2013.

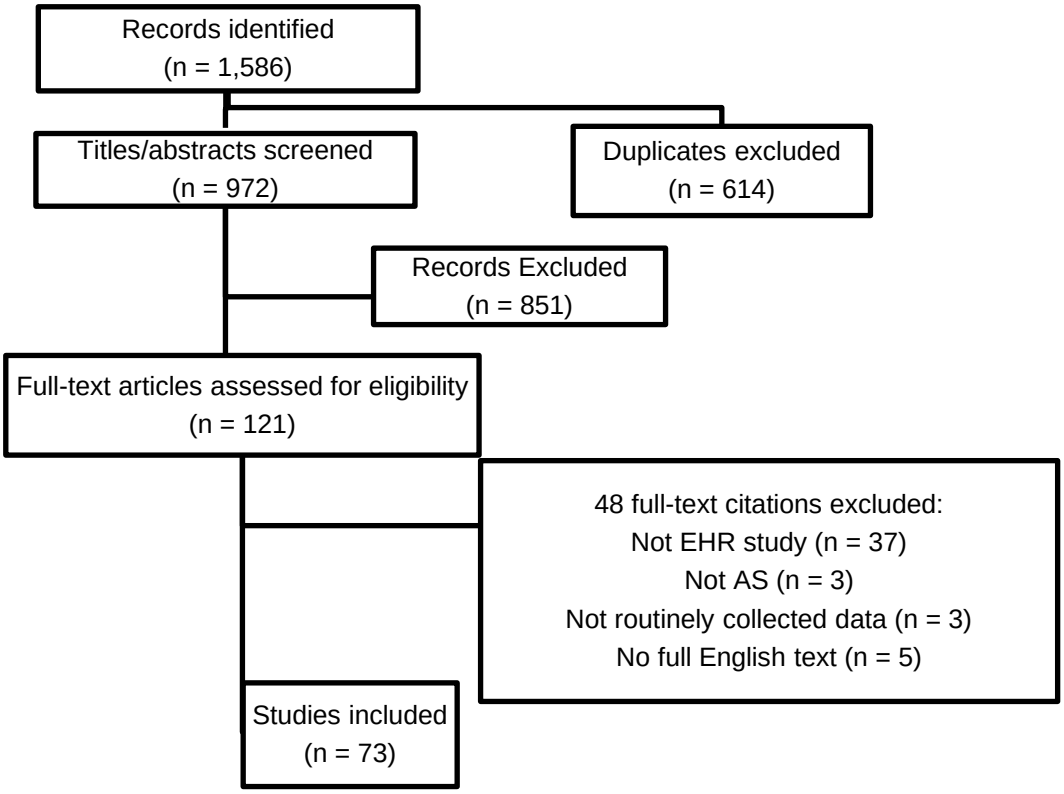
5.2.4 Study Outcomes

The primary outcome was to describe the areas of investigation (study themes) in EHR-based studies of AS and key findings. A secondary outcome was to identify the sources of UK EHR data used in these studies (e.g. GP, hospital), given the focus in this thesis on UK EHR data.

5.3 Results

From 972 screened articles, 121 full-text articles were reviewed and 73 met the study criteria (Figure 18). The articles were published between 1979 and 2017 and the publication rate increased over time as described in Chapter 4 (e.g. 58% were published in the last 5 years, 2013-2017). The median time between study end and publication was 4 years (IQR = 3-5). Only 13 (18%) studies used GP EHR data.

Figure 18. Flowchart of AS EHR study identification and selection



5.3.1 Studies in the UK

Among 15 UK studies, 5 used primary care data (4 of these CPRD, 1 THIN), 5 used GP data linked to survey and/or hospital data in a geographic sub-region of the UK (e.g. Oxford), 4 used HES and ONS death registry data and 1 used hospital data (Table 20). The median AS cohort size was 482 (IQR = 149-3736).

Table 20. Summary of information extracted on UK studies (N = 15)

Theme	Title; publication year (reference)	Objective	Key AS finding	Country	AS therapy	Data source	AS patients	Start	End	Gaps	Methods
AS phenotyping	No increased rate of acute myocardial infarction or stroke among patients with ankylosing spondylitis-a retrospective cohort study using routine data; 2012 (400)	AMI and stroke risk among patients with versus without AS	No increase in AMI or stroke in AS patients, though higher prevalence of diabetes and hypertension (HR 1.27, 1.65). 12% patients with AS diagnosis in GP had a different diagnosis in Rheumatology and 24% with AS in Rheumatology had no AS code in GP data.	Wales	No	SAIL: GP and hospital admissions	1,686	1999	2010	Does not assess impact of NSAID use on hypertension, AMI and stroke	Cox regression for hazards, logistic regression for prevalence, sensitivity analysis using those with AS coded in GP plus another database
AS phenotyping	Validity of ankylosing spondylitis diagnoses in The Health Improvement Network; 2016 (62)	Validity of AS diagnoses in THIN, via 85 surveys to GPs for a patient with AS in the EHR	Questionnaires yielded PPV 72% for AS, or 89% where 2+ codes >7 days apart, or 86% for AS plus DMARD/biologic, and high sensitivity if no OA/RA codes were present (98%). The recommended selection criteria is 2+ codes >7d apart. DMARD use was 29.5%, 98% used NSAIDs	UK	NSAID, DMARD, both, biologic (0 in EHR)	THIN	61	2000	2013		PPV, sensitivity
AS phenotyping	HERALD (health economics using routine anonymised linked data); 2012 (401)	Value of linking data for health economics analysis	Linked data offers unique opportunity for longitudinal health economic analysis	Wales	Paracetamol, ibuprofen, naproxen - patient recall; GP Read codes for drugs for	SAIL: GP EHRs, inpatient, outpatient, A&E, surveys	183	NS	NS		Descriptive

					many disease groups					
Epidemiology	Differences in the prevalence of ankylosing spondylitis in primary and secondary care: only one-third of patients are managed in rheumatology; 2016 (402)	Prevalence of AS and proportion managed in rheumatology	AS in GP and rheumatology: prevalence 1.34 and 0.47 per 1000, age at diagnosis 38 and 35, uveitis history 22% and 34%, IBD history 6% and 12%, Ps 6% and 14%. This indicates 35% are managed in rheumatology.	Scotland	No	Primary care database and hospitals registry	3,664	2010	2013	Descriptive statistics with t-tests or chi^2 tests, prevalence
Epidemiology	Immune-related disease before and after vasectomy: an epidemiological database study; 2007 (403)	Any association between vasectomy and immune-related disease	No long term risk of AS or other diseases following a vasectomy, or risk of vasectomy following these diseases	England	No	Hospitals and death registrations in Oxford Linkage Study	204	1963	1999	Rates of occurrence compared with control
Epidemiology	Prevalence and incidence rates of cardiovascular, autoimmune, and other diseases in patients with psoriatic or psoriatic arthritis: a retrospective study using Clinical	Disease prevalence in PsA and Ps patients	PsA patients had higher rates of AS than Ps patients (HR 6.98, 2.37-20.58)	UK	In relation to Ps/PsA: NSAID, DMARDs including MTX, SSZ, cyclosporin, leflunomide, azathioprine,	CPRD	115	2006	2010	Comparison of characteristics via t-tests or Chi^2, IRRs with comparison via Wilcoxon or Chi^2, cox proportional hazards to compare comorbidities, with / without adjustment

	Practice Research Datalink; 2015 (404)				Acitretin, hydroxyurea						
Epidemiology	Extra-gastrointestinal manifestations of inflammatory bowel disease may be less common than previously reported; 2016 (405)	Odds of extra-intestinal manifestations in IBD vs non-IBD patients	OR of AS was 7.07 in IBD patients, though still very rare (1.1% compared to 0.2%)	UK	No	CPRD	1,415	1987	2011		Conditional logistic regression with adjustment for age, BMI and tobacco use
Epidemiology	Hospital admissions for vitamin D related conditions and subsequent immune-mediated disease: record-linkage studies; 2013 (406)	Assess association between vitamin D deficiency and immune-mediated disease	The rate of AS is doubled in patients with rickets, osteomalacia or vitamin D deficiency	England	No	HES and death registrations	16	1999	2011		Rates, o/e, rate ratio, Chi^2
Epidemiology	Severe flare as a predictor of poor outcome in ankylosing spondylitis: a cohort study using questionnaire and routine data linkage; 2015 (407)	Explore flare and remission patterns as predictors of poor outcomes in AS	72% experienced flares pre-diagnosis. 58% reported severe flares. (69% of these had pre-diagnosis flares). Flares associated with worse function, disease activity, work impairment, anxiety and more GP visits. Patients with unremitting disease had worse outcomes and were more likely to smoke. 17% received TNFi	Wales	TNFi (though not sure if from EHR or survey data)	Survey and SAIL	348	1999	2009	Uncertain as to what is self-reported or from EHR	Descriptive statistics, regression analysis, thematic analysis of written responses

AS management	Good outcomes of percutaneous fixation of spinal fractures in ankylosing spinal disorders; 2014 (408)	Outcomes of percutaneous fixation of spinal fractures in ankylosing spinal disorders	Surgery for spinal fracture did not lead to complications	UK	No	Hospital database	10	2009	2013		Descriptive
Work disability and cost	The Cost of Ankylosing Spondylitis in the UK Using Linked Routine and Patient-Reported Survey Data; 2015 (189)	Cost of routine AS care from questionnaires and routine data (where linkable)	AS care costs ~£19k per year (over 80% work-related): GP visits, administration, hospital costs, patient-reported costs, early retirement, absenteeism, presenteeism, unpaid assistance	Wales	NSAID/painkiller, DMARD/TNFi	Survey, hospital data and SAIL	482 survey & HES, 150 also GP db	2009	2010	Larger cohort as only ~200 linked EHRs; therapy costs	Descriptive statistics with bootstrapped CIs (1000 iterations), regression analysis
Comorbidity and mortality	Associations between selected immune-mediated diseases and tuberculosis: record-linkage studies; 2013 (403)	Association between diseases and TB	TB risk increased after hospitalisation for AS, but AS risk did not increase following TB	England	No	HES admissions, ONS Death, Oxford Record Linkage Study	30,287	1963	2011		Rate ratio of TB (o/e) and chi^2 test for its significance, sensitivity analyses
Comorbidity and mortality	Risk of subarachnoid haemorrhage in people admitted to hospital with selected immune-mediated diseases: record-linkage studies; 2013 (409)	Risk of subarachnoid haemorrhage in patients hospitalised with immune-mediated disease	SAH rate was increased following hospitalisation for AS (RR 1.64) and rate ratio within 12 months was 1.99	England	No	HES and death registrations	29,136	1999	2011		Rate ratio based on person-days, Chi^2 for significance of the confidence intervals

Comorbidity and mortality	Ankylosing spondylitis and risk of ischaemic heart disease: a population-based cohort study; 2016 (410)	Risk of IHD in AS patients	IRRs and HRs for IHD and AMI were non-significant. NSAID use tended to increase IHD risk (non-significant) (recorded in 3m pre-diagnosis and follow-up)	UK	NSAID, Cox-2 inhibitors, naproxen	CPRD	3,809	1987	2012	NSAID - IHD: linked through disease activity? Could control via resource utilisation?	IRs, IRRs, HRs, baseline characteristics compared with controls via Chi^2, Cox hazard of first IHD or AMI
Comorbidity and mortality	The epidemiology of extra-articular manifestations in ankylosing spondylitis: a population-based matched cohort study; 2015 (411)	Incidence and risks of acute anterior uveitis, Ps and IBD in AS patients	Prevalence at AS diagnosis: 11.4% AAU, 4.4% Ps, 3.7% IBD. Incidence rate per 1000 person-years: 8.9, 3.4 and 2.4 respectively. HR: 15.5, 15. and 3.3. Ps and IBD risk was greatest in the first year after diagnosis while AAU risk remained high even after 10 years. 46.9% took NSAID in 6m pre-diagnosis and follow-up compared to 8.6% controls	UK	NSAID, antidepressants	CPRD	4,101	1987	2012		IR, cumulative IR, adj HRs compared to control, with time-dependent adjustments for age, sex, comorbidity and medications

AS = Ankylosing spondylitis; AMI = acute myocardial infarction; HR = hazard ratio; GP = general practice; SAIL = Secure Anonymised Information Linkage; NSAID = non-steroidal anti-inflammatory; THIN = The Health Improvement Network; EHR = electronic health record; PPV = positive predictive value; DMARD = disease-modifying anti-rheumatic drug; OA = osteoarthritis; RA = rheumatoid arthritis; NS = not stated; IBD = inflammatory bowel disease; Ps = psoriasis; PsA = psoriatic arthritis; MTX = methotrexate; SSZ = sulfasalazine; CPRD = Clinical Practice Research Datalink; OR = odds ratio; o/e = observed / expected; TNFi = anti-tumour necrosis factor; HES = hospital episode statistics; CI = confidence interval; TB = tuberculosis; ONS = Office for National Statistics; RR = rate ratio; SAH = subarachnoid haemorrhage; IRR = incidence risk ratio; IHD = ischaemic heart disease; IR = incidence rate; AAU = acute anterior uveitis

5.3.2 Study Themes

Five themes emerged:

1. Diagnosis (9 studies)
2. Epidemiology (16 studies)
3. Disease impact (10 studies)
4. Comorbidity and mortality (21 studies)
5. Management (17 studies)

The following sub-sections review the topics investigated and key findings in each theme.

5.3.2.1 Diagnosis

Studies examined the validity and benefits of defining AS in EHR-based studies. Dubreuil et al. used GP questionnaires and the Assessment of SpondyloArthritis International Society classification criteria to estimate the PPV of AS diagnostic codes in the THIN database as 71.8% (95% CI 61.0-81.0) and 72% respectively for ≥1 code and 88.6% and 77% for ≥2 codes separated by ≥7 days (62). In the study, GPs completed a questionnaire and reviewed the EHRs as part of the validation process to determine PPV (62). Codes for AS in administrative data were reported as having 83% PPV for ≥1 code and 100% PPV for ≥2 codes, using rheumatologist diagnosis as the gold standard although the cohort size was small (n = 6) (412). No improvement was found from additionally using DMARD prescriptions. Curtis et al. reported high PPV (81%) for ≥2 codes assigned by a rheumatologist but found that 53% of suspected AS cases did not see a rheumatologist (251).

The relevance of data linkage was also highlighted: 12% of patients having AS in GP records had a different diagnosis in rheumatology and 24% of rheumatology-based diagnoses were missing from GP records, while linkage was shown to facilitate patient stratification and health economic analysis and offer similar additional insight (400, 401, 413). In addition, in one study, for 37.7% of patients with an AS diagnosis recorded in their GP EHR, symptoms of IBP were unknown to the GP (62).

Only a few EHR-based studies explored diagnostic delay and these have been small-scale hospital studies. A study (n = 14) identified spinal injury as a factor associated

with delayed diagnosis (414). A study of 677 patients reported that diagnostic delay contributed to mortality (176). Delay to diagnosis of spinal fracture was also reported in AS patients (n = 119), with 28% of cases initially misdiagnosed, which associated with risk of severe neurological complications (415).

5.3.2.2 Epidemiology

EHR-based research highlighted risk factors for AS and worse disease severity. A registry study identified 50% concordance of AS in monozygotic twins, suggesting a genetic component (416). AS diagnosis was reported as three-times more common in men and at a younger age than in women (417). One study reported no difference in incidence risk in men following vasectomy (403). A study in Argentina reported that AS patients had a similar sociodemographic pattern to the general population although functional capacity and disease activity were worse in more deprived patients (418). Chronic inflammation, vitamin-D deficiency and autoimmune response associated with the risk of AS, while human immunodeficiency virus was reported as protective in one study (405, 406, 419, 420). Patients with psoriasis had 13 times greater risk of AS (odds ratio [OR] = 13.34, 12.02-14.81) (421). Risk factors for worse disease severity included hip involvement in AS, which in one study associated with early disease onset and impaired functioning (422). Flares pre-diagnosis associated with poor long-term outcomes including worse function and anxiety (407).

Incidence and prevalence were relatively under-explored. Annual incidence was reported in secondary care as 0.7 per 10,000 persons in Norway (1982-1993), and higher in Canada where primary care diagnoses were included (1.4-1.6 per 10,000 persons between 1995 and 2010) suggesting that not all patients receive rheumatology referral (182, 249). One study found rising AS incidence in women (249). Prevalence estimates for AS varied, ranging from 0.09% to 0.49% in European regions and lower in American regions e.g. 0.11% in California (1996-2009), 0.13% in Minnesota (1935-1973) (250, 251, 417, 423). A study reported variation in prevalence between GP and rheumatology clinic settings, 0.13% and 0.05% respectively, indicating that 65% of cases are not managed in rheumatology (402). Some studies report rising prevalence e.g. in Canada from 0.08% in 1995 to 0.21% in 2010 and in Norway from 0.04% in 1970 to 0.21% in 1990 (182, 249).

5.3.2.3 Comorbidity and Mortality

The incidence of extra-articular manifestations in AS has been investigated using longitudinal EHR data. AS patients were 15 times more likely to develop anterior uveitis and risk of psoriasis and IBD were raised in the year post-diagnosis (180).

EHR-based studies reported raised comorbidity and mortality risk in AS patients, compared with the general population, and examined treatment safety and effectiveness. As with other autoimmune rheumatic diseases, rates of diabetes, periodontitis and hypertension were higher (220, 400, 419, 424). Large datasets have facilitated analysis of uncommon occurrences such as spontaneous pneumothorax and uncovered raised risk of nephrolithiasis (AHR=2.1, 95% CI 1.8-2.4) despite only 715 events occurring in 57,830 (425, 426). Mortality may be 14.5 times higher, with 40% of deaths attributed to CVD (176). Anti-TNF inhibitors were shown to have no association with all-cause mortality, birth defects or cancer (427-429). Increased risk in neck, lung, head cancer and myeloma, although no elevated lymphoma risk, were reported (430-433).

Variation in disease severity or therapy in the study population may have contributed to conflicting reports of cardiovascular and cerebrovascular risk, with three studies reporting raised morbidity or mortality from these while two studies reported no difference in acute myocardial infarction or ischaemic heart disease risk (220, 400, 410, 434). High stroke risk was reported in a cohort hospitalised with AS but not in all AS patients and suboptimal therapy associated with mortality (176, 400, 409, 435, 436). Increased risk of tuberculosis and RA were reported for hospitalised cohorts (403, 431) but this requires comparison with less severe AS cohorts.

5.3.2.4 Disease Impact

Linked data was used in EHR-based studies of the societal cost of AS. Lost income and early retirement were the main costs, as identified using survey and EHR data in Wales (n = 150) and Hong Kong (n = 148) (189, 437). Risk of sick leave in the under 45s was 2.1 times higher (95% CI 1.6-2.8) than in a control population and in Italy disability pensions form 54% of the social security cost of AS (438, 439). Patients with work disability were older and had more radiological damage (440). Effective AS management may limit functional impairment that is the most important driver of work-related costs and anti-TNF inhibitors may half the risk of sick leave (437, 441). However, the specific cost-effectiveness of medication would depend on medication

costs, which vary between countries. For example, adalimumab had the greatest annual cost of anti-TNF inhibitors in Austria but not in the USA and the cost-effectiveness for etanercept was higher in the UK than Germany (442-444).

5.3.2.5 Management

Therapeutic management trends have been described using registry, health insurance and hospital data. Rates of therapeutic management in rheumatology clinics varied across countries, e.g. 24% and 22% having NSAIDs and biologics respectively in Turkey (2012-2013) and 58% and 28% having had biologics and non-biologics in Portugal, 2008-2010 (252, 445). Common first-prescribed DMARDs were sulfasalazine, with a median drug survival time of 4.5-years, and more recently etanercept and adalimumab especially among younger patients (443, 446). Biologics were initiated late in the disease course (mean 13.4 [±10] years post-diagnosis) (252). Further, non-adherence to biologics, defined as PDC <0.8, was reported among 28% of patients and in one study only 48% of patients prescribed anti-TNF inhibitors had continuous prescription for ≥12-months (447, 448).

Longitudinal data facilitated comparison of outcomes between management options. Outcomes from anti-TNF inhibitors were examined following their introduction and therapeutic benefit included a mean reduction in the disease activity, measured using the Bath Ankylosing Spondylitis Disease Activity Index score, from 6 [±2] to 2.7 [±2.2] following six-months of biologic therapy (252). The risk of anterior uveitis was reported to be reduced in patients with adalimumab or infliximab therapy, yet raised in patients with etanercept (449). EHR data were also used to compare approaches to identifying patients that respond well to therapy (450).

The ankylosed spine is prone to trauma and spinal fractures are common in AS, which requires suitable management. One study reported that NSAID therapy might reduce the fracture risk in AS patients (451). Hospital EHRs have facilitated comparisons of surgery options. Early surgery with percutaneous fixation showed improved outcomes for spine fractures in AS compared with braces and associated with fewer complications and shorter hospitalisation than open stabilisation techniques (408, 452, 453).

5.4 Discussion

This thematic scoping review of EHR-based studies in AS identified an increasing number of studies published over time. Most (87%) studies used secondary care data, from a single hospital EHR system, an insurance claims database, or from summary data extracted from clinical systems and submitted to a regional or national database (e.g. Oxford Linkage Study; HES). These capture the involvement of rheumatology in AS management, however the single site and claims-based studies may not be representative of the general population and studies using summary datasets used only a subset of the relevant data fields available in EHRs (e.g. prescribing data was not included in HES). In the UK, only 4/15 studies used a population-based GP EHR database representative of the general population (CPRD). Five major themes were identified in the studies. Definitions of AS in EHR-based studies were shown to have high validity, especially when ≥2 AS codes were employed. Studies have investigated factors that might contribute to diagnostic delay, and identified risk factors for AS and factors associating with worse disease outcomes. Population-level estimates of the incidence and prevalence of AS are uncertain. Comorbidity and mortality risk were significantly raised in AS patients. Studies have used linked data to investigate the high levels of work disability in AS patients. Hospital-based studies have reported on prescribing patterns and compared the effectiveness of medication and surgery options.

Studies reported high validity in defining AS in EHR-based studies, suggesting that EHR data are suitable for investigating AS. Appropriate validity is required to enable the benefits offered by population-based EHR databases for large-scale study of this less common disease, to uncover infrequent events or secular trends otherwise overlooked in small studies. Improved validity was reported from definitions of AS that required ≥2 diagnostic codes, although not from requiring additional DMARD prescriptions (62, 251, 412), which is an important consideration for future AS studies. Differing or missing diagnoses between GP and rheumatology data were reported (400), which suggests that future studies should consider record linkage where possible and the impact of the AS definition on results, e.g. through sensitivity analyses. Significant diagnostic delay was also reported in hospital studies. These studies commonly reported on the mean time to diagnosis, however, normal distribution of the data cannot be assumed. Correspondingly, the median rather than mean time to diagnosis will be calculated in the investigation into diagnostic delay in Chapter 8 of this thesis.

Studies of the risk factors for AS, and of the factors associating with worse disease outcomes, have contributed to important epidemiologic understanding. Investigation of the risk following vasectomy was relevant given that AS is most common in men (403). The study of AS incidence among siblings (416) suggested the contribution that routine recorded data can make to understanding genetic and environmental factors in disease risk. The factors contributing to worse outcomes in more deprived patients require investigation (418); given the role of early biologic prescribing in disease outcomes, differential healthcare access might be a factor. The association noted between flares pre-diagnosis and poor outcomes (407) may have been confounded by delays in diagnosis given that these may be considerable.

The variation in prevalence reported between GP and rheumatology clinic settings, (0.13% and 0.05%) (402), suggests the importance of using GP data in estimating AS incidence and prevalence if not all patients receive rheumatology referral. The factors contributing to reports of rising AS prevalence and rising incidence in women (182, 249), require further investigation. It may be that recognition of AS has improved over time.

Long-term investigation of comorbidity, birth defect and mortality risk are particularly pertinent in the AS population where disease onset is early in adulthood and lifelong treatment is initiated early in the disease course. EHR-based studies have reported on long-term comorbidity and mortality risk in AS patients, reporting higher risk of diabetes, hypertension and CVD (176, 220). This increase risk suggests the importance of comorbidity screening in patients with AS. The large cohort size that is possible in EHR studies facilitated the investigation of uncommon but important occurrences (425, 426). Some risks were reported in hospitalised cohorts and require investigation in the non-hospitalised AS population as disease severity may also affect risk. Comorbidity and fracture risk was affected by medication therapy, yet studies did not always report on prescribing or make adjustments for this in investigating risk. The effectiveness of AS therapy in reducing comorbidity risk highlights the importance of appropriate disease management.

EHR data was used to determine the direct and indirect costs of biologic treatment for the healthcare system in different countries (442-444). The huge societal cost of AS through functional impairment was identified through studies with linked data. Biologic therapy was reported to be effective in reducing functional impairment and sick leave (437, 441), further suggesting the importance of efforts to reduce diagnostic delay and promote the initiation of biologics in early AS.

The advantage of long follow-up in EHRs has facilitated the investigation of prescribing patterns and the long-term safety and effectiveness of treatment, in hospital settings. The finding of biologics to avoid in patients with underlying susceptibility to anterior uveitis was important, given that this is a common EAM (449). The lengthy delay reported between diagnosis and biologics initiation (13.4 years), and suboptimal medication adherence, is of concern (252, 447, 448). It is uncertain whether factors contributing to optimal pharmacologic management, such as timely diagnosis and early biologic therapy, have improved over time following efforts to promote biologic use in early AS (e.g. BSR guidelines published in 2005 (240)).

5.4.1 Strengths and Limitations

This thematic scoping review evaluated the areas of investigation and key findings in EHR-based studies of AS, and described the UK data sources used. A strength of the study was the breadth of synonyms and MeSH that were used to enhance the sensitivity of the literature search. A conservative screening approach was adopted whereby abstracts were included for full-text review if their eligibility was uncertain. The systematic approach of the study was strengthened by using a study selection protocol and data extraction form. The comprehensiveness of reporting was strengthened by following the PRISMA statement.

A limitation of the study was that non-English publications were excluded and EHR studies may have been missed where they did not report on using EHR-derived data, given the lack of a standardised term for EHRs. Registry and health insurance database studies were included given the underlying source may be EHR data and the methodological similarity in studies, although terms for these were not searched. The literature search could have been performed in addition on the Scopus literature database as Chapter 4 determined this returned the highest numbers of studies for the EHR search term, although Web of Science returned the second highest counts and a higher number for the health term (gout). Recruitment of a second reviewer in study selection and data extraction could have improved the quality of the review. The RoB was not assessed although the low RoB reported for EHR-based studies in Chapter 4 suggests this may be low.

5.4.2 Conclusion

The breadth of study themes and increasing number over time suggest the contribution of EHR data to a real-world understanding of AS, including its epidemiology and management. This is facilitated by the high validity of AS diagnostic definitions derived using EHR data. Studies have identified risk factors for AS and comorbidities, and the importance of pharmacologic management, yet have also reported significant diagnostic delay and suboptimal prescribing. Use of GP EHR data is less common, despite the role of GPs in both the timely diagnosis and referral to rheumatology for biologic therapy, and in the screening and management of comorbidities. The limited understanding of incidence, prevalence and time to diagnosis will be addressed using GP data in Chapters 7 and 9.

Chapter 6 Rheumatoid Arthritis: An Initial Thematic Scoping Literature Review of Electronic Health Record-based Research

6.1 Introduction

This chapter informs the first objective of the thesis with focus on EHR-based research in RA. As described in Chapter 2, RA is a chronic inflammatory musculoskeletal disease in which treatment with DMARDs from the early stages of disease is important for preventing joint damage, and annual review of treatment and comorbidity screening is recommended (280). As noted in Chapter 2, the diagnosis of RA and pharmacologic prescribing are recorded in EHRs, and EHR data has informed a real-world understanding of RA. However, a literature review is required to understand the study themes examined using EHR data in this field. As also described in Chapter 2, principles of pharmacologic management of RA differ between countries, related to national guidelines and historic practice, and particularly regarding corticosteroid prescribing and its initiation and duration of tapering (280, 297, 298). Given this, and the use of UK EHR data in subsequent chapters of this thesis, this chapter aims to provide a thematic scoping review of all RA studies using UK EHR data.

6.2 Methods

6.2.1 Literature Search

A protocol was defined for the literature searches and for selecting EHR-based studies of RA. The search terms were' EHR' and 'RA' and the synonyms and MeSH were informed by the knowledge base in Chapter 2 and the terms used in Chapters 4 and 5. Synonyms relating to CPRD were included in the EHR term to ensure inclusion of studies using the main UK EHR database. As described in Chapter 4, in database searches the synonyms and MeSH for a term were combined with 'OR' and then the terms combined with 'AND'. The literature search was performed using PubMed, Web of Science and Google Scholar and by hand-searching the references of selected manuscripts (Table 19, Table 22, Table 23). In PubMed and Web of Science, MeSH terms were used and filters were applied to exclude citations published prior to 1970 (the first decade with established EHRs), of non-English language, or with non-human subjects. The search of PubMed was performed on 30 October 2017, with searches of

Web of Science Core Collection and Google Scholar performed on 8 December 2017 (Table 19, Table 22, Table 23). Harzing's Publish or Perish (version 6) was used to collect the first 1000 manuscripts returned by Google Scholar (316). Given the constraint on returns in Publish or Perish, to ensure the most relevant return, 'UK' was included as a search term and manuscripts published pre-2000 were excluded. Manuscripts were imported into EndNote for de-duplication and then further de-duplication was performed in Rayyan, as described in Chapter 3 (318, 320).

Table 21. Search terms and synonyms used in PubMed for the RA EHR search, with count of returned citations (individually and in combination)

Search Term	Return
Rheumatoid arthritis (RA)	134,852
Rheumatoid arthritis+	134,852
Electronic health record (EHR)	1,026,420
(Health OR patient OR clinic* OR medic* OR care) AND (compute* OR system OR electronic OR warehouse OR link* OR dataset OR network OR database)	994,547
Medical records systems+	36,083
Electronic health record+	33,861
eHealth	26,738
EPR [electronic patient record]	19,826
Record-linkage	6,518
EMR [electronic medical record]	5,855
EHR	4,705
e-health	2,625
Datalink	977
Routin* n5 data	7
(Electronic OR link* OR compute* OR anonymi*) n5 record	2
General Practice Research Database	3,855
Clinical Practice Research Datalink	735
GPRD	456
CPRD	324
EHR and RA	6,647
Limit to humans	5,141
Since 1970	4,695
Endnote 'Discard Duplicates'	4,692
Rayyan manual de-duplication	3,691

Note: MeSH terms are indicated by '+' and a wildcard by '*'; PoP = Publish or Perish

Table 22. Search terms and synonyms used in Web of Science for RA EHR search, with count of returned citations (individually and in combination)

Search Term	Return
Rheumatoid arthritis (RA)	157,532
Electronic health record (EHR)	1,877,199
(Health OR patient OR clinic* OR medic* OR care) AND (compute* OR system OR electronic OR warehouse OR link* OR dataset OR network OR database)	1,785,710
EPR	58,204
(Electronic OR link* OR compute* OR anonymi*) NEAR/5 record	46,461
Medical records systems+	24,846
Routin* NEAR/5 data	18,415
Electronic health record+	18,098
EMR	6,437
e-health	4,434
Record-linkage	4,358
EHR	4,261
eHealth	3,137
Datalink	1,310
General Practice Research Database	3,018
Clinical Practice Research Datalink	743
GPRD	471
CPRD	391
EHR and RA	20,031
Endnote 'Discard Duplicates' (added to PubMed and PoP returns)	2,796
Rayyan manual de-duplication (added to PubMed and PoP returns)	2,781

Note: MeSH terms are indicated by '+' and a wildcard by '*' ; PoP = Publish or Perish

Table 23. Search applied to Google Scholar using Publish or Perish for RA EHR search

Search selection	Search terms
All of	Rheumatoid arthritis, UK
The phrase	Electronic health record
Timeframe	2000 – 2017

6.2.2 Study Selection

In addition to the selection criteria stated in Chapter 3, inclusion criteria were that studies must report on RA in adults (aged >17), using UK-based EHR data. One reviewer (SSRC) performed the study selection in Rayyan to select English-language studies using UK EHR data to report on the adult human RA population (318). Manuscripts with uncertain relevance were marked for inclusion during title-abstract screening. Excluded full-text manuscripts were assigned a pre-defined exclusion label based on the exclusion criteria.

6.2.3 Data Extraction

A data extraction form was created using Microsoft Excel 2013 and information was extracted for each study regarding the title; study timeframe; data source; study objective/s; methods / measures; RA medications studied; key RA-related findings. Study limitations and opportunities for further work were also noted. The form was based on the recommendations of the CRD regarding general information, study and participant characteristics, setting and results (337). The data was abstracted by SSRC and descriptive statistics were performed using Microsoft Excel 2013.

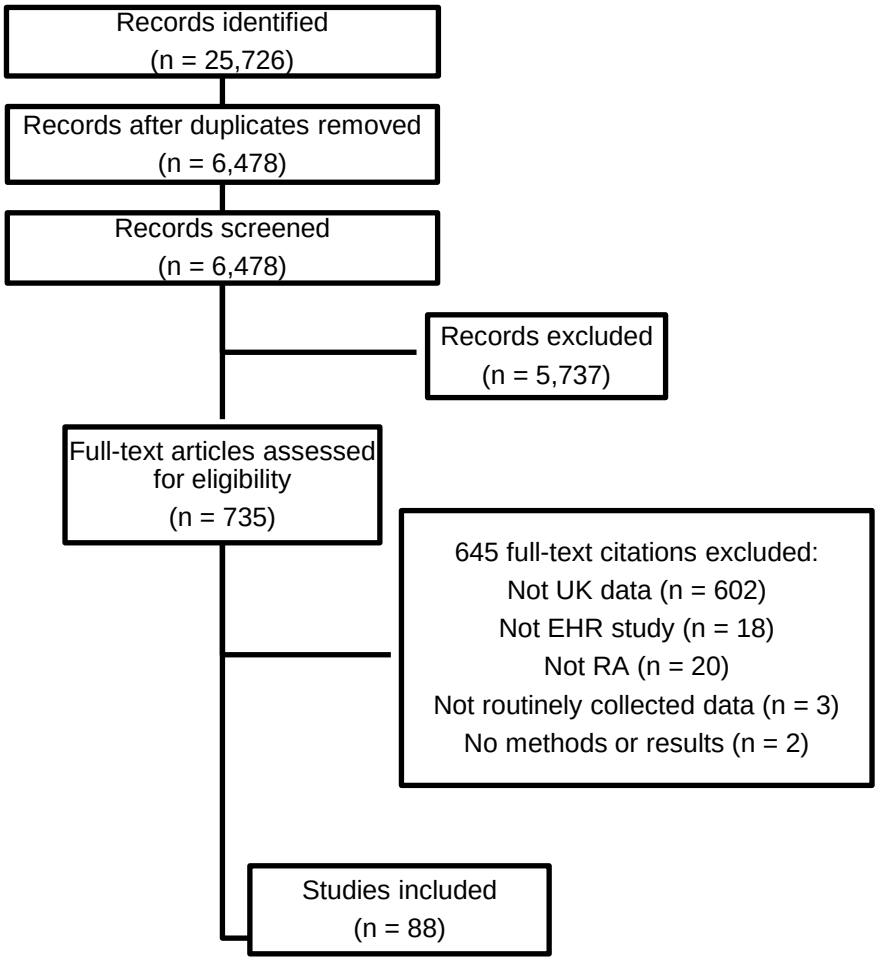
6.2.4 Study Outcomes

The primary outcome was to describe the areas of investigation (study themes) in EHR-based studies of RA, including methodology and key findings. A secondary outcome was to identify the sources of data used in these studies (e.g. GP, hospital).

6.3 Results

There were 6,478 manuscripts identified from which 88 were selected (Figure 19). The earliest study period commenced in 1981 (ending in 2000) and the latest commenced in 2012 (ending in 2016). The median study duration was 14 years (IQR = 10-17).

Figure 19. Flowchart of study identification and selection



6.3.1 Data Source

Of the studies, 66 used GP data from the CPRD, QResearch and THIN and 8 used other GP or GP and linked hospital datasets, while 14 used hospital or hospital registry data (Table 24).

Table 24. Sources of EHR data in the studies (N = 90)

Source	Type	Count
CPRD	GP, HES and national	40
CPRD and US databank	GP and US data	2
CALIBER (CPRD + MINAP)	GP, HES and registry	1
CPRD and biologics register	GP, hospital, registry	1
THIN	GP	14
QResearch	GP	7
THIN and QResearch	GP	1

Other GP group	GP	5
Other GP and hospital	GP and hospital	3
HES or HES and mortality	Hospital and national	4
Hospital cohort	Hospital or hospital and registry	10
	Total	88

Note: CPRD = Clinical Practice Research Datalink; GP = general practice; HES = Hospital Episode Statistics, MINAP = Myocardial Ischaemia National Audit Project; THIN = The Health Improvement Network

6.3.2 Study Themes

Four themes emerged from the studies:

1. Diagnosis (7 studies)
2. Epidemiology (12 studies)
3. Comorbidity and mortality (40 studies)
4. Management (30 studies)

6.3.2.1 Diagnosis

In RA, EHR-based studies ascertained cases of RA by identifying associated traits in a patient's data, including diagnostic or symptom codes and disease-specific medication, with accuracy estimated through the review of records by a clinician. The approaches to select diagnostic indicators usually combined clinical knowledge with data-driven methods, performing code-frequency comparisons via random forest, Chi^2, Mann-Whitney U tests or Kruskal-Wallis, in one instance with comparison to an algorithm solely based on clinical expertise.

The RA phenotyping algorithms developed using EHRs were generally highly reflective of diagnostic coding (454, 455). Rheumatoid arthritis phenotyping and diagnostic studies reported the existence of related arthritic symptoms, rheumatology referrals and tests in the year prior to coded diagnosis in up to 83.5% of EHRs, which may reflect a period of diagnostic uncertainty or earlier implicit diagnosis (456, 457). Definitions of RA diagnosis based on DMARD prescribing, existence of subsequent RA codes during follow-up and having no later alternative diagnostic codes had >80% sensitivity (458). However, while all records with RA references in free-text were reported in one study to have an RA diagnostic code, there was evidence for a delay in

RA diagnostic coding, including from 22% of GP records having free-text references to DMARDs more than two weeks prior to diagnosis (459).

6.3.2.2 Epidemiology

In studies of incidence and prevalence, the reported measures were incidence rate (IR), incidence rate ratio (IRR) and percentage prevalence, with standardisation for length of data contribution (duration of registration) and join-points analysis to test for trend changes. In studies of risk factors for developing RA, the measures reported included IR and IRR, with Chi^2 test or Wilcoxon signed rank for significance. Time-to-event measures through Cox proportional hazards and log-log survival plots were common, with Schoenfeld residuals to test model assumptions of proportionality.

Incidence and prevalence estimates were reported up to 2014, with some studies reporting regional variation. Two CPRD studies estimated national levels of crude incidence, using phenotyping algorithms with differing specificity: Abhishek et al. defined RA by diagnostic codes, while Rodriguez et al. defined RA by diagnostic code plus a specialist referral, diagnostic test, specific treatment or confirmation from the GP (265, 460). The reported incidence from these studies was 3.9 and 1.5 per 10,000 person-years respectively, in 1996 (265, 460). Both reported higher incidence with age and in women than men (265, 460). Abhishek et al. also reported decreasing incidence across 1990-2014 at a rate of 1.6% per year, to 3.8 per 10,000 person-years in 2014, with the greatest change in the East Midlands and Yorkshire, and higher incidence among older patients and women (265). Silman et al. used GP-notified instances of first onset of joint swelling and found no evidence of time or space clustering in RA incidence in East Anglia (266). In this study, it was not specified that EHRs were used by GPs to identify the cases, although this study was included in the scoping review given that GP practices in the UK predominantly use EHRs (16, 266). Crude prevalence was also reported as decreasing by 1.1% per year from 2005 to 0.67% in 2014 using the CPRD, and being higher when estimated from databases that mandated or prompted diagnostic coding (265, 461).

Longitudinal and large population EHR databases facilitated studies identifying risk factors for RA. Studies using the CPRD reported that hyperlipidaemia increases RA risk and statins may be protective, particularly at high doses (462, 463). Rodriguez et al. found no associated risk of RA with alcohol use, obesity, or prescribing of hormone replacement therapy or low-dose aspirin (460). Severe psoriasis, psoriatic arthritis, Klinefelter's syndrome, vitamin D deficiency and Th1- and Th2-mediated diseases also

associate with increased risk of RA (404, 406, 464, 465). Conversely, EHR-based studies suggest that physical trauma and multiple sclerosis are not risk factors (466, 467).

6.3.2.3 Comorbidity and Mortality

EHR-based research has contributed to investigations of the lifetime risk of comorbidity and mortality among RA patients. Reported associations include increased risk of developing comorbidities, particularly cardiovascular and metabolic diseases, and higher rates of fractures and mortality (468-473). The risk of cardiac arrest was twice more high in RA than in non-RA patients (HR = 2.26, 95% CI 1.69-3.02) and cardiovascular deaths were also comparably elevated (AHR = 1.55, 95% CI 1.44–1.66), as well as respiratory deaths (AHR = 1.85, 95% CI 1.72–2.01) (471, 473). Based on GP diagnostic coding, one-third of patients had pre-existing comorbidities at RA diagnosis, which associated with increased risk of subsequent cardiovascular events though not with RA disease activity or structural damage (468). Although suicide deaths were rare, a GP EHR-based study reported elevated rates in RA compared with non-RA patients (AHR = 2.47, 95% CI 1.51-4.04) (471). Carpal tunnel syndrome and tuberculosis were also more common among RA patients (AOR=2.23, 95% CI 1.57-3.17; and rate ratio 3.2, 95% CI 3.0-3.5, respectively) (403, 474). In one study, an apparent association between RA and incident diabetes was removed following adjustment for body mass index, smoking status, alcohol intake and corticosteroid prescribing (475).

6.3.2.4 Management

Pharmacologic and non-pharmacologic aspects of management are important in RA, and these had been investigated in 21 and 10 studies respectively.

6.3.2.4.1 Pharmacologic Management

All studies of pharmacologic management were longitudinal (the minimum study duration was four months) and reported measures included population summary statistics (e.g. proportion of years with each prescription, mean dose and mean count of courses per year) (476). Patient characteristics among therapy versus non-therapy groups via t-tests; prescribing variation among demographic groups was described and tested with logistic regression; prescription durations were assessed using log rank tests. Medication outcomes were generally assessed using IR, relative risk estimates,

logistic regression, multivariate Cox regression models and segmented linear regression in time-series analysis. From reporting on drug definitions used, it was not clear how medication was selected in the studies and if this was based on name (generic or brand) or drug group / chapter.

Most EHR-based studies into routine pharmacologic management in RA used the CPRD GP database (13/21 studies) and therefore reported on GP prescribing, predominantly of non-biologic DMARDs (non-bDMARDs). The proportion of RA patients prescribed non-bDMARDs was stable between 1987 and 2002 and then rose, from 19-49% in 1995-1999 to 45-74% by 2006-2010 (477, 478). Between 1995 and 2010, methotrexate replaced sulfasalazine as the most common prescription (4% to 60%) and there was a slow growth in combination prescribing, to 17% (477). The stepwise increase in GP-led non-bDMARD prescribing follows BSR guideline recommendations published in 2006, which saw the annual increase in prescribing rate rise from 1.64% to 3.55% (479, 480). Studies have also assessed prescription duration and prescribing patterns, with methotrexate having the longest median duration (8 years) and being favoured in first-line therapy in patients with multiple swollen joints and poor prognosis (478, 481).

Studies of the effectiveness and side-effects of DMARDs used GP data. Abatacept did not associate with increased risk of malignancies, in a study which used registry as well as EHR data (CPRD and Swedish EHRs), but was included because it reported on the EHR-derived cohort separately to the registry-derived cohorts (482). As biologics are prescribed in secondary care, the impact of prescribing was indirectly estimated in a CPRD study that compared knee replacement surgery rates before and after the publication of national guidelines on biologics, to attribute a 34% reduction to the use of biologics (483). The role of GPs in non-biologic DMARD prescribing enabled a CPRD study to investigate septic arthritis risk, reporting an increased risk in RA patients with certain non-biologic DMARDs (e.g. penicillamine, IRR=2.51 [1.29-4.89]; sulfasalazine, IRR=1.74 [1.04-2.91]) (484).

Corticosteroid and NSAID prescribing and side effects were reported using GP data. While Chapter 4 described 56.0% (n = 42) of gout pharmacologic management studies reporting on NSAID prescribing, in AS pharmacologic management studies across a similar time-period, this was only 33.3% (n = 7) (485). A 9-fold increase in NSAID prescribing was noted in the six-months pre-diagnosis of RA (477). Estimates of treatment intent are limited given that NSAIDs are prescribed for a range of inflammations; one study of GP prescribing reported that 4.7% of NSAID prescriptions

were for RA management (476, 486). In the UK, 47% of RA patients received ≥1 prescription of corticosteroid from their GP, with persistent prescribing rates across 1987-2002 though prescription durations and dosages vary (478, 486). Studies reported that corticosteroid prescribing associated with increased risk of diabetes by 30-60%, heart failure and ischaemic heart disease, while acute myocardial infarctions increased during the first eight weeks post-NSAID use but was not affected by current NSAID use (487-489). However, in one study the risk of acute thromboembolic cardiovascular events was lower in RA patients with naproxen prescribed (compared with non-naproxen NSAID and non-NSAID cohorts), even after adjustment for comorbidity and cardiovascular risk score, which may result from the inhibition of platelet aggregation (490). One study investigated prophylaxis prescribing (which aims to reduce risk of toxicity), and reported that concomitant prescribing of bisphosphonates with corticosteroids were cost-effective in reducing fracture and mortality risk (491).

EHR-based studies also evaluated comorbidity prevention, screening and management strategies. Statins showed effectiveness in reducing total cholesterol, mortality and cardiovascular events among RA patients (492). Higher rates of antiviral prescribing for herpes zoster were reported among patients with RA, particularly those prescribed DMARDs, which may suggest appropriate management of immunosuppression (493). However, despite vaccination guidelines for patients treated with DMARDs, only 80% and 50% received ≥1 influenza or pneumococcal vaccine respectively in five years, with very low re-vaccination rates (494). One study reported differences in the time to RA diagnosis and the timeliness of DMARD initiation among RA patients, based on autoantibody status (495).

6.3.2.4.2 Non-Pharmacologic Management

Statistical methods were applied in evaluating non-pharmacologic management. Measurements including IR, IRR and expected and observed counts were compared between groups or modelled using Cox proportional hazards regression. Further methods included competing risk regression analysis, logistic regression, Kaplan-Meier, random forest, time-specific risk ratios, sequential regression and log-log survival plots. Model fit was commonly validated through C-index and the Akaike information criterion (AIC), R^2 and D statistics. Some studies imputed missing data.

Studies of the non-pharmacologic aspects of RA management reported on comorbidity monitoring and the impact of lifestyle factors. Unfortunately, a study reported less

accuracy for RA than non-RA patients in predicting CVD risk using an EHR-based risk tool, while CRP measurements did not significantly improve the accuracy of scores (496). Studies using GP and hospital or death certificate data reported a two-fold risk of hospitalisation for cardiovascular events in smokers and that smoking cessation reduced hospitalisation and mortality risk in RA patients by 10-15% in each year following cessation (497, 498). One study of hospital and survey data reported no association between alcohol consumption and alanine transaminase levels in patients prescribed methotrexate or leflunomide (499). A CPRD study reported that alcohol consumption below 14 units per week did not increase the risk of transaminitis in patients prescribed methotrexate (500).

6.4 Discussion

This thematic scoping review of UK EHR-based studies in RA identified four research themes. Most studies (84.1%; n = 74) used GP EHR-derived data, which may suggest that the availability and accessibility of widespread secondary care data is comparably lower. Useful data for RA studies was reported to be held in EHRs, including symptom, referral, laboratory test, diagnosis, vaccination and prescribing data, and high validity was reported for EHR-based definitions of RA. The risk of RA associated with a number of factors has been explored. One study reported declining incidence in RA, although this may have been affected by changes in RA coding incentives, which will be investigated in Chapter 7 (265). Pre-existing comorbidities are common in RA patients and risk of comorbidity and mortality was found to be higher than in the general population. The suboptimal influenza and pneumococcal vaccine uptake is of concern given the elevated levels of respiratory deaths in RA patients (471, 494). Support for smoking cessation may be important in non-pharmacologic management as cessation associated with reduced hospitalisation and mortality risk (497, 498). Longitudinal studies of pharmacologic management reported rising non-bDMARD prescribing with slow growth in combination prescribing (477, 478).

The high diagnostic validity of RA codes was confirmed through clinical reviews of EHRs, although prior DMARD prescribing suggested a lag in diagnostic recording (459). Studies of the early disease course may therefore benefit from considering factors that may be recorded prior to diagnostic codes such as DMARD prescribing or free-text references to RA in defining early RA. The investigations of sensitivity and specificity in RA diagnostic definitions, and the trade-off between these, can aid future studies in determining the most appropriate definition for addressing a given research question. Studies should consider applying more specific definitions in sensitivity

analyses for the validation of findings. Further research could ascertain whether the sensitivity and specificity of definitions varies based on ethnicity, frailty, pharmacotherapy, weight and socio-economic deprivation. For example, the aspects recorded in EHRs, the route to diagnosis and the time to diagnosis, may differ between patients in a way that can be identified through cluster analysis. It would be important to identify any risk factors for diagnostic delay in RA.

Incidence and prevalence estimates were shaped by the method of case ascertainment used in the study and EHR utilisation or coding practices during the study period. Incidence was lower for confirmed cases than for all patients with an RA diagnostic code (265, 460). Recording practice in organisations contributing to the data source should also be considered as indicated by the higher rates of prevalence reported from GP consultation databases derived through processes that prompt or mandate diagnostic coding (461). The identification of risk factors for RA in EHR-based studies may inform monitoring and prophylaxis regimes, in addition to understanding of the underlying mechanisms of disease, and factors that contribute to the reported variations in prevalence. Further research could examine incidence, prevalence and risk with more detailed stratification by factors commonly derived from EHRs, including ethnicity, polypharmacy, multi-morbidity, body mass index (BMI) and socio-economic deprivation. The impact of duration of exposure and risk of RA, and the impact of pre-existing risk-factors on disease severity and routes of progression and treatment response, could be explored, clustering by patient demographics or frailty. The association between vitamin D deficiency and RA, noted from hospital data, could be explored more comprehensively in GP data with information on supplements, latitude and follow-up (406).

In addition to EHR data on diagnosis and death, demographic and prescription data recorded in EHRs has facilitated investigations of comorbidity and mortality risk in RA. Statistical conditioning is important in such analyses and in one study adjustment based on such factors removed a crude association between RA and risk of diabetes (475). The increased risk of comorbidity, especially CVD, in RA had implications not only for disease burden but for mortality. The higher rates of comorbidity, mortality and suicide is of great concern and suggests the importance of regular patient review, effective pain control, and comorbidity screening. Further work could investigate any association between pre-existing comorbidities and RA severity and progression.

Pharmacologic studies commonly reported on GP non-bDMARD prescribing and reported rising non-bDMARD prescribing between 1987 and 2002 (477, 478). The rise

increased following the publication in 2006 of BSR guideline recommendations, and methotrexate was a favoured first-line treatment, suggesting the efficacy of pharmacologic management guidelines (478-481). The studies reported on the effectiveness of DMARDs in aiding remission and reducing disease progression and knee replacement surgery rates (483, 484). A side effect however was increased risk of infection and septic arthritis, suggesting the importance of medication review, ongoing monitoring and providing patient information on risks. The investigations of pharmacologic management using GP data were limited in that biologic prescribing is led by rheumatologists. However, indirect assessments were made, e.g. between knee replacement surgery rates before and after the publication of national guidelines in 2002 prompting biologic therapy initiation (483). The impact of the publication in 2009 of national guidelines on RA management with DMARDs requires investigation (280).

The investigation of temporal trends in corticosteroid and NSAID prescribing were limited, with studies mainly describing the proportion with prescribed medication at any point in the study period and risks associated with prescribing. Both the 9-fold increase in NSAID prescribing pre-diagnosis, and the low disease specificity reported for NSAIDs, show its use in the relief of general symptoms (476, 477). This non-disease specificity in NSAID prescribing suggests that NSAIDs are not suitable for use in RA diagnostic definitions and that investigations of RA NSAID prescribing should utilise a comparator cohort to aid in discerning RA-specific NSAID prescribing. The proportion of patients with ≥1 corticosteroid prescription was stable across 1987-2002, however, future research could investigate changes in the prescribing duration in recent years and following the 2018 NICE guidance on <3 months corticosteroid prescribing, given the cumulative risk of toxicity (280, 478, 501). Management of RA with DMARDs should limit the symptoms that require ongoing corticosteroid and NSAID prescribing. The study in Chapter 9 will explore whether DMARD prescribing continued to rise and whether this modern treatment paradigm has associated with reduced corticosteroid and NSAID prescribing, with comparison to a non-RA cohort to facilitate understanding of RA-specific prescribing.

Further EHR-based research could evaluate interactions between RA medication and other diseases or medications (individual or combinations) and determine potential long-term impacts over the patient life-course. Large EHR databases provide opportunity to evaluate differences in prescribing and dosage titration, adherence and medication switches in association with factors such as ethnicity, region, frailty, visit frequency or comorbidities.

Data in EHRs was used to describe prophylaxis prescribing with bisphosphonates in corticosteroids, suggesting that EHR-based research could inform evaluations of change in prophylaxis prescribing following the publication of guideline recommendations (491). In the UK, national guidelines published in 2008 and 2010 recommend prophylaxis prescribing with corticosteroids and NSAIDs, as described in Chapter 2 (238, 302). Chapter 9 will investigate temporal trends in corticosteroid and NSAID prophylaxis prescribing between 1998 and 2017. Given the raised CVD risk from corticosteroids (501), future EHR-based research could also investigate CVD risk screening and anti-hypertensive medication prescribing in hypertensive patients, or up-titration of existing anti-hypertensive medication upon initiation of corticosteroid prescribing. Further research could evaluate the associated impact, of any prophylaxis prescribing changes, on health outcomes. Research could identify any demographic or spatio-temporal variation in guideline-recommended practices.

The studies of non-pharmacologic management showed the importance of smoking cessation support but not alcohol avoidance (beyond national limits) and the difficulties in monitoring cardiovascular risk in RA patients. Smoking cessation reduced both comorbidity risk and mortality, which is important given the higher comorbidity and mortality in RA patients (497, 498). The reported evidence supported a relaxation on advice for alcohol avoidance in the product characteristics for methotrexate and leflunomide (499, 500). Cardiovascular risk is raised in RA patients and annual monitoring is recommended in the national guidelines for RA management (280). As such, it is concerning that EHR-based CVD risk tools were reported to be less accurate for RA than non-RA patients, which was not improved by CRP measurements (496). Future research of non-pharmacologic management could explore change over time in the proportion receiving lifestyle advice and poor outcomes including cardiovascular events and knee replacement surgery, with comparison to a non-RA cohort. Regional and demographic variation in management practice could also be investigated in EHR-based research.

6.4.1 Strengths and Limitations

A strength of this thematic scoping review is that it was conducted in line with the PRISMA statement and evaluated the key study themes in EHR-based studies of RA (336). The breadth of synonyms and MeSH terms enhanced the sensitivity of the literature search and the search was applied to three literature databases. A conservative screening approach was adopted whereby abstracts were included for

full-text review if their eligibility was uncertain. The use of a study selection protocol and a data extraction form strengthened the systematic approach of the study.

The study limitations included that publications were excluded where a full text in English was not available, and the risk of publication bias in academic research. Further, the lack of a standardised term for EHRs may have resulted in some studies being not being identified in the literature database searches. Only studies using UK data were reviewed. This UK focus provided a useful review of the incidence, prevalence and management of RA in the UK, which is pertinent background for the subsequent chapters in this thesis, given that guideline and historic practice differences exist between countries. However, other factors of epidemiology as well as the pathogenesis of RA and its relationship to comorbidities, are unlikely to be country-specific and so relevant international studies in these areas may have been missed. As with the review of all AS EHR literature in Chapter 5, a second reviewer could improve the quality of the review and the study quality could have been assessed, although the RoB in EHR-based studies was deemed to be low when it was assessed in Chapter 4.

6.4.2 Conclusion

In the UK, a large number of EHR-based RA studies (n = 88) have been conducted, suggesting the success of UK-based initiatives such as the CPRD to facilitate access to EHR data for research, and the relevance of EHR data in research. These studies reported information on the epidemiology, comorbidities and management of RA, as well as showing high validity of RA diagnostic definitions used in EHR-based studies. Unlike in AS (Chapter 5), most studies used GP data and informed understanding of the diagnosis and management of RA and comorbidities in general practice. The studies identified the incidence and prevalence of RA although not in more recent calendar years, which the next chapter aims to address through an EHR-based study using CPRD data. Risk factors for RA and comorbidities in RA were investigated, highlighting a substantial cardiovascular burden in RA, with implications for mortality which require addressing. DMARD prescribing has increased following growing evidence of their efficacy in treating RA, though there was limited study of the trends in corticosteroid and NSAID prescribing and prophylaxis, which will be addressed in Chapter 9.

Chapter 7 The Epidemiology of Ankylosing Spondylitis and Rheumatoid Arthritis in the UK

7.1 Introduction

This chapter addresses the second thesis objective of describing the trends in disease epidemiology using EHR data. The uncertainty regarding the incidence and prevalence of AS in the UK was highlighted in Chapter 2 (190). In the same chapter, it was noted that a recent population-based study provided evidence of a decline in the incidence and prevalence of RA in the UK across 1990-2014 (20, 265); and yet payment for RA coding, introduced in 2013, may have led to changes in RA coding over time. Chapter 4 highlighted the importance of considering such changes in coding practice across the study time-frame (485). Further, these estimates of RA incidence and prevalence could be extended to more recent calendar years. It is also important to describe any spatio-temporal and socio-economic variation in the incidence and prevalence of AS and RA.

This chapter aims to investigate regional and demographic variation in the incidence and prevalence of AS and RA between 1998 and 2017, using GP EHR data from the CPRD GOLD as described in Chapter 3. This study provides information about the epidemiology of AS and RA in the UK, providing the context for further evaluation of the diagnosis and management of these diseases in the following chapters. In addition, it contributes to the understanding of the impact of payment tariffs on coding practices that needs to be considered in EHR-based research, with relevance for policy-making and the design of study methodology.

7.2 Methods

A population-based retrospective longitudinal observational study using GP EHR data was reported, following the STROBE guidelines (322). The study protocol approval, data source, study population, study timeframe and period of study follow-up were described in Chapter 3.

7.2.1 Ankylosing Spondylitis Cohort

The RCV2 code ‘N100.’, used to record AS diagnosis, was used to identify the AS cohort from the study population. Dubreuil et al. (2016) had previously validated the AS

code on GP practice data (72% positive predictive value; 89% for two AS codes >7 days apart) (62). Follow-up commenced on the date of diagnosis or study follow-up start (defined in Chapter 3), whichever was latest.

Sensitivity analyses were performed, given the variation between primary and secondary care-based estimates of AS incidence and prevalence that was highlighted in Chapter 2, and the potential for cohort specificity to impact on study findings that was identified in Chapter 4. These required the use of additional codes to improve diagnostic certainty. Sensitivity analyses AS1 and AS2 identified two sub-cohorts with an additional AS diagnostic or AS activity measurement code (N100., 2377., 388p., 388p0, 38QL) recorded >7 days (AS1) and ≥180 days (6 months; AS2) later. The medications prescribed in primary care for AS management are not disease-specific and so were not used in defining AS.

7.2.2 Rheumatoid Arthritis Cohort

Read Code Version 2 codes used to record RA, were identified and compared to those used in another CPRD study by Abhishek et al (2017) (265), with a code review performed by an epidemiologist with expertise in CPRD research (MP-R) (Appendix B: Table B 1). RA codes in CPRD have been previously validated (~80% positive predictive value) (264, 502). Patients with RA were identified by having ≥1 instance of these codes in their clinical data. Patients with an RA diagnosis while aged below 18 years, or juvenile RA diagnosis, were excluded.

As for the AS cohort, sensitivity analyses were also performed employing more specific diagnostic definitions. These excluded patients in which RA diagnosis was not confirmed or followed by an RA-specific prescription. Sensitivity analysis 1 (‘RA1’) identified patients with ≥2 RA diagnoses at least 6 months apart, to exclude patients with suspected RA diagnosis where diagnosis was subsequently discarded. Sensitivity analysis 2 (‘RA2’) selected patients with an RA diagnosis and a subsequent DMARD prescription (on or after the date of the patient’s first RA code, before April 2018). This definition is more specific because RA is the most common disease for which DMARDs are prescribed. DMARDs are recommended upon diagnosis of RA and the DMARD methotrexate is the common initial treatment for RA (503). The sensitivity analyses were based on approaches used in previous studies which showed that DMARD medication or multiple diagnosis codes increased the validity of RA diagnosis by GPs (458, 504). Serological tests were not used in defining RA because testing rates change over time and are often recorded in secondary rather than primary care.

DMARD prescriptions were identified for sensitivity analysis RA2 based on having a product name or drug substance name containing a term listed in Table 25 and one of the following administration routes: gastroenteral, intraarterial, intravenous, oral, subcutaneous. Term selection was informed by review of BNF section 8.2 'Drugs Affecting the Immune Response', the RA drug lists on the NHS and Versus Arthritis websites and the Yorkshire DMARD guidelines (304, 505-507).

Table 25. Drugs used to determine prescribed DMARDs

Medication
Abatacept; Adalimumab; Azathioprine; Baricitinib; Certolizumab; Ciclosporin / cyclosporine; Cyclophosphamid/e; Etanercept; Gold injections / injectable gold / sodium aurothiomalate; Golimumab; Hydroxychloroquine; Infliximab; Leflunomide; Methotrexate; Mycophenolate / mycophenalte mofetil; Penicillamine; Rituximab; Sarilumab; Sulfasalazine; Tocilizumab; Tofacitinib; Ustekinumab

Follow-up commenced on the date of RA diagnosis or study follow-up start (defined in Chapter 3), whichever was latest. However, in a sub-analysis of sensitivity analysis RA1, the date of the subsequent code was used to assign the date of RA diagnosis. This was to assess any pattern in the timing of the subsequent RA code ≥6 months after the first.

7.2.3 Outcomes

The outcomes were the period and annual incidence and prevalence of AS and RA, overall and stratified by sex, age-group and geographical area over two decades (1998-2017).

7.2.4 Statistical Analysis

Baseline cohort characteristics were described for the RA and AS cohorts, both for the full cohort (incident and prevalent cases) and incident cohort (patients diagnosed during follow-up). Outcome measures were stratified by sex (female, male), age-group (18-19 and 20-29 or 18-29 then 10-year bands up to 99) and geographical area (GP practice region as defined in CPRD (89)). Data were suppressed where there were ≤5 cases or patient representation from <5 GP practices in a geographic area. For age calculations, the year of birth recorded in CPRD was used with the day and month of birth set as 1 July. Age was calculated as on 1 July in a given year. Annual measures

were reported per calendar year between 1 January 1997 and 31 December 2017. Difference between calendar years was deemed significant where the 95% confidence intervals did not overlap.

Crude annual and period incidence rates of AS and RA were calculated per 10,000 person-years with 95% CIs for patients 'at-risk', i.e. having no diagnosis at the start of that time period and having ≥1 year of prior GP registration. This excluded any prevalent cases that were incorrectly recorded as an incident diagnosis instead of as medical history during GP registration (396). The follow-up duration of incident cases was divided by the total person-years of follow-up in the same time period. The annual percentage change (APC) was also calculated.

Crude point and period prevalence were calculated as percentages with 95% CI, the former being calculated on 1 July of each calendar year, with APC also calculated. For point prevalence calculations, the person-years of follow-up (denominator) was computed as follows:

Days of follow-up in a given year was calculated per patient:

1. 365.25 days if the start of follow-up was before or on 01 January and end of follow-up was on or after 31 December
2. Number of days between 01 January and end of follow-up (inclusive) if the start of follow-up was before or on 01 January and end of follow-up was before 31 December
3. Number of days between the beginning of follow-up and the end of follow-up (inclusive) if the start of follow-up was after 01 January and end of follow-up was before 31 December
4. Number of days between the beginning of follow-up and 31 December (inclusive) if the start of follow-up was after 01 January and end of follow-up was on or after 31 December

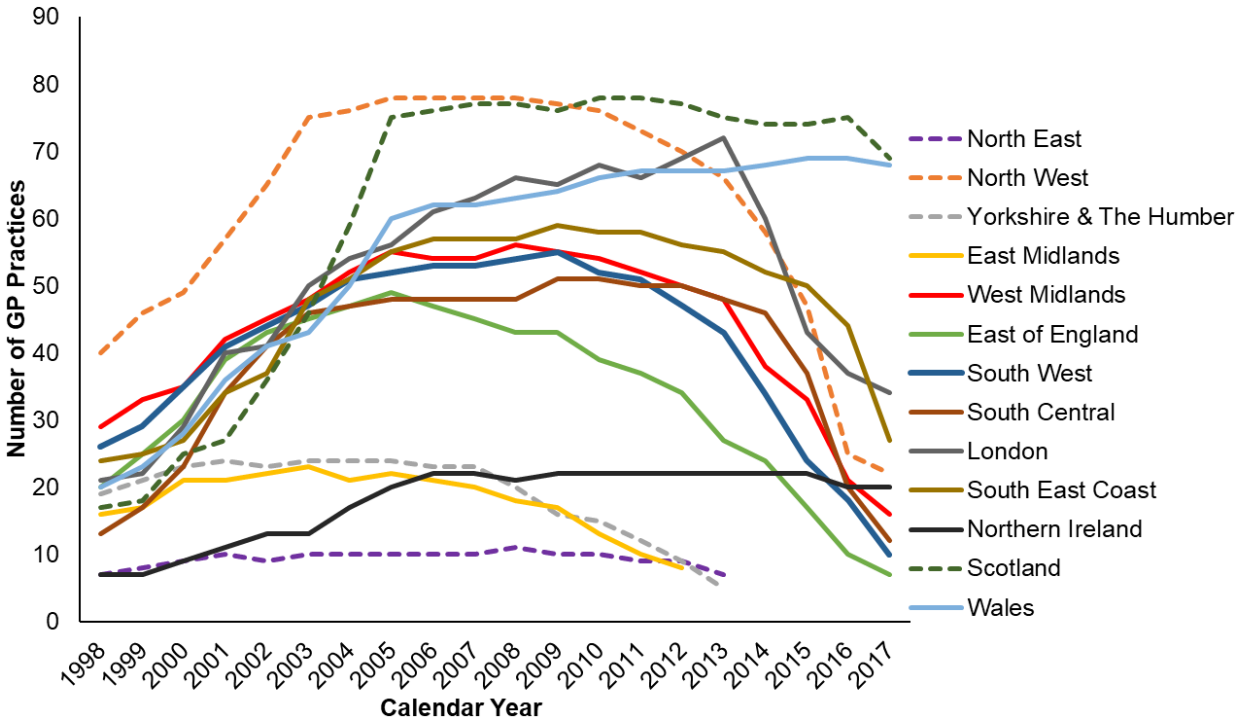
The days of follow-up for the cohort were summed and divided by 365.25 to calculate the total number of person-years of follow-up in that year.

In sensitivity analyses, the alternative definitions of AS and RA were applied for the 1998-2016 period to allow for follow-up in which the further diagnostic coding and prescribing could occur.

7.3 Results

The dataset included 707 practices that had ≥5 patients in the AS or RA cohorts. Of these, the number of practices per region varied over the study period (Figure 20).

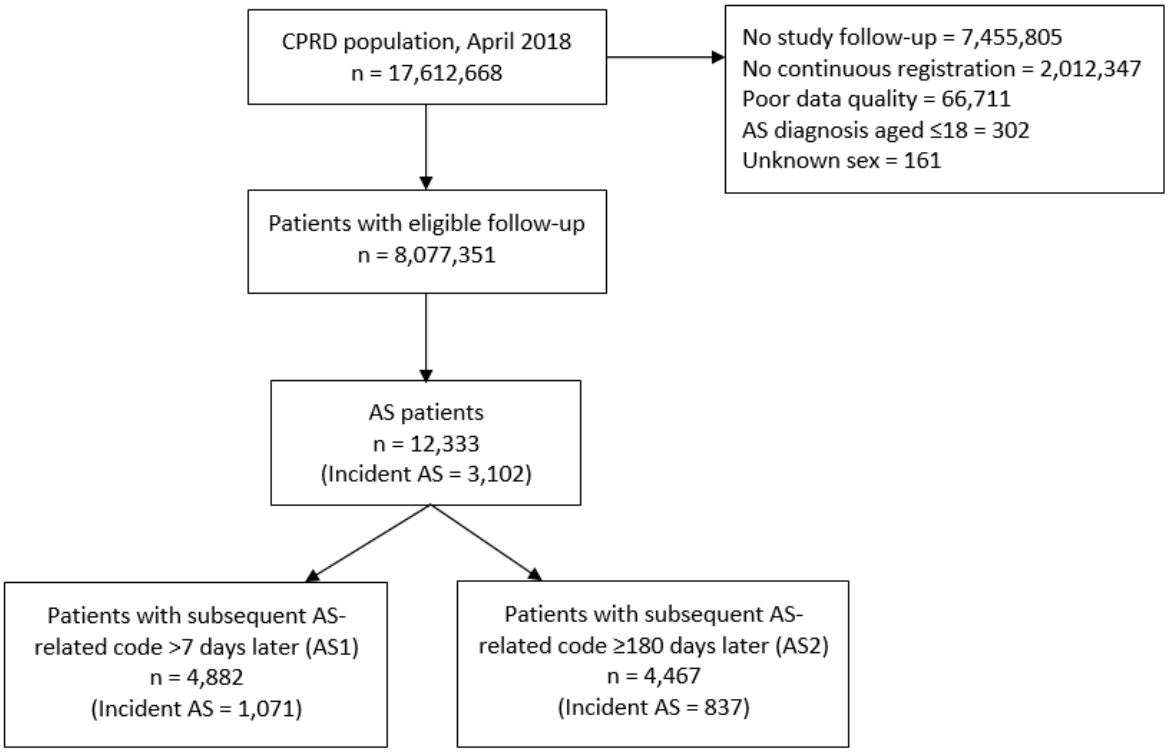
Figure 20. Annual count of GP practices with ≥5 AS or RA patients, per geographic region (N = 707 practices)



7.3.1 Ankylosing Spondylitis

A cohort of 12,333 AS patients was identified (excluding 302 patients diagnosed aged ≤18), 4,882 and 4,467 having a subsequent diagnostic or measurement code >7 and ≥180 days later (sensitivity analyses AS1 and AS2) (Figure 21). Of these, 4,850 and 4,435 had ≥2 AS codes (N100.) >7 and ≥180 days apart. The median duration of follow-up was 8.85 (interquartile range, IQR = 3.63-14.22) years.

Figure 21. Study flow diagram of cohort selection



In the full AS cohort (incident and prevalent cases), the proportion of women was 26.0% (n = 3,209) and was lower in sensitivity analyses. The median age at diagnosis was 36 (IQR = 28-47) years and was higher in women (38 [30-50]) than men (35 [28-46]). The age at diagnosis was lower, with a narrower sex gap in sensitivity analyses (Table 26). In the incident cohort (patients diagnosed during follow-up), the proportion of women was higher (30.8%, n = 954) and there was no sex difference in age at diagnosis.

Table 26. Number of AS patients, overall and diagnosed during follow-up, and median age at diagnosis, by sex

	All AS patients		Sensitivity analysis AS1		Sensitivity analysis AS2	
	Full cohort	**Incident cohort**	**Full cohort**	**Incident cohort**	**Full cohort**	**Incident cohort**
Patient count	12,333	3,102	4,882	1,071	4,467	837
Women (%)	3,209 (26.0)	954 (30.8)	1,095 (22.4)	299 (27.9)	994 (22.3)	246 (29.4)
Men (%)	9,124 (74.0)	2,148 (69.2)	3,787 (77.6)	772 (72.1)	3,473 (77.7)	591 (70.6)
Median age (Q1, Q3)	36 (28-47)	43 (33-56)	34 (27-44)	40 (32-51)	34 (27-43)	40 (32-50)
Women (%)	38 (30-50)	43 (33-55)	35 (28-46)	40 (33-51)	35 (28-45)	40 (33-51)
Men (%)	35 (28-46)	43 (33-56)	34 (27-43)	40 (31-50)	34 (27-42)	40 (31-49)

Note: Q1 = quartile 1, Q3 = quartile 3

7.3.1.1 Incidence

The period incidence of AS was 0.54 (±0.02) per 10,000 person-years; 0.19 (±0.01) and 0.15 (±0.01) in sensitivity analyses AS1 and AS2. The person-years in each analysis were 57,539,311; 55,859,034; and 55,861,261 (Figure 22). Incidence was greatest among patients aged 30-39 (0.80 [±0.08]; 0.33 [±0.04]; and 0.26 [±0.03]) and 2.3 times higher in men (0.76 [±0.03]; 0.21 [±0.02]; and 0.21 [±0.02]) than women (0.33 [±0.02]; 0.09 [±0.06]; and 0.09 [±0.01]) (Table 27).

Figure 22. Person-years (million) in the incidence 'at-risk' cohort per year, 1998-2017 (N = 8,052,546)

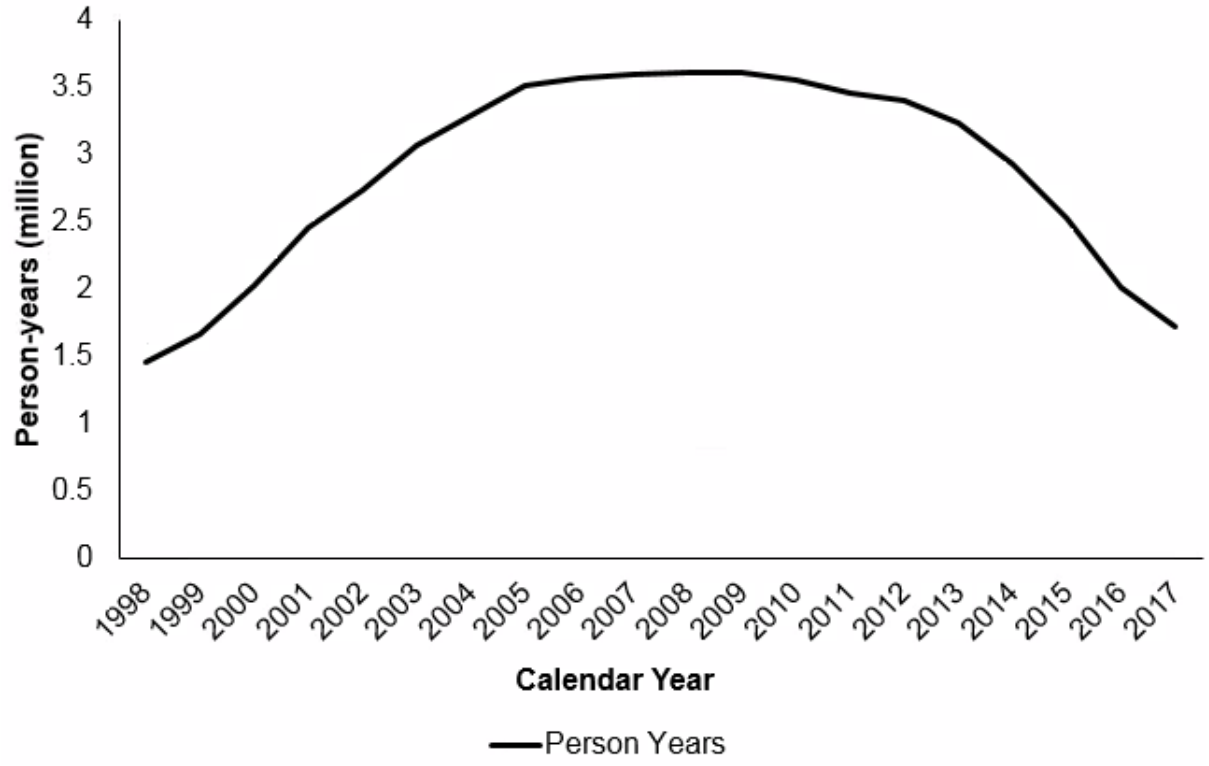


Table 27. Incidence of AS by calendar year and stratified by sex, age-group and geographical area (N = 8,052,546 in the main analysis, N = 7,919,770 in sensitivity analysis AS1, N = 7,918,922 in sensitivity analysis AS2)

	Number of events (person-years at risk / 10,000)			Incidence rate per 10,000 person-years (±95% CI)		
	Main analysis	**Sensitivity analysis AS1**	**Sensitivity analysis AS2**	**Main analysis**	**Sensitivity analysis AS1**	**Sensitivity analysis AS2**
Overall	3,102 (5,753.9)	1,071 (5,585.9)	845 (5,586.1)	0.54 (0.02)	0.19 (0.01)	0.15 (0.01)
Year						
1998	106 (146.3)	50 (146.4)	45 (146.4)	0.72 (0.14)	0.34 (0.09)	0.31 (0.09)
1999	119 (166.9)	58 (167.0)	49 (167.0)	0.71 (0.13)	0.35 (0.09)	0.29 (0.08)
2000	125 (202.6)	61 (202.8)	50 (202.8)	0.62 (0.11)	0.30 (0.08)	0.25 (0.07)
2001	142 (246.3)	64 (246.5)	53 (246.5)	0.58 (0.09)	0.26 (0.06)	0.22 (0.6)
2002	145 (273.5)	61 (273.7)	49 (273.7)	0.53 (0.09)	0.22 (0.06)	0.18 (0.05)
2003	181 (307.7)	81 (308.0)	73 (308.0)	0.59 (0.09)	0.26 (0.06)	0.24 (0.05)
2004	202 (329.6)	62 (329.9)	56 (329.9)	0.61 (0.08)	0.19 (0.05)	0.17 (0.04)
2005	177 (351.6)	74 (351.9)	59 (351.9)	0.50 (0.07)	0.21 (0.05)	0.17 (0.04)
2006	186 (357.5)	79 (357.8)	64 (357.8)	0.52 (0.07)	0.22 (0.05)	0.18 (0.04)
2007	140 (359.3)	46 (359.6)	39 (359.6)	0.39 (0.06)	0.13 (0.04)	0.11 (0.03)
2008	183 (361.1)	73 (361.4)	47 (361.4)	0.51 (0.07)	0.20 (0.05)	0.13 (0.04)
2009	176 (361.6)	68 (362.0)	55 (362.0)	0.49 (0.07)	0.19 (0.04)	0.15 (0.04)
2010	183 (356.0)	60 (356.3)	45 (356.3)	0.51 (0.07)	0.17 (0.04)	0.13 (0.04)
2011	166 (346.7)	45 (347.0)	36 (347.0)	0.48 (0.07)	0.13 (0.04)	0.10 (0.03)
2012	167 (341.1)	54 (341.5)	37 (341.5)	0.49 (0.07)	0.16 (0.04)	0.11 (0.03)
2013	198 (323.9)	63 (324.2)	48 (324.2)	0.61 (0.09)	0.19 (0.05)	0.15 (0.04)

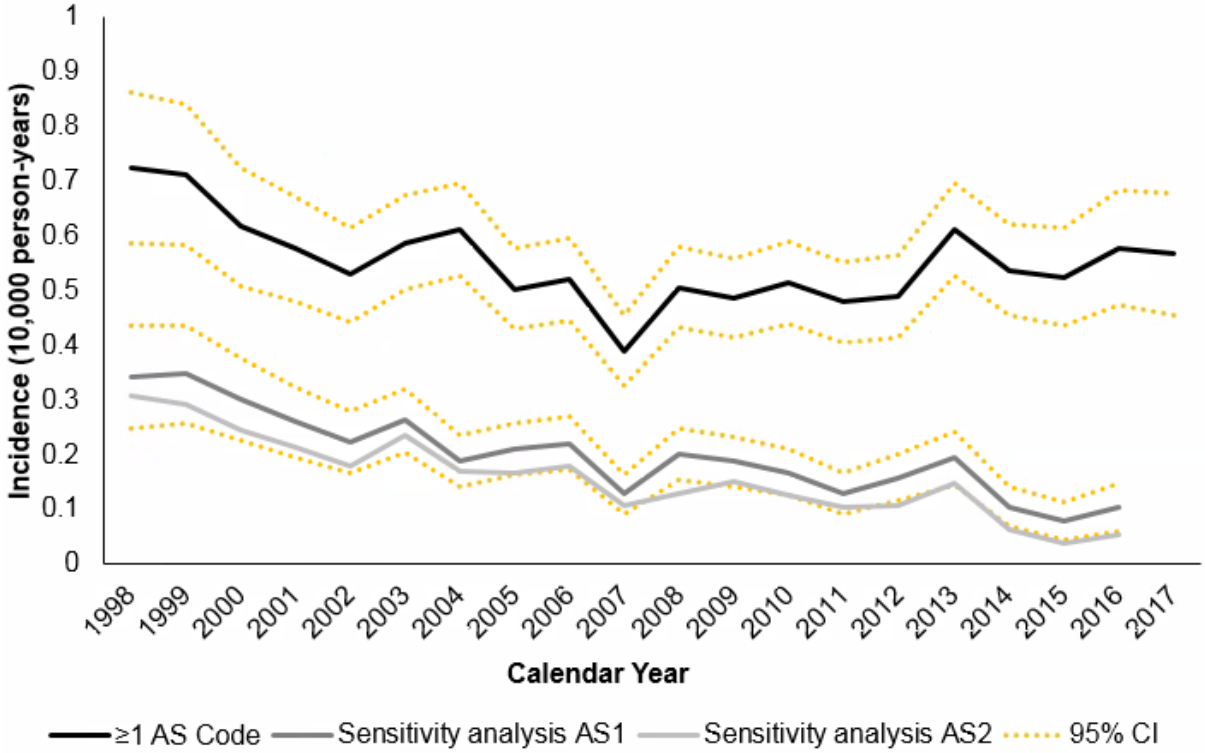
2014	158 (293.8)	31 (294.1)	19 (294.1)	0.54 (0.08)	0.11 (0.04)	0.06 (0.03)
2015	133 (253.2)	20 (253.5)	10 (253.5)	0.53 (0.08)	0.08 (0.03)	0.04 (0.02)
2016	117 (202.3)	21 (202.5)	11 (202.5)	0.58 (0.10)	0.10 (0.04)	0.05 (0.03)
2017	98 (172.9)			0.57 (0.11)		
Sex						
Female	954 (2,919.6)	299 (2,833.3)	246 (2,833.4)	0.33 (0.02)	0.09 (0.01)	0.09 (0.01)
Male	2,148 (2,834.3)	772 (2,752.6)	599 (2,752.7)	0.76 (0.03)	0.21 (0.02)	0.21 (0.02)
Age-group						
18-19	31 (115.7)	11 (112.2)	7 (112.2)	0.27 (0.09)	0.09 (0.06)	0.06 (0.05)
20-29	476 (773.7)	186 (750.6)	150 (750.6)	0.62 (0.06)	0.24 (0.04)	0.19 (0.03)
30-39	792 (987.7)	318 (960.2)	253 (960.3)	0.80 (0.06)	0.33 (0.04)	0.26 (0.03)
40-49	704 (1,102.1)	264 (1,072.2)	219 (1,072.2)	0.64 (0.05)	0.25 (0.03)	0.20 (0.03)
50-59	507 (999.4)	183 (968.4)	150 (968.4)	0.51 (0.04)	0.19 (0.03)	0.15 (0.03)
60-69	333 (814.7)	68 (790.9)	48 (790.9)	0.41 (0.04)	0.09 (0.02)	0.06 (0.02)
70-79	188 (581.4)	32 (563.5)	16 (563.5)	0.32 (0.05)	0.06 (0.02)	0.03 (0.01)
80-89	67 (313.1)	9 (218.9)	≤5	0.21 (0.05)	0.04 (0.03)	
90-99	≤5	≤5	≤5			
Geographical area						
North East	53 (91.1)	25 (91.2)	18 (91.2)	0.58 (0.16)	0.27 (0.11)	0.20 (0.09)
North West	385 (648.3)	155 (637.5)	125 (637.6)	0.59 (0.06)	0.24 (0.04)	0.20 (0.03)
Yorkshire & The Humber	101 (173.8)	43 (174.0)	33 (174.0)	0.56 (0.11)	0.25 (0.07)	0.18 (0.06)
East Midlands	82 (186.2)	25 (186.4)	18 (186.4)	0.44 (0.10)	0.13 (0.05)	0.10 (0.04)
West Midlands	260 (542.8)	98 (531.2)	74 (531.2)	0.48 (0.06)	0.18 (0.04)	0.14 (0.03)

East of England	275 (472.3)	92 (466.8)	70 (466.8)	0.58 (0.07)	0.20 (0.04)	0.15 (0.03)
South West	269 (485.1)	88 (478.0)	69 (478.0)	0.55 (0.07)	0.18 (0.04)	0.14 (0.03)
South Central	307 (596.3)	123 (586.6)	93 (586.7)	0.51 (0.06)	0.21 (0.04)	0.15 (0.03)
London	271 (541.2)	74 (524.2)	58 (524.2)	0.50 (0.06)	0.14 (0.03)	0.11 (0.03)
South East Coast	322 (552.6)	84 (533.7)	73 (533.7)	0.58 (0.06)	0.16 (0.03)	0.13 (0.03)
Northern Ireland	149 (196.7)	65 (186.5)	57 (186.5)	0.76 (0.12)	0.35 (0.08)	0.31 (0.08)
Scotland	312 (598.1)	92 (562.0)	74 (562.1)	0.52 (0.06)	0.16 (0.03)	0.13 (0.03)
Wales	308 (656.2)	107 (616.7)	83 (616.7)	0.47 (0.05)	0.17 (0.03)	0.13 (0.03)

Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later

Incidence declined significantly from 0.72 (±0.14) in 1998 to 0.39 (±0.06) in 2007, then rose significantly to 0.57 (±0.11) in 2017 (overall mean APC -0.44) (Figure 23). The initial decline was steeper in sensitivity analyses AS1 and AS2 (0.34 [±0.09] and 0.31 [±0.09] in 1998, 0.13 [±0.04] and 0.11 [±0.03] in 2007) but stabilised thereafter (0.10 [±0.04] and 0.05 [±0.03] in 2016) (overall mean APC: -2.91; -4.71). The mean APC across five-year intervals was -7.42 in 1998-2002, -4.90 in 2003-2007, +5.42 in 2008-2012 and +3.72 in 2012-2017. In sensitivity analyses, this was -9.89 and -12.92 in 1998-2002, -7.14 and -5.57 in 2003-2007, +1.19 and +7.90 in 2008-2012 and +1.07 - 1.32 in 2012-2016. The initial decline was significant among men but not women (1.12 [±0.25] and 0.35 [±0.13] in 1998, 0.59 [±0.11] and 0.20 [±0.06] in 2007, 0.79 [±0.19]), then both sexes showed a rising trend (0.79 [±0.19] and 0.35 [±0.12] in 2017) (Figure 24). In sensitivity analysis AS1 the initial decline was significant in men and women (0.55 [±0.17] and 0.15 [±0.09] in 1998, 0.19 [±0.07] and 0.06 [±0.04] in 2007) with no significant change thereafter (Figure 25). Sensitivity analysis AS2 had comparable trends but the decline in women was not significant. Incidence was stable among patients aged ≥60 and declined in younger patients until 2007, before rising (Figure 26). Sensitivity analyses showed similar results although incidence stabilised from 2007 in younger patients (Figure 27). There was no clear regional trend (Figure 28, Figure 29).

Figure 23. Annual incidence rate of AS defined as having ≥1 AS diagnostic code, 1998-2017 (N = 8,052,546), and in sensitivity analyses AS1 and AS2, 1998-2016 (N = 7,919,770; N = 7,918,922)



Note: 95% confidence intervals (CI) for sensitivity analysis AS2 are not shown, for clarity. AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later

Figure 24. Annual incidence rate of AS in women and men, 1998-2017 (N = 8,052,546)

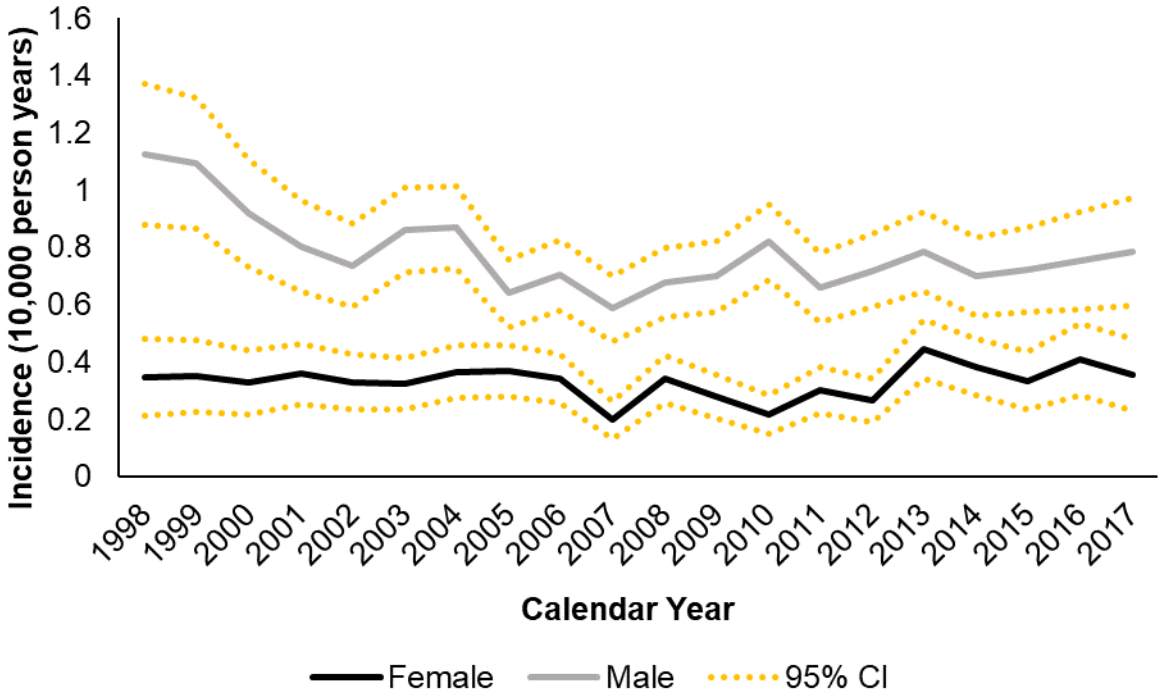
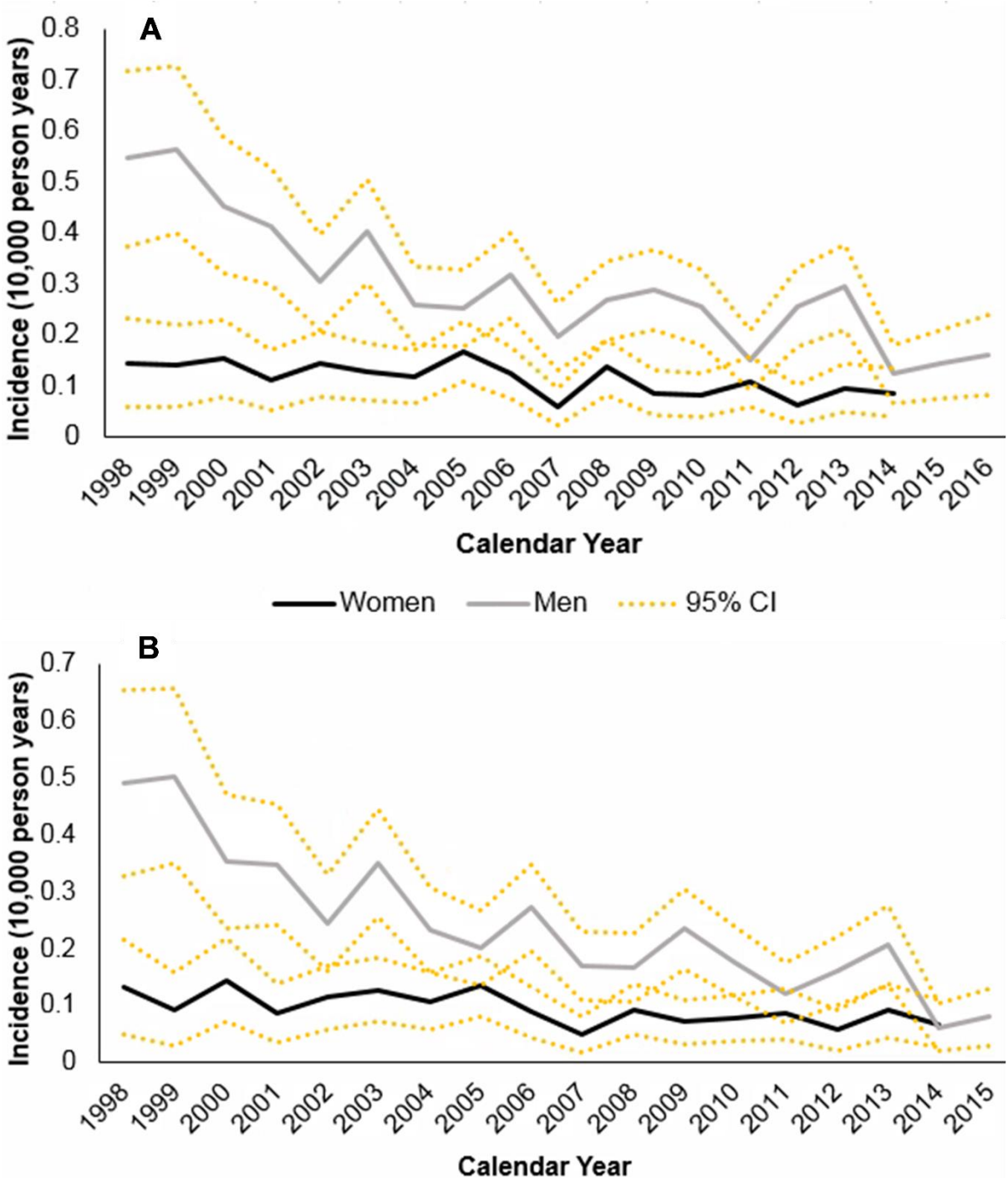
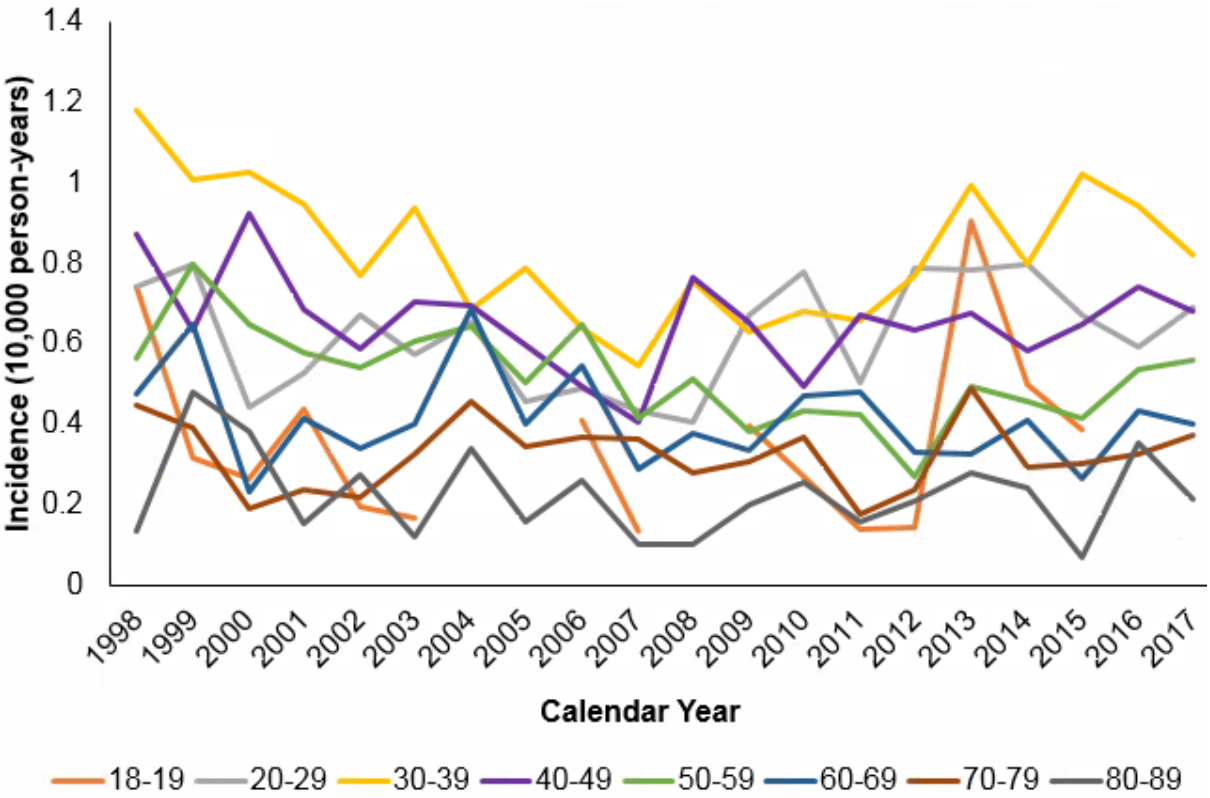


Figure 25. Annual incidence rate of AS in women and men in sensitivity analyses, 1998-2016: A) AS1 (N = 7,919,770); B) AS2 (N = 7,918,922)



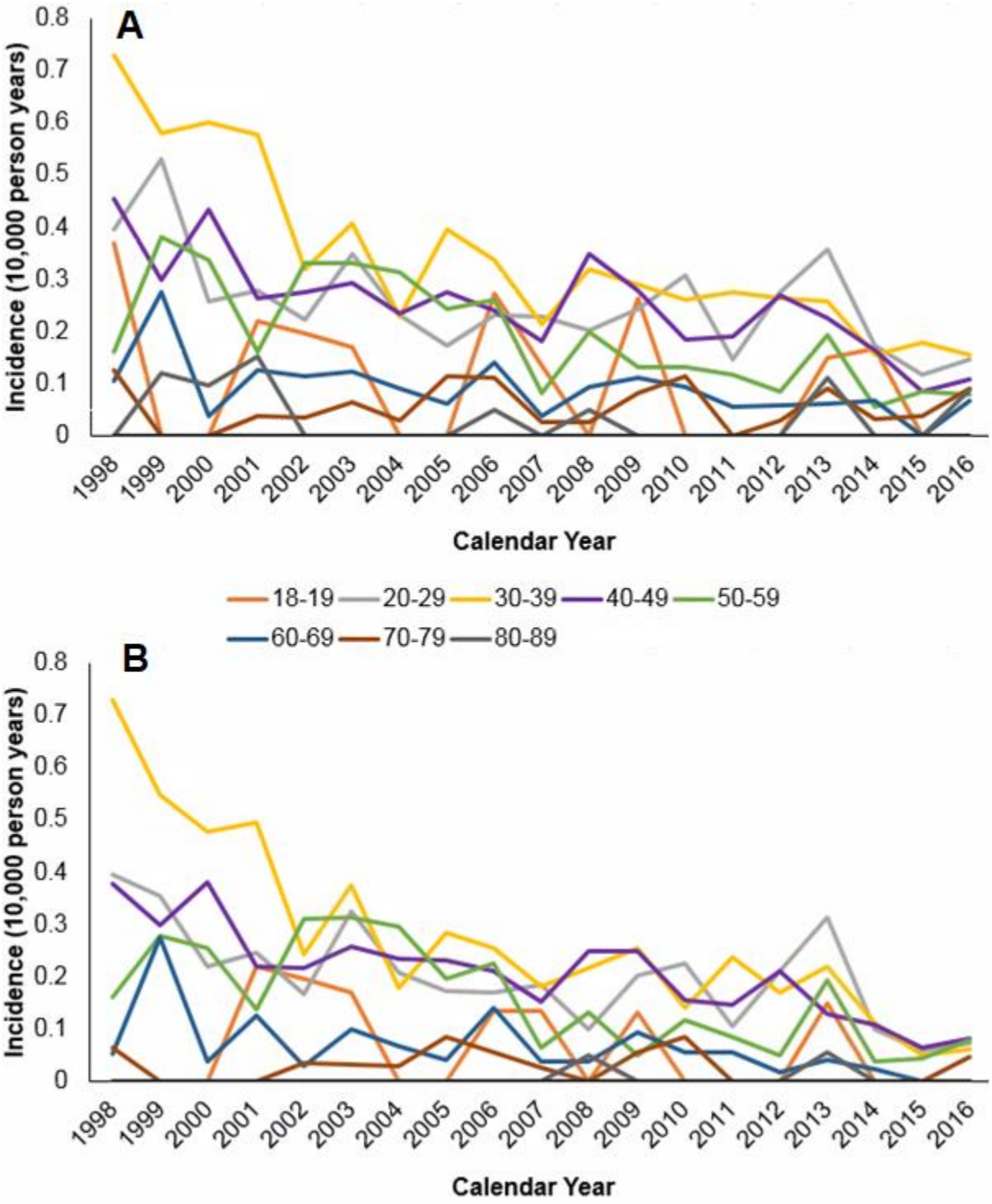
Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later. Data display is suppressed where ≤5 cases

Figure 26. Annual incidence rate of AS by age-group, 1998-2017 (N = 8,051,097)



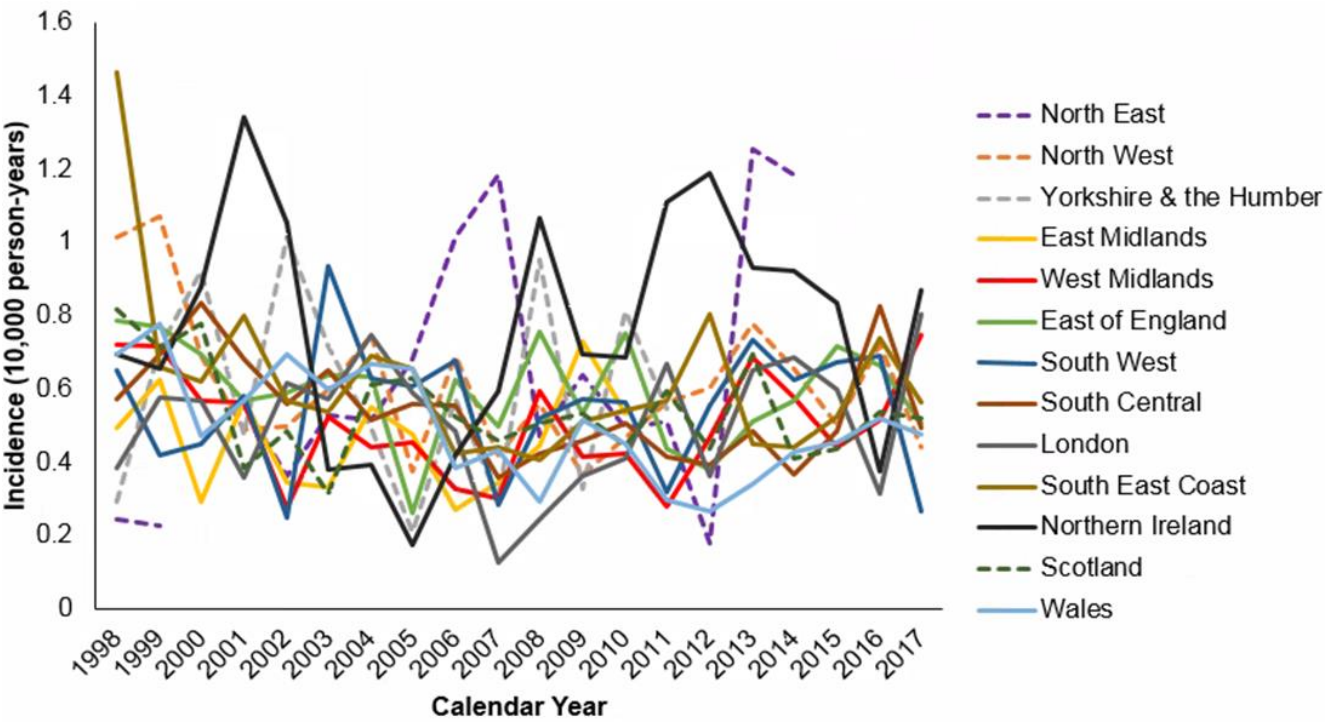
Note: 90-99 is suppressed (≤5 cases)

Figure 27. Annual incidence rate of AS in the sensitivity analyses by age-group, 1998-2016: A) AS1 (N = 7,918,339); B) AS2 (N = 7,917,491)



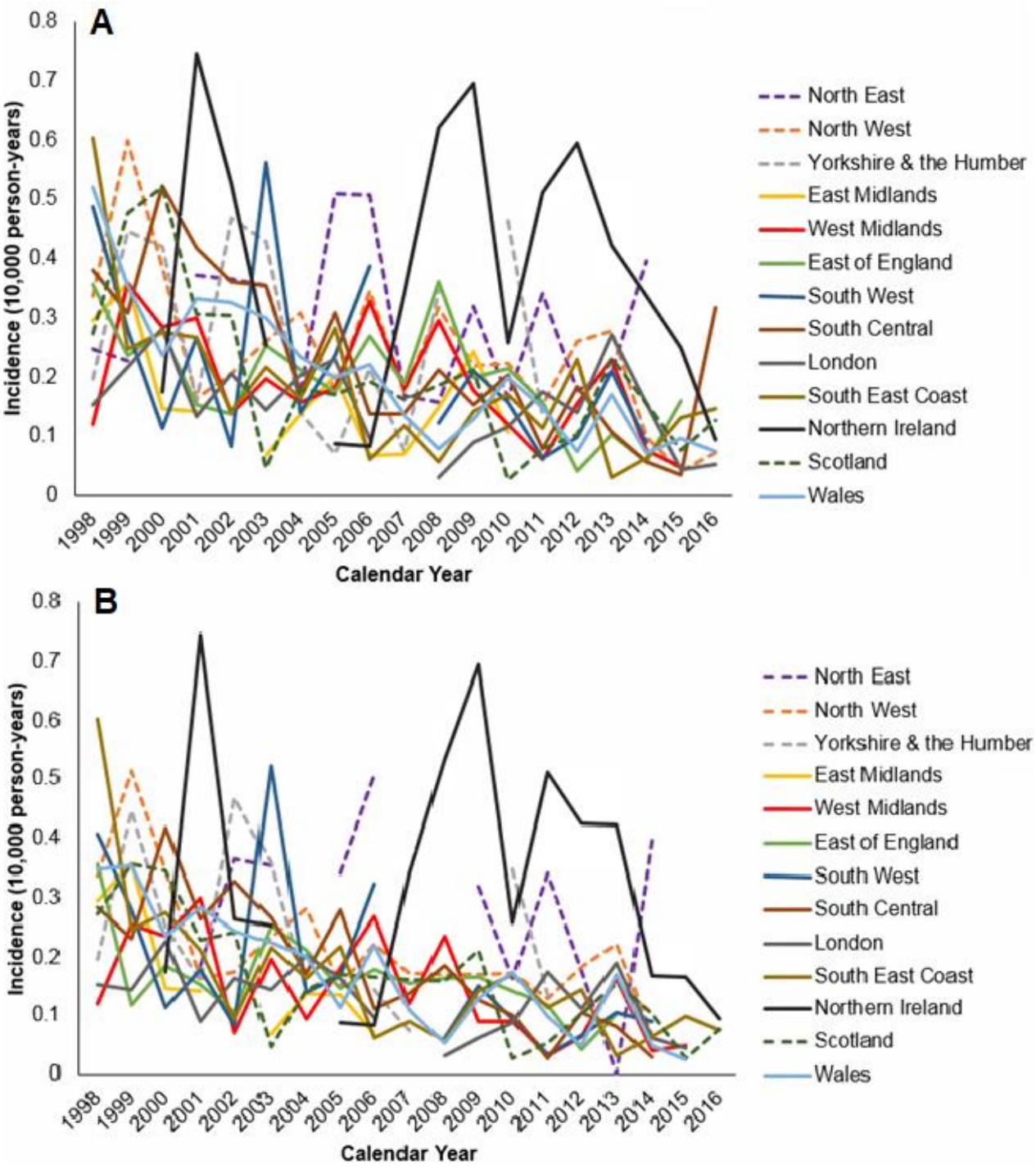
Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later; 90-99 is suppressed (≤5 cases)

Figure 28. Annual incidence rate of AS by geographic region, 1998-2017 (N = 8,044,388)



Note: Data display is suppressed where ≤5 cases

Figure 29. Annual incidence rate of AS in the sensitivity analyses by geographic region, 1998-2016: A) AS1 (N = 7,913,069); B: AS2 (N = 7,912,222)



Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later. Data display is suppressed where ≤5 cases

7.3.1.2 Prevalence

The period prevalence of AS was 0.15% (±0.003); 0.06 (±0.002) in sensitivity analysis AS1 and AS2. Prevalence was greatest among patients aged 60-69 (0.26 [±0.008]) and 2.9 times higher among men (0.23 [±0.005]) than women (0.08 [±0.003]).

Prevalence was lowest in London (0.11 [±0.01]). Patterns were consistent in the sensitivity analyses (Table 28).

Table 28. Percentage prevalence of AS by calendar year and stratified by sex, age-group and geographical area (N = 7,532,980 in the main analysis, N = 7,413,674 in sensitivity analysis AS1, N = 7,412,859 in sensitivity analysis AS2)

	Percentage prevalence (±95% CI)		
	Main analysis	**Sensitivity analysis AS1**	**Sensitivity analysis AS2**
Overall	0.15 (0.003)	0.06 (0.002)	0.06 (0.002)
Year			
1998	0.13 (0.006)	0.07 (0.004)	0.07 (0.004)
1999	0.14 (0.006)	0.07 (0.004)	0.07 (0.004)
2000	0.14 (0.005)	0.07 (0.004)	0.07 (0.004)
2001	0.14 (0.005)	0.07 (0.003)	0.07 (0.003)
2002	0.15 (0.005)	0.07 (0.003)	0.07 (0.003)
2003	0.15 (0.004)	0.07 (0.003)	0.07 (0.003)
2004	0.16 (0.004)	0.07 (0.003)	0.07 (0.003)
2005	0.16 (0.004)	0.07 (0.003)	0.07 (0.003)
2006	0.16 (0.004)	0.07 (0.003)	0.07 (0.003)
2007	0.16 (0.004)	0.07 (0.003)	0.07 (0.003)
2008	0.17 (0.004)	0.07 (0.003)	0.07 (0.003)
2009	0.17 (0.004)	0.08 (0.003)	0.07 (0.003)
2010	0.17 (0.004)	0.08 (0.003)	0.07 (0.003)
2011	0.17 (0.004)	0.08 (0.003)	0.07 (0.003)
2012	0.17 (0.004)	0.08 (0.003)	0.07 (0.003)
2013	0.17 (0.005)	0.08 (0.003)	0.07 (0.003)
2014	0.17 (0.005)	0.08 (0.003)	0.07 (0.003)
2015	0.18 (0.005)	0.08 (0.003)	0.07 (0.003)
2016	0.18 (0.006)	0.08 (0.003)	0.07 (0.003)
2017	0.18 (0.006)	0.07 (0.004)	0.07 (0.004)
Sex			
Female	0.08 (0.003)	0.03 (0.002)	0.03 (0.002)
Male	0.23 (0.005)	0.10 (0.003)	0.09 (0.003)
Age-group			
18-19	0.001 (0.001)	0.001 (0.0003)	0.001 (0.0003)
20-29	0.03 (0.001)	0.02 (0.001)	0.01 (0.003)

30-39	0.11 (0.002)	0.06 (0.002)	0.05 (0.004)
40-49	0.18 (0.003)	0.09 (0.002)	0.09 (0.005)
50-59	0.23 (0.003)	0.11 (0.002)	0.11 (0.005)
60-69	0.26 (0.003)	0.11 (0.002)	0.10 (0.005)
70-79	0.21 (0.004)	0.07 (0.002)	0.07 (0.004)
80-89	0.14 (0.004)	0.03 (0.002)	0.03 (0.004
90-99	0.09 (0.007)	0.01 (0.003	0.01 (0.002)
Geographical area			
North East	0.16 (0.02)	0.09 (0.02)	0.08 (0.02)
North West	0.17 (0.01)	0.07 (0.01)	0.07 (0.01)
Yorkshire & The Humber	0.15 (0.02)	0.07 (0.01)	0.06 (0.01)
East Midlands	0.15 (0.02)	0.06 (0.01)	0.06 (0.01)
West Midlands	0.14 (0.01)	0.06 (0.01)	0.06 (0.01)
East of England	0.16 (0.01)	0.07 (0.01)	0.06 (0.01)
South West	0.17 (0.01)	0.07 (0.01)	0.06 (0.01)
South Central	0.16 (0.01)	0.07 (0.01)	0.06 (0.01)
London	0.11 (0.01)	0.04 (0.004)	0.04 (0.004)
South East Coast	0.17 (0.01)	0.06 (0.01)	0.06 (0.01)
Northern Ireland	0.18 (0.02)	0.08 (0.01)	0.08 (0.01)
Scotland	0.17 (0.01)	0.05 (0.01)	0.05 (0.01)
Wales	0.15 (0.01)	0.06 (0.01)	0.05 (0.01)

Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later

Prevalence rose significantly from 0.13% (±0.01) in 1997 to 0.18% (±0.01) in 2017; more steeply in women than men (mean APC +2.69 and +1.07) (Figure 30, Figure 31). Annual prevalence was stable in sensitivity analyses (0.07% [±0.004] in 1998 and 2016), although there was a rising trend in women (mean APC +1.43 in AS1; +1.41 in AS2). In patients aged <60, annual prevalence was stable but declined in sensitivity analyses (Figure 32, Figure 33). Prevalence rose in older cohorts, notably in patients aged 70-79 (0.12 [±0.02] in 1998, 0.30 [±0.02] in 2017). There was little regional variation over time (Figure 34, Figure 35).

Figure 30. Annual percentage prevalence of AS, 1998-2017 (N = 8,052,980), and in sensitivity analyses AS1 and AS2, 1998-2016 (N = 7,413,674; N = 7,412,859)

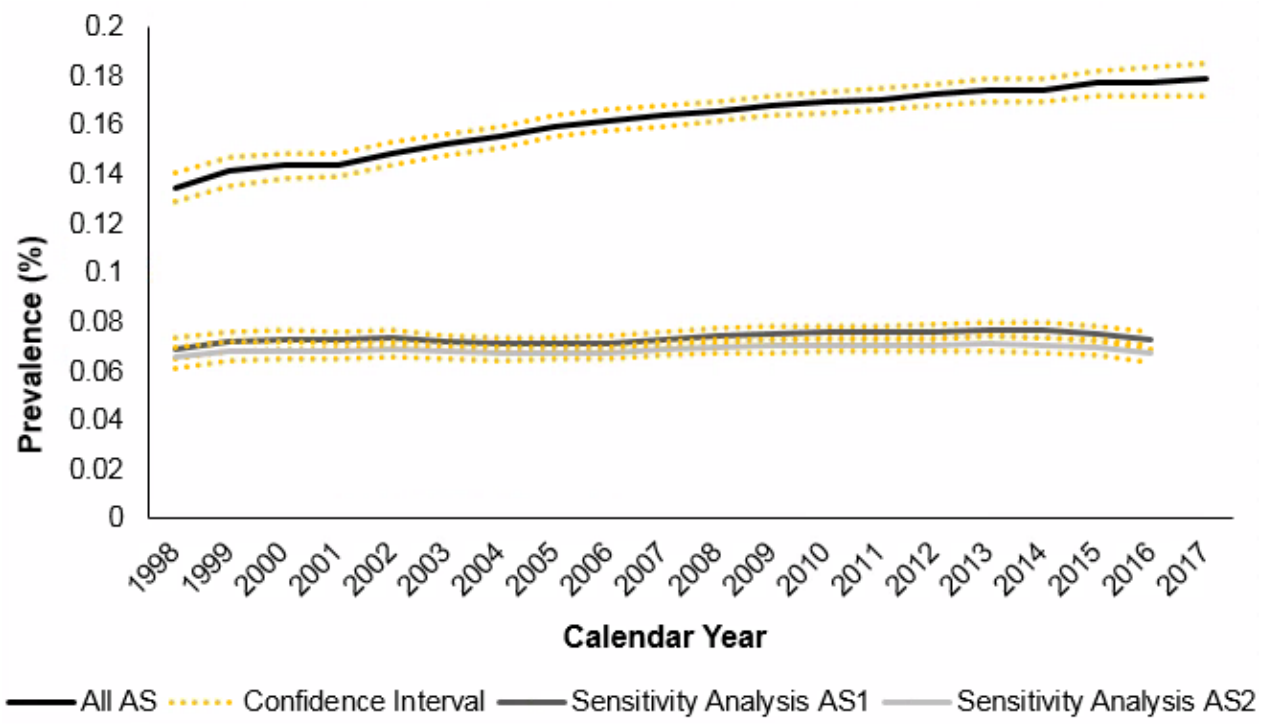
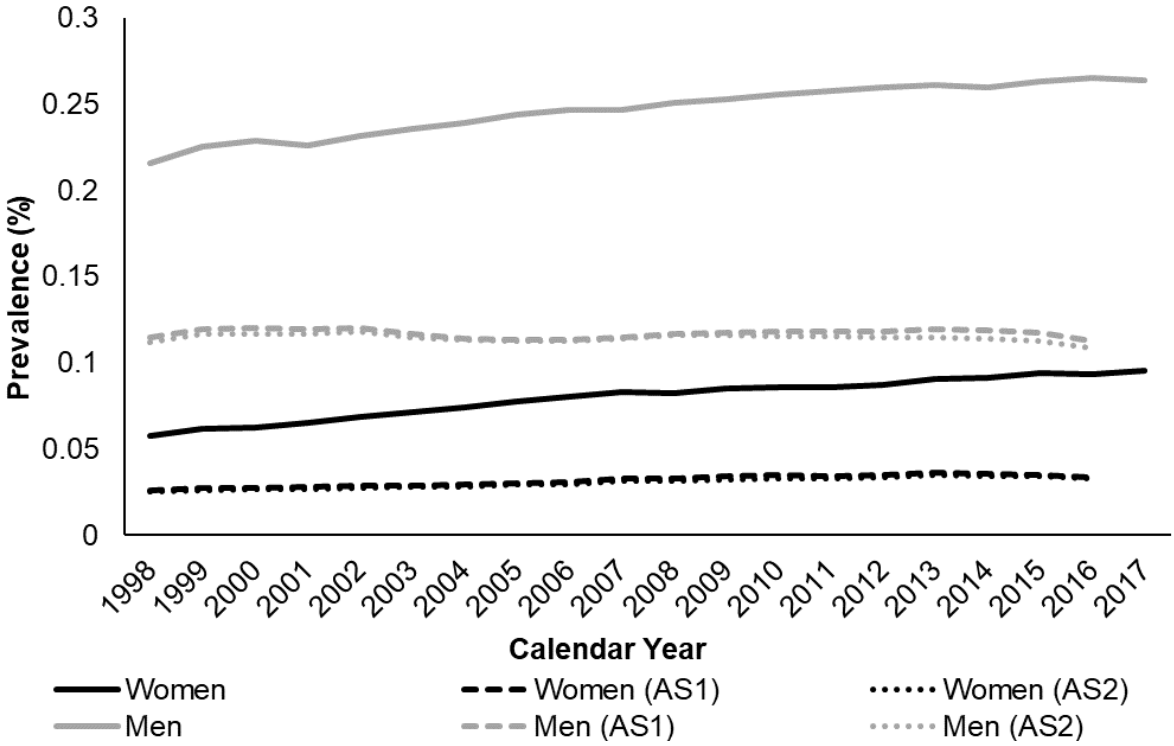


Figure 31. Annual percentage prevalence of AS in women and men, 1998-2017 (N = 8,052,980), and in sensitivity analyses AS1 and AS2, 1998-2016 (N = 7,413,674; N = 7,412,859)



Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later

Figure 32. Annual percentage prevalence of AS per age-group in patients aged 18-99, 1997-2017 (N = 7,532,700)

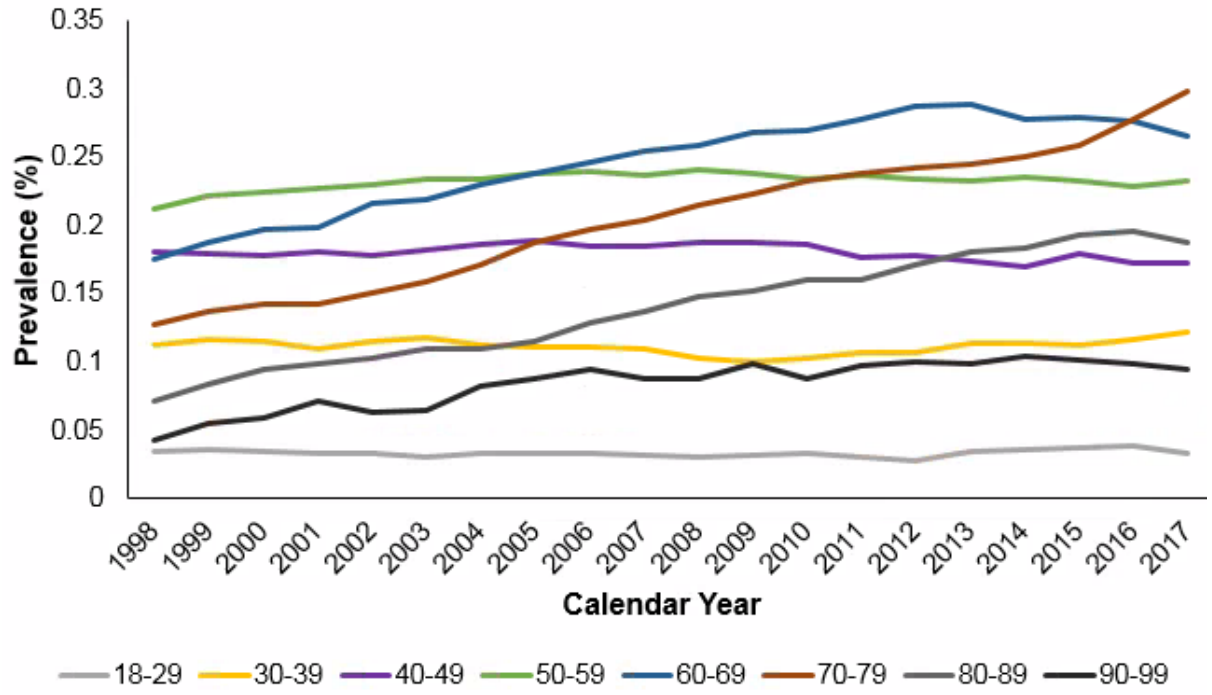
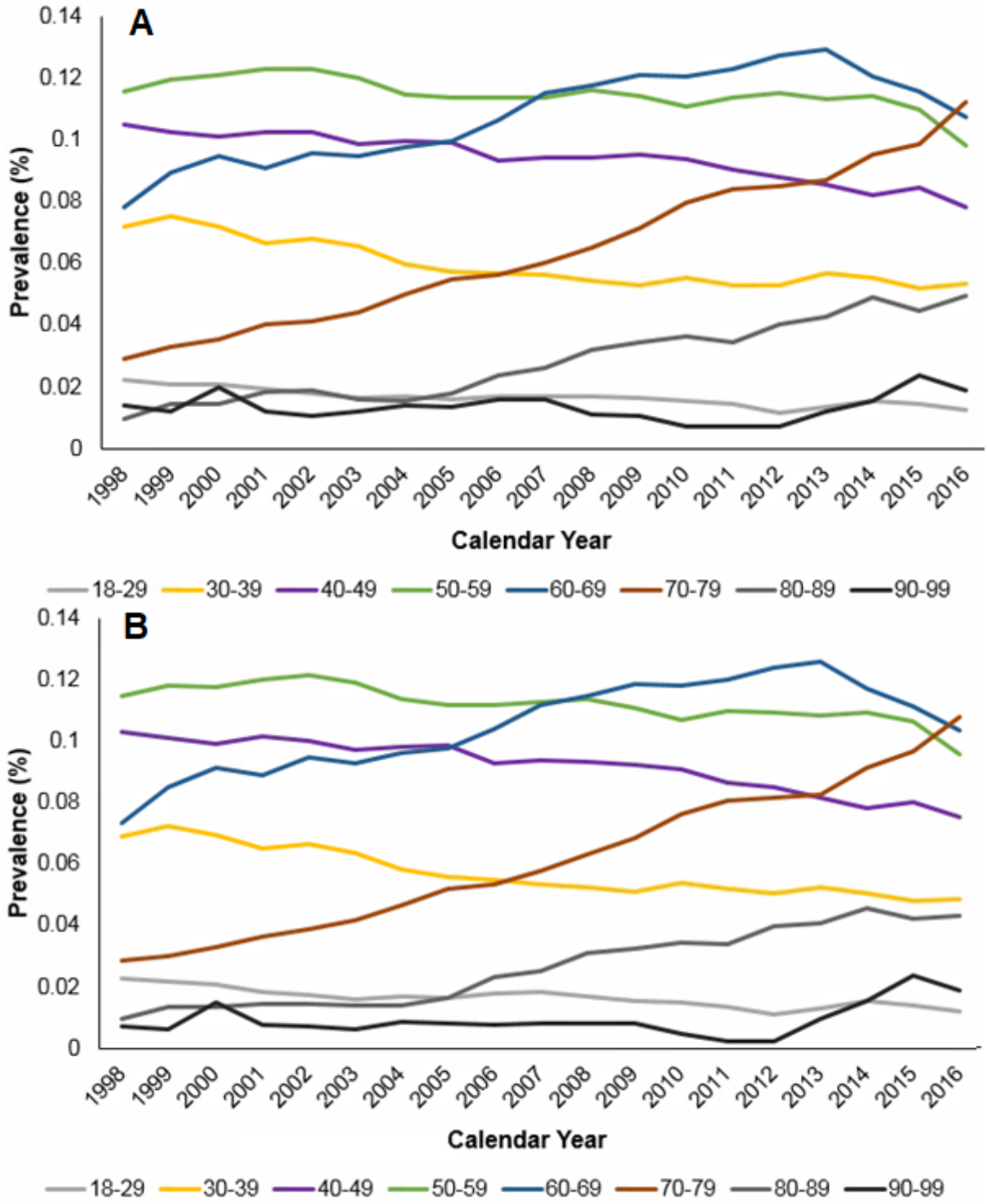


Figure 33. Annual percentage prevalence of AS per age-group in patients aged 18-99 in sensitivity analyses, 1997-2016: A) AS1 (N = 7, 413,674)); B) AS2 (N = 7,412,859)



Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later

Figure 34. Annual percentage prevalence of AS per geographic region, 1997-2017 (N = 7,522,334)

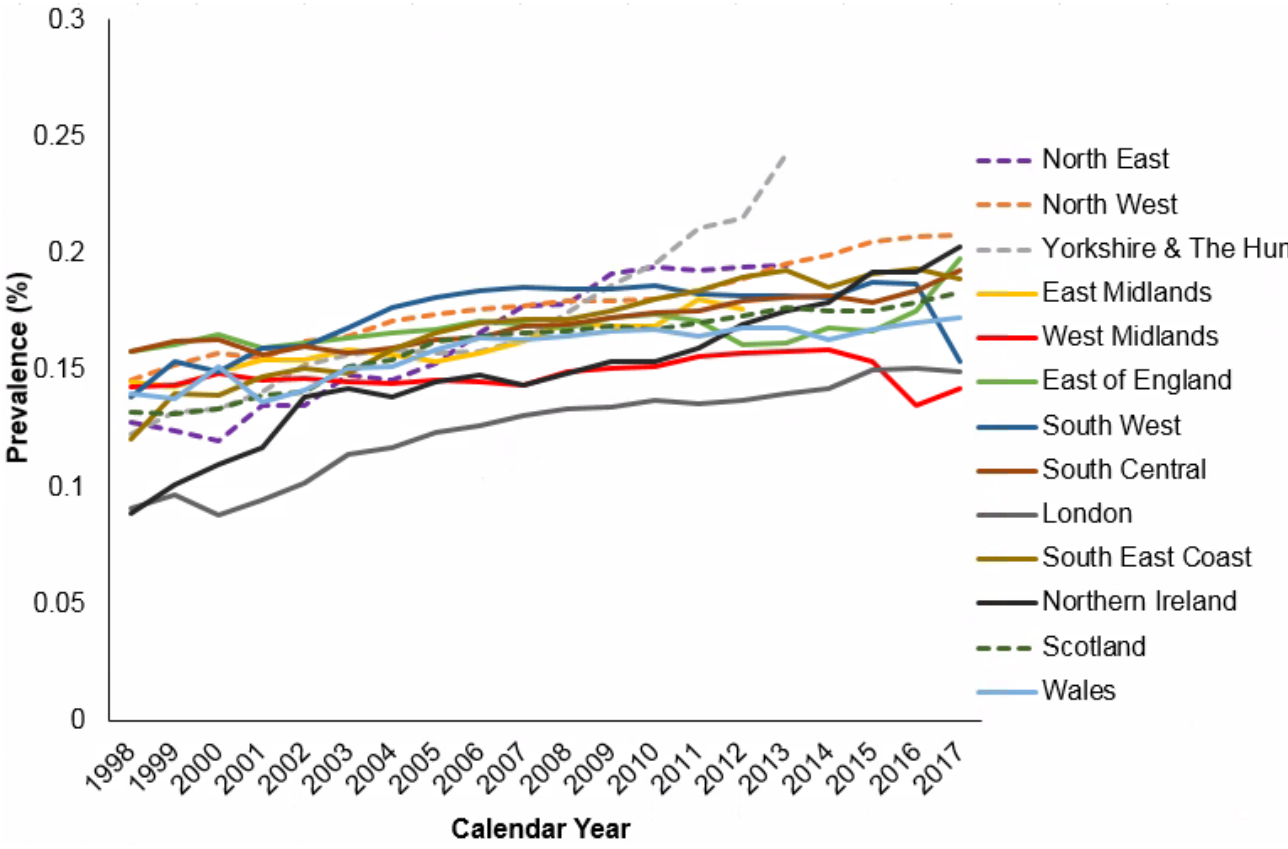
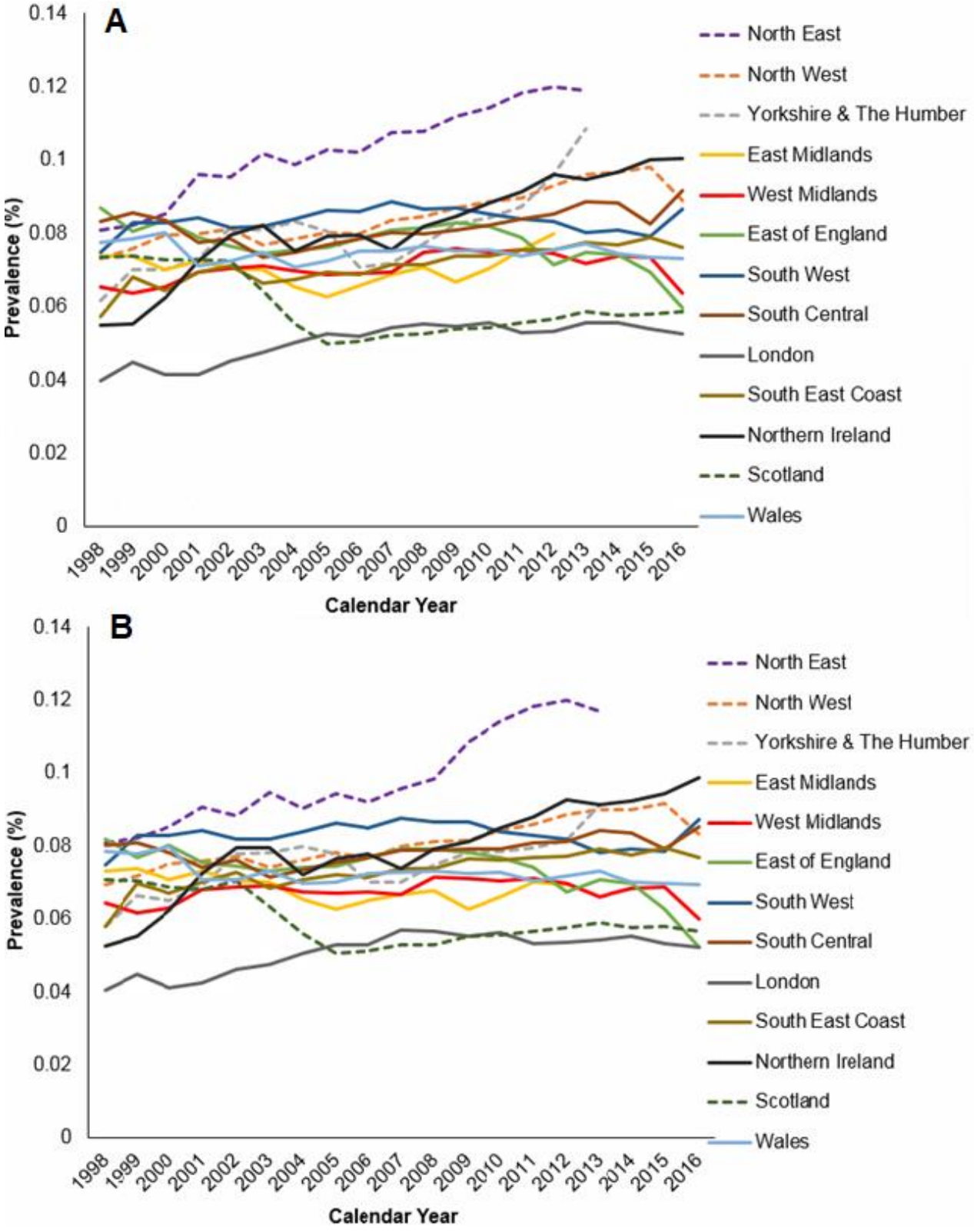


Figure 35. Annual percentage prevalence of AS per geographic region in the sensitivity analyses, 1997-2016: A) AS1 (N = 7,404,732); B) AS2 (N = 7,403,920)



Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later

7.3.2 Rheumatoid Arthritis

A cohort of 71,411 RA patients were identified (excluding 1,164 juvenile RA patients), 44,426 (62.2%) with ≥2 diagnoses (RA1); 45,438 (63.6%) with diagnosis and prescribed DMARD (RA2) (Figure 36). Of the RA patients, 54,685 (76.6%) were in RA1 or RA2 and 35,179 (49.3%) were in RA1 and RA2. The median duration of follow-up was 5.1 (±7.6) years; 4.4 (±6.1) years for the incident cohort (patients diagnosed during 1998-2017). In the full RA cohort (incident and prevalent cases), the median age at diagnosis was 57 (±23) and 70.0% (n = 49,974) were female **(**Table 29**)**. In the incident cohort (n = 31,838), the median age at diagnosis was 61 [IQR: 22] and 67.58% [21,464] were female. The sensitivity analyses showed consistent results.

Figure 36. Study flow diagram of cohort selection

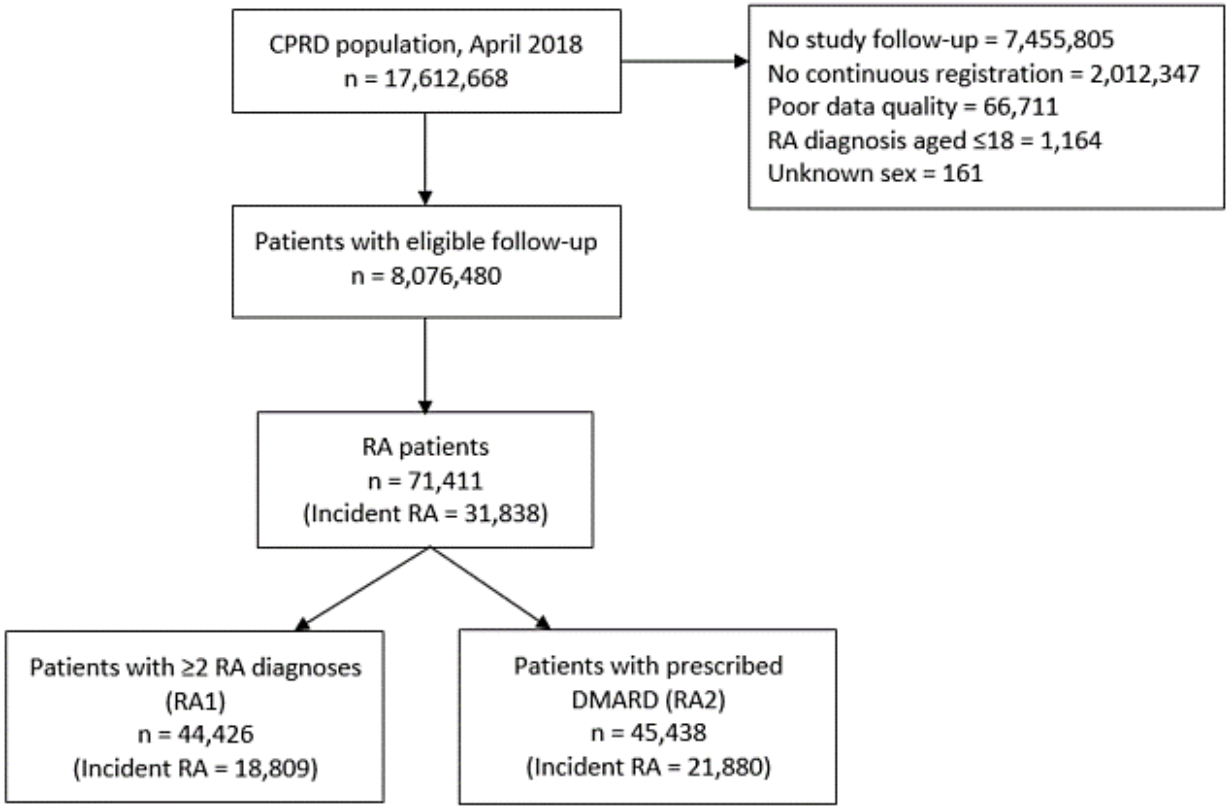


Table 29. Number of RA patients, overall and diagnosed during follow-up, and median age at diagnosis, by sex

	All RA patients		Sensitivity analysis RA1		Sensitivity analysis RA2	
	Full cohort	Incident cohort	Full cohort	Incident cohort	Full cohort	Incident cohort
Patient count	71,411	31,838	44,426	18,809	45,438	21,880

Women (%)	49,974 (70.0)	21,509 (67.6)	31,442 (70.8)	12,862 (68.4)	31,512 (69.4)	14,716 (67.3)
Men (%)	21,437 (30.0)	10,329 (32.4)	12,984 (29.2)	5,947 (31.6)	13,926 (30.6)	7,164 (32.7)
Median age (Q1, Q3)	57 (45-68)	61 (50-72)	55 (44-66)	60 (50-70)	56 (45-66)	60 (49-70)
Women (%)	57 (45-68)	61 (50-72)	55 (43-65)	59 (49-70)	55 (44-66)	59 (48-69)
Men (%)	58 (46-68)	63 (52-72)	56 (46-66)	62 (52-71)	57 (47-67)	62 (52-71)

Note: Q1 = quartile 1, Q3 = quartile 3

7.3.2.1 Incidence

RA incidence was calculated using an at-risk cohort of 8,022,645 patients (approximately 1-3 million in each year). During the study 31,838 patients were diagnosed with RA (729 in 1998, 1,027 in 2017) and the period incidence was 5.57 (±0.06) cases per 10,000 (Table 30). In sensitivity analyses RA1 (≥2 RA codes) and RA2 (DMARD prescribed), 18,657 and 21,295 patients respectively were diagnosed during 1998-2016 and the period incidence was 3.34 (±0.05) and 3.82 (±0.05). Incidence was higher in women (7.44 [±0.10]) than men (3.65 [±0.07]) and increased with age until 70-79 years (11.40 [±0.28]), with little regional variation and similar patterns in the sensitivity analyses.

Table 30. Incidence of RA and mean APC for the full RA cohort from 1998-2017 (N = 8,022,645) and sensitivity analyses from 1998-2016 (N = 7,922,544)

	Number of events (person-years at risk / 10,000)			Incidence rate per 10,000 person-years (95% CI)		
	Main analysis	**Sensitivity analysis RA1**	**Sensitivity analysis RA2**	**Main analysis**	**Sensitivity analysis RA1**	**Sensitivity analysis RA2**
Overall	31,838 (5,718.3)	18,657 (5,578.0)	21295 (5,577.9)	5.57 (0.06)	3.34 (0.05)	3.82 (0.05)
Year						
1998	729 (145.5)	392 (146.4)	422 (146.4)	5.01 (0.36)	2.68 (0.27)	2.88 (0.27)
1999	863 (166.0)	485 (167.1)	532 (167.0)	5.20 (0.35)	2.90 (0.26)	3.18 (0.27)
2000	1,108 (201.5)	581 (202.8)	697 (202.7)	5.50 (0.32)	2.86 (0.23)	3.44 (0.26)
2001	1,369 (244.9)	738 (246.5)	885 (246.4)	5.59 (0.30)	2.99 (0.22)	3.59 (0.24)
2002	1,502 (271.9)	851 (273.6)	1,003 (273.6)	5.52 (0.28)	3.11 (0.21)	3.67 (0.23)
2003	1,692 (305.9)	938 (307.9)	1,155 (307.8)	5.53 (0.26)	3.05 (0.19)	3.75 (0.22)
2004	1,853 (327.6)	1,024 (329.7)	1,281 (329.6)	5.66 (0.26)	3.11 (0.19)	3.89 (0.21)
2005	1,795 (349.4)	1,031 (351.6)	1,261 (351.5)	5.14 (0.24)	2.93 (0.18)	3.59 (0.20)
2006	1,710 (355.3)	995 (357.5)	1,226 (357.4)	4.81 (0.23)	2.78 (0.17)	3.43 (0.19)
2007	1,664 (357.1)	1,049 (359.2)	1,291 (359.1)	4.66 (0.22)	2.92 (0.18)	3.59 (0.20)
2008	1,595 (358.9)	1,038 (360.9)	1,279 (360.8)	4.44 (0.22)	2.88 (0.17)	3.54 (0.19)
2009	1,734 (359.4)	1,216 (361.4)	1,438 (361.3)	4.82 (0.23)	3.36 (0.19)	3.98 (0.21)
2010	1,605 (353.8)	1,139 (355.6)	1,314 (355.5)	4.54 (0.22)	3.20 (0.19)	3.70 (0.20)
2011	1,642 (344.6)	1,197 (346.3)	1,307 (346.2)	4.77 (0.23)	3.46 (0.20)	3.77 (0.20)
2012	1,663 (339.0)	1,268 (340.6)	1,286 (340.6)	4.91 (0.24)	3.72 (0.20)	3.78 (0.21)
2013	2,729 (321.9)	1,814 (323.3)	1,646 (323.3)	8.48 (0.32)	5.61 (0.26)	5.09 (0.25)

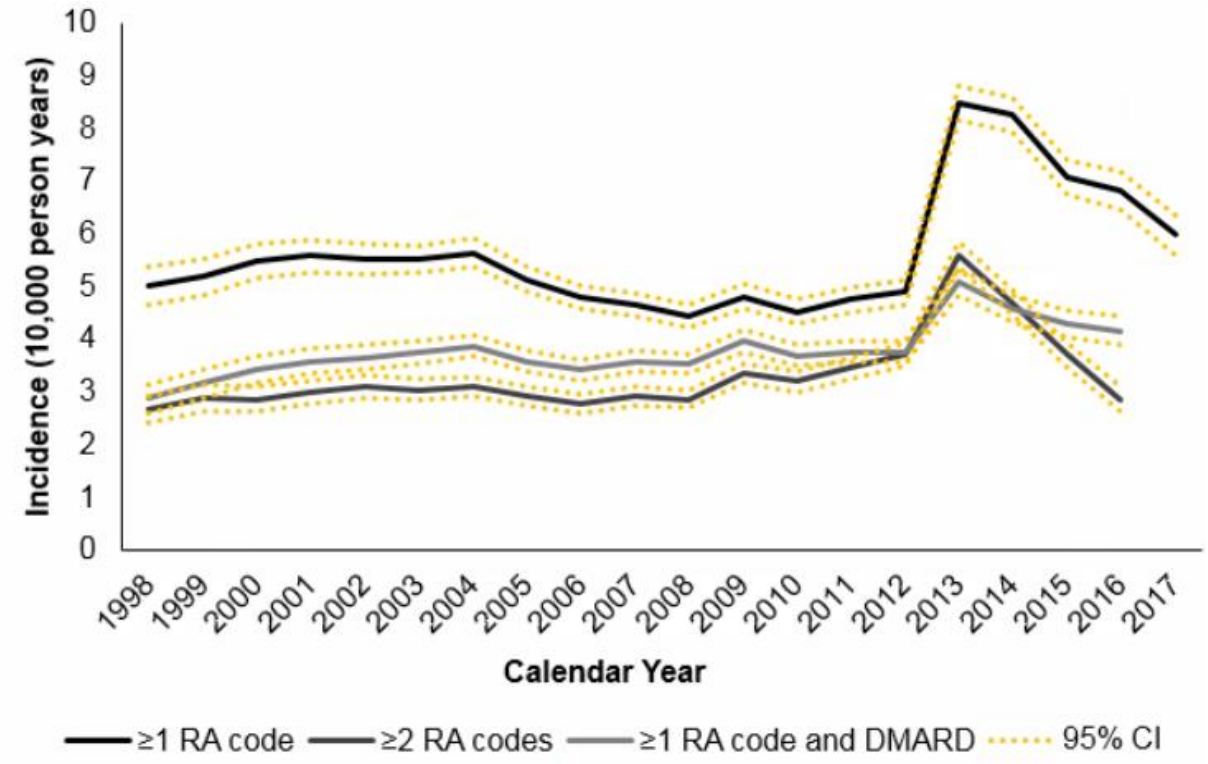
2014	2,411 (291.9)	1,382 (293.2)	1,345 (293.2)	8.26 (0.33)	4.71 (0.25)	4.59 (0.25)
2015	1,779 (251.4)	939 (252.5)	1,085 (252.6)	7.08 (0.33)	3.72 (0.24)	4.29 (0.26)
2016	1,368 (200.8)	580 (210.8)	842 (201.8)	6.81 (0.36)	2.87 (0.23)	4.17 (0.28)
2017	1,027 (171.6)			5.98 (0.37)		
Sex						
Female	21,509 (2,889.8)	12,764 (2,825.7)	14,341 (2,825.8)	7.44 (0.10)	4.52 (0.08)	5.07 (0.08)
Male	10,329 (2,828.5)	5,893 (2,752.2)	6,954 (2,752.1)	3.65 (0.07)	2.14 (0.05)	2.53 (0.06)
Age-group						
18-19	52 (115.7)	27 (112.1)	30 (112.1)	0.45 (0.12)	0.24 (0.09)	0.27 (0.10)
20-29	749 (773.5)	411 (750.6)	506 (750.4)	0.97 (0.07)	0.55 (0.05)	0.67 (0.06)
30-39	2,189 (987.1)	1,327 (960.1)	1,598 (960.0)	2.22 (0.10)	1.38 (0.07)	1.66 (0.08)
40-49	4,453 (1,100.0)	2,865 (1,072)	3,271 (1,071.)	4.05 (0.12)	2.67 (0.10)	3.05 (0.10)
50-59	6,990 (993.5)	4,481 (966.9)	5,038 (966.8)	7.04 (0.16)	4.63 (0.13)	5.21 (0.14)
60-69	7,589 (805.0)	4,694 (788.5)	5,339 (788.5)	9.43 (0.21)	5.95 (0.17)	6.77 (0.18)
70-79	6,516 (571.4)	3,605 (561.3)	4,153 (561.2)	11.40 (0.28)	6.42 (0.21)	7.40 (0.22)
80-89	2,947 (307.2)	1,193 (302.9)	1,307 (302.9)	9.59 (0.35)	3.94 (0.22)	4.31 (0.23)
90-99	350 (64.0)	54 (62.7)	53 (62.72)			
Geographical area						
North East	546 (90.6)	307 (90.6)	398 (90.88)	6.03 (0.51)	3.37 (0.38)	4.24 (0.42)
North West	3,496 (644.3)	2,170 (644.3)	3,206 (646.6)	5.43 (0.18)	3.38 (0.14)	3.94 (0.15)
Yorkshire & The Humber	1,054 (172.5)	399 (172.5)	620 (173.4)	6.11 (0.38)	2.30 (0.23)	3.76 (0.29)
East Midlands	983 (185.0)	305 (185.0)	542 (185.8)	5.31 (0.34)	1.64 (0.18)	3.47 (0.27)
West Midlands	2,887 (539.4)	1,757 (539.4)	2,775 (541.3)	5.35 (0.20)	3.29 (0.15)	3.65 (0.16)

East of England	2,880 (469.3)	1,405 (469.3)	2,242 (471.2)	6.14 (0.23)	3.00 (0.16)	3.73 (0.18)
South West	2,948 (481.8)	1,587 (481.8)	2,592 (483.9)	6.12 (0.23)	3.30 (0.16)	3.79 (0.17)
South Central	3,041 (593.3)	1,813 (593.3)	2,756 (595.1)	5.13 (0.19)	3.08 (0.14)	3.57 (0.15)
London	2,696 (538.4)	1,580 (538.4)	2,643 (540.2)	5.01 (0.20)	2.99 (0.15)	3.11 (0.15)
South East Coast	3,156 (549.3)	2,134 (549.3)	3,271 (551.2)	5.75 (0.21)	3.97 (0.17)	4.03 (0.17)
Northern Ireland	1,155 (195.4)	836 (195.4)	1,146 (196.0)	5.91 (0.35)	4.40 (0.30)	4.39 (0.30)
Scotland	3,074 (594.3)	1,887 (594.3)	3,088 (596.8)	5.17 (0.19)	3.33 (0.15)	4.17 (0.17)
Wales	3,843 (651.7)	2,583 (651.7)	3,934 (654.3)	5.90 (0.19)	4.16 (0.16)	4.15 (0.16)

Note: RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug

The annual incidence was 5.01 (±0.36) per 10,000 person-years in 1998 and 5.98 (±0.37) in 2017. It rose slightly to 5.59 (±0.30) in 2001 before declining to 4.77 (±0.23) in 2011 (mean APC -1.78), before sharply increasing to 8.48 (±0.32) in 2013 and then declining towards pre-2012 levels. The temporal variation was reduced in sensitivity analyses RA1 and RA2; 2.68 (±0.27) and 2.88 (±0.27) in 1998 and 2.87 (±0.14) and 4.17 (±0.28) in 2016 respectively (Figure 37). The mean APC was -0.36 pre-2012 but positive overall (+2.09). In sensitivity analyses RA1 and RA2 the mean APC pre-2012 was +2.17 and +2.27 respectively.

Figure 37. Annual incidence rate of RA, using three definitions of RA (N = 8,022,645): ≥1 RA diagnostic code (1998-2017); ≥2 RA diagnostic codes at least 6 months apart (1998-2016); ≥1 RA diagnostic code plus a subsequent DMARD prescription (1998-2016)



The incidence among women was approximately double that of among men (6.92 (±0.60) and 3.01 (±0.40) in 1998; 7.86 (±0.58) and 4.33 (±0.44) in 2017 respectively), with comparable results in sensitivity analyses (Figure 38, Figure 39). Annual incidence did not seem to vary based on age and remained highest in patients aged 70-79 (10.53 (±1.62) in 1998 and 11.09 (±1.52) in 2017), with comparable results in sensitivity analyses (Figure 40). There was little regional variation except for two peaks: in 2015 (20.03 (±2.44)) and 2016 (20.03 (±3.21)) in the East of England; in 2016 (8.64 (±1.46)) and 2017 (11.31(±2.08)) in South Central England (Figure 41). Regional variation in the sensitivity analyses reveal variation in coding practice, and to a far lesser degree, GP

DMARD prescribing. In sensitivity analysis RA1, the East Midlands deviates with falling incidence, while the West Midlands has a peak in 2016 (Figure 42). The peaks in the East of England and South Central England are not present in sensitivity analysis RA2. The annual incidence returned toward pre-2013 levels more sharply when regions with <5 practices in a given year and the East of England from 2014 were excluded as outliers (Figure 43).

Figure 38. Annual incidence rate of RA in women and men, 1997-2017 (N = 8,022,645)

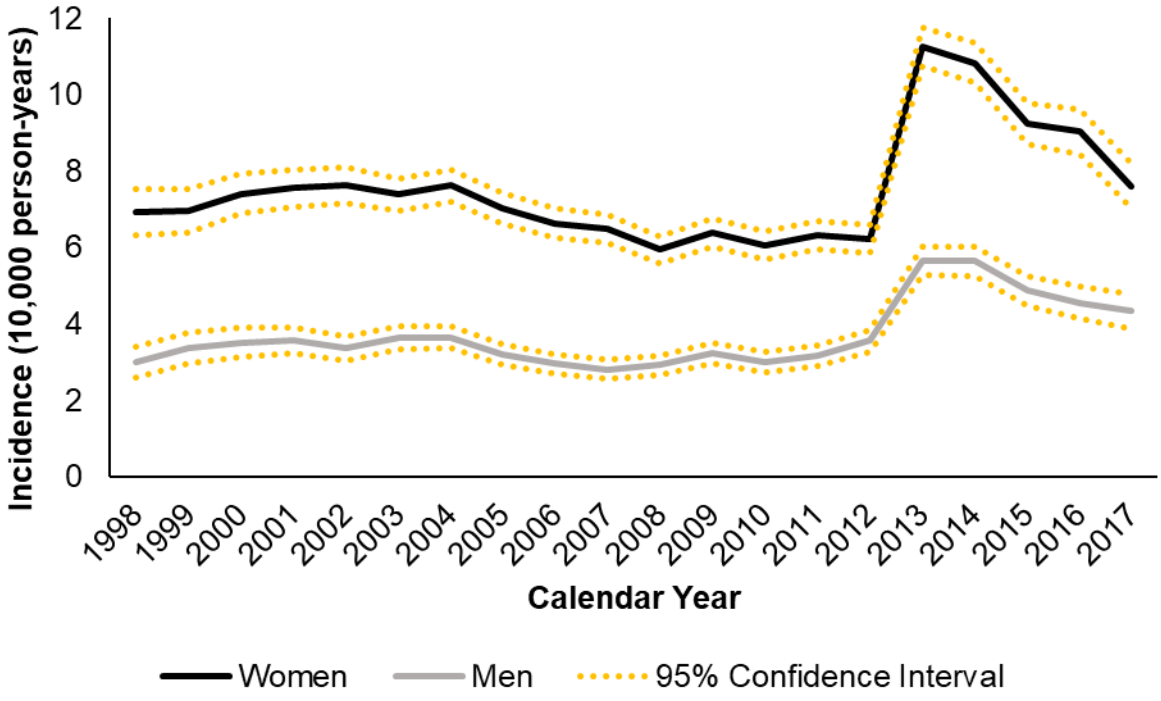
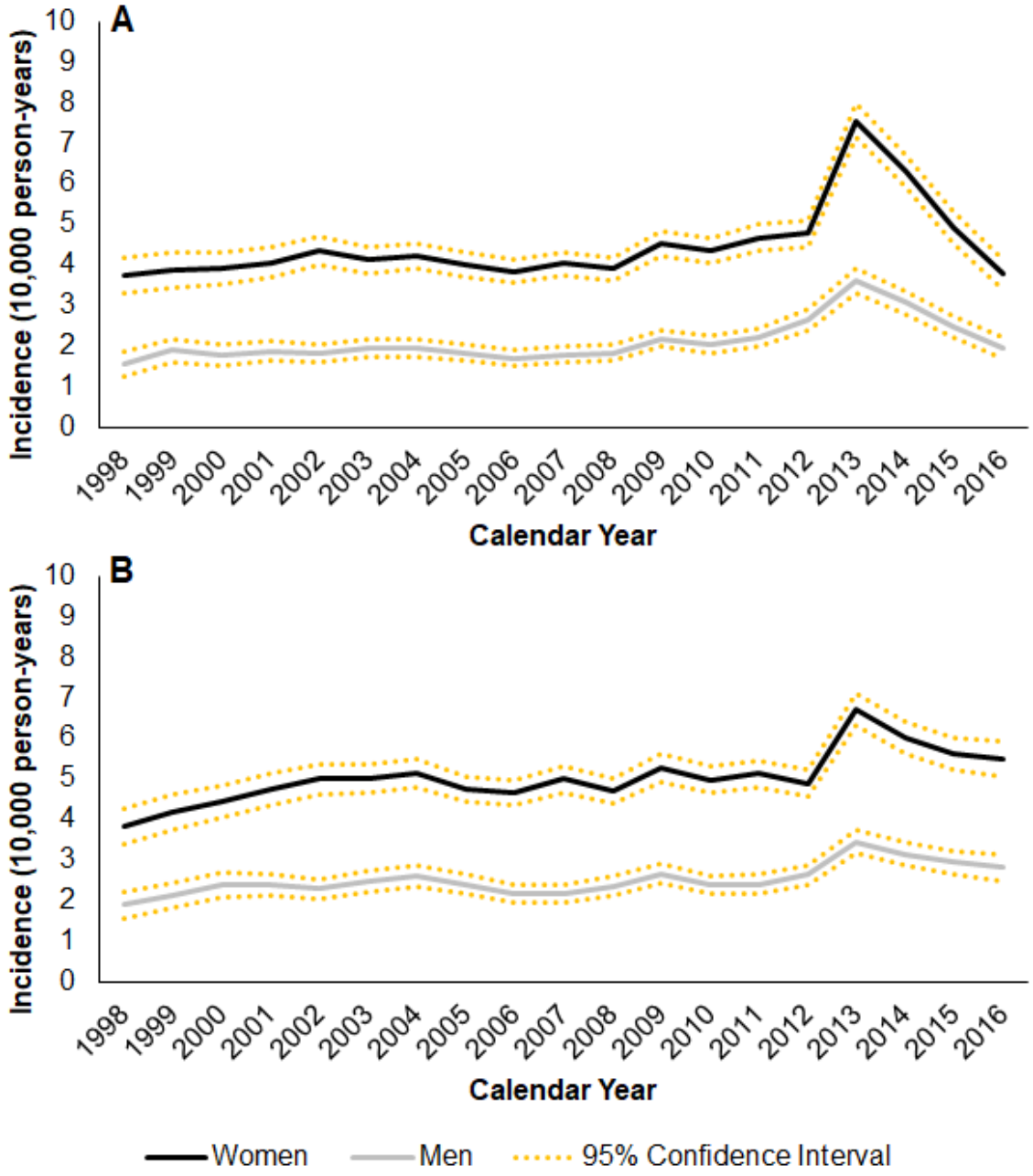


Figure 39. Annual incidence rate of RA in women and men in sensitivity analyses, 1997-2016 (N = 7,922,544): A) RA1; B) RA2



Note: RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug

Figure 40. Annual incidence rate of RA by age-group, 1998-2017 (N = 8,021,209)

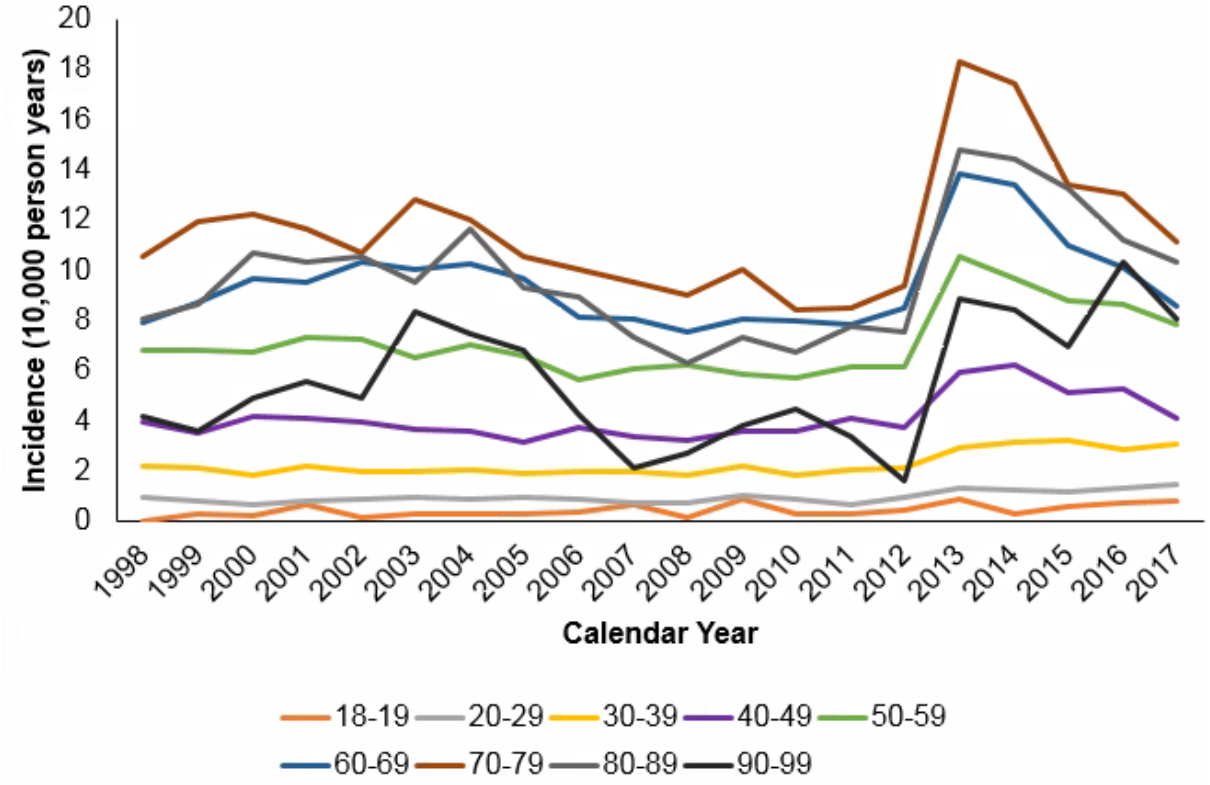


Figure 41. Annual incidence rate of RA by geographic region, 1998-2017 (N = 8,014,524)

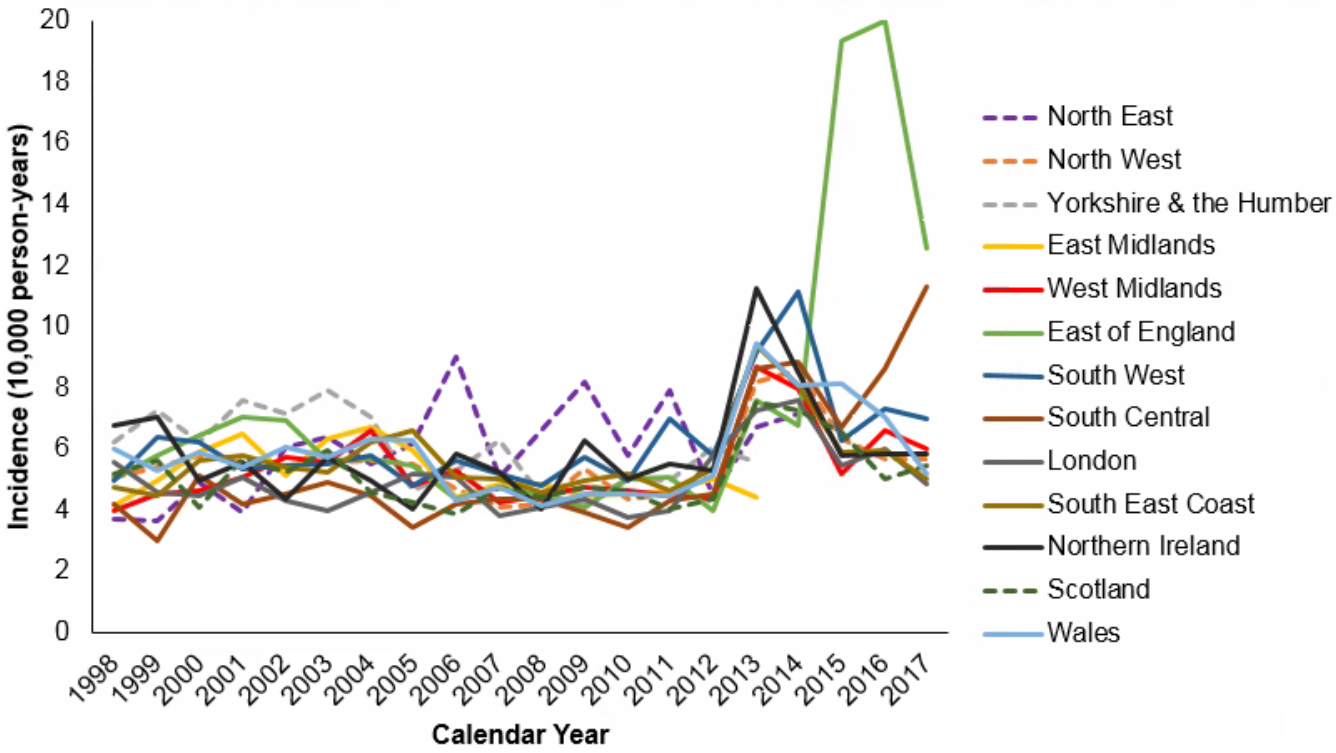
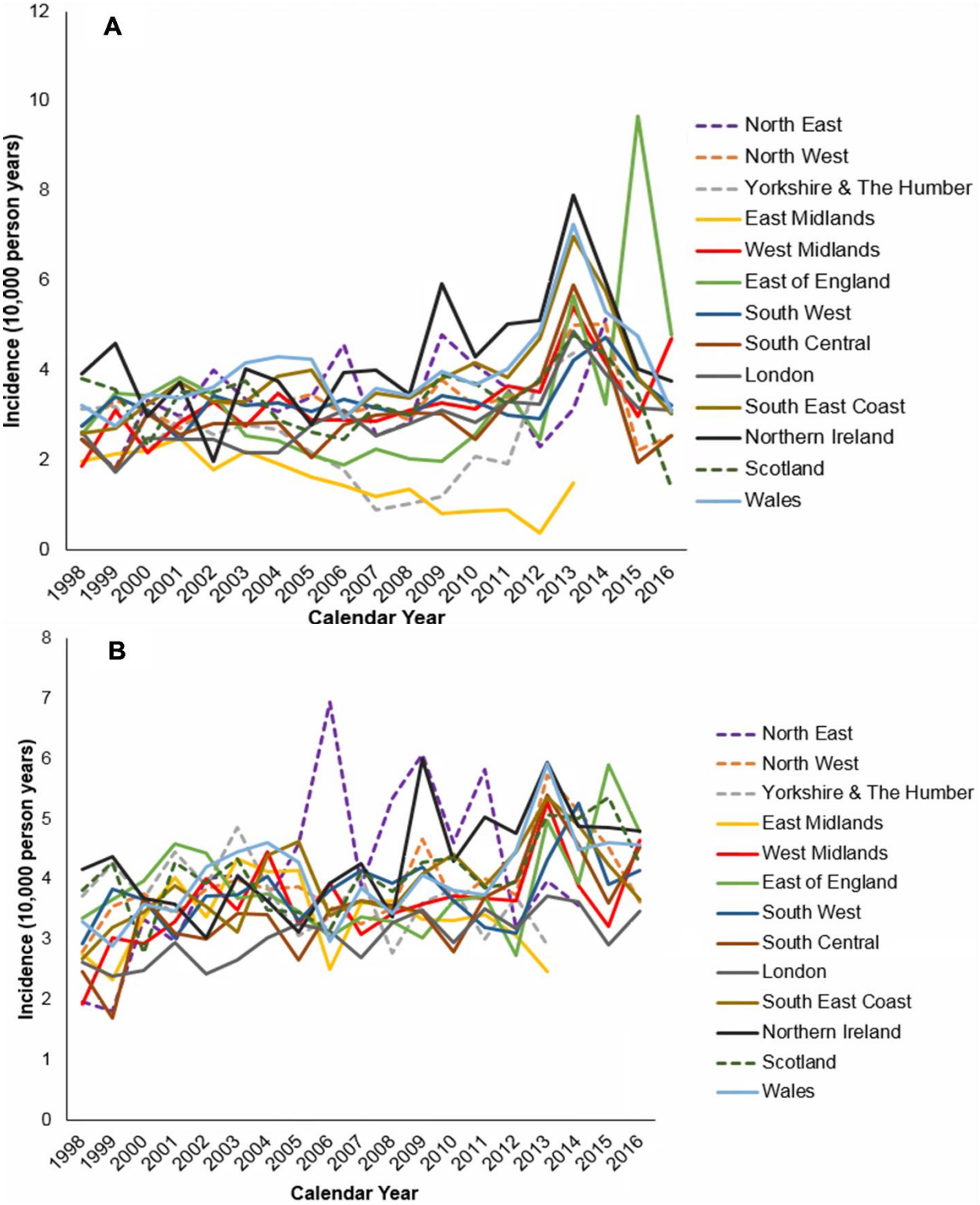
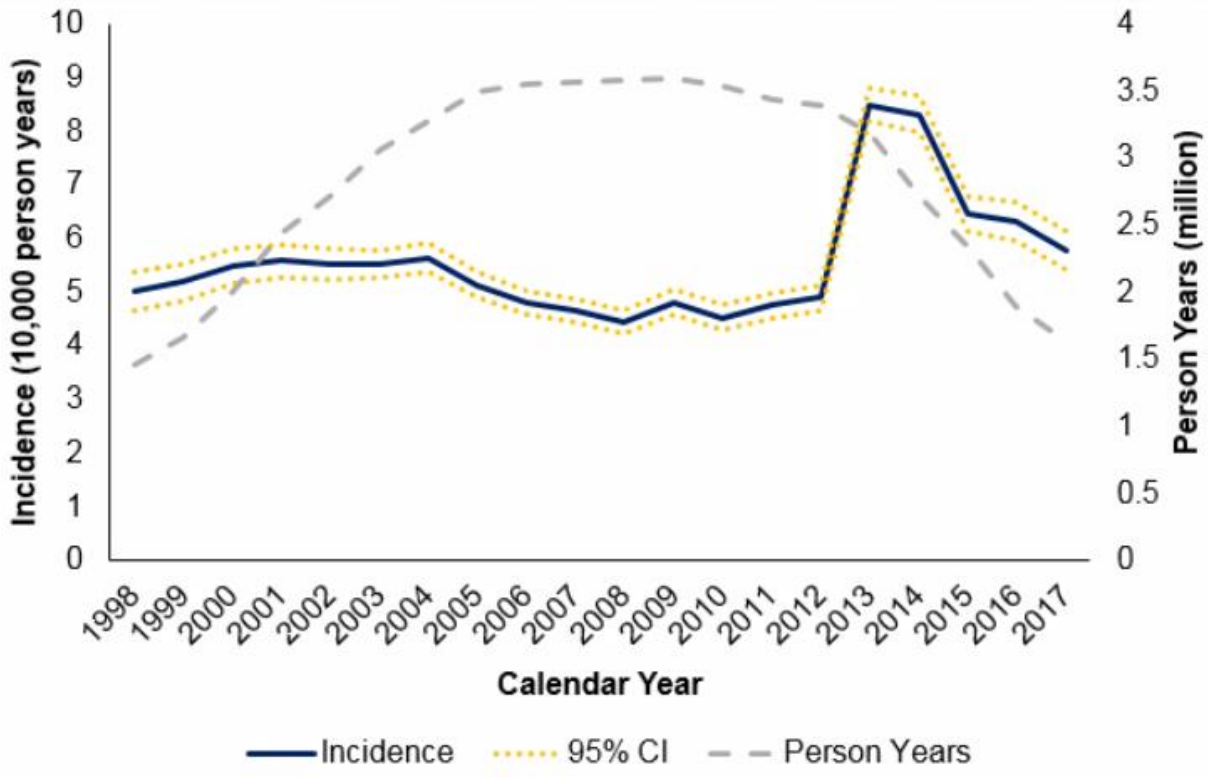


Figure 42. Annual incidence rate of RA by geographic region in the sensitivity analyses, 1998-2016: A) RA1 (N = 7,916,842); B) RA2 (N = 7,915,842)



Note: RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug. Data suppressed where ≤5 cases

Figure 43. Annual incidence rate of RA and person-years of follow-up in the at-risk cohort (1998-2017), excluding regions with <5 GP practices in a given year and the East of England post-2013 (N = 7,981,915)



In the sub-analysis of sensitivity analysis RA1 where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis, the incidence was lower (e.g. 3.40 [±0.30] in 1998; 4.64 [±0.32] in 2016) excepting for a substantial peak in 2013 (23.13 [±0.52]).

7.3.2.2 Prevalence

RA prevalence was calculated using a cohort of 7,532,147 patients (7,412,859 in sensitivity analyses). The crude period prevalence across 1998-2017 was 0.89% (±0.01); 0.58% (±0.01) across 1998-2016 in both sensitivity analyses (Table 31).

Table 31. Percentage prevalence of RA by calendar year and sociodemographic factors, in the main analysis (N = 7,532,147) and sensitivity analyses (N = 7,412,859)

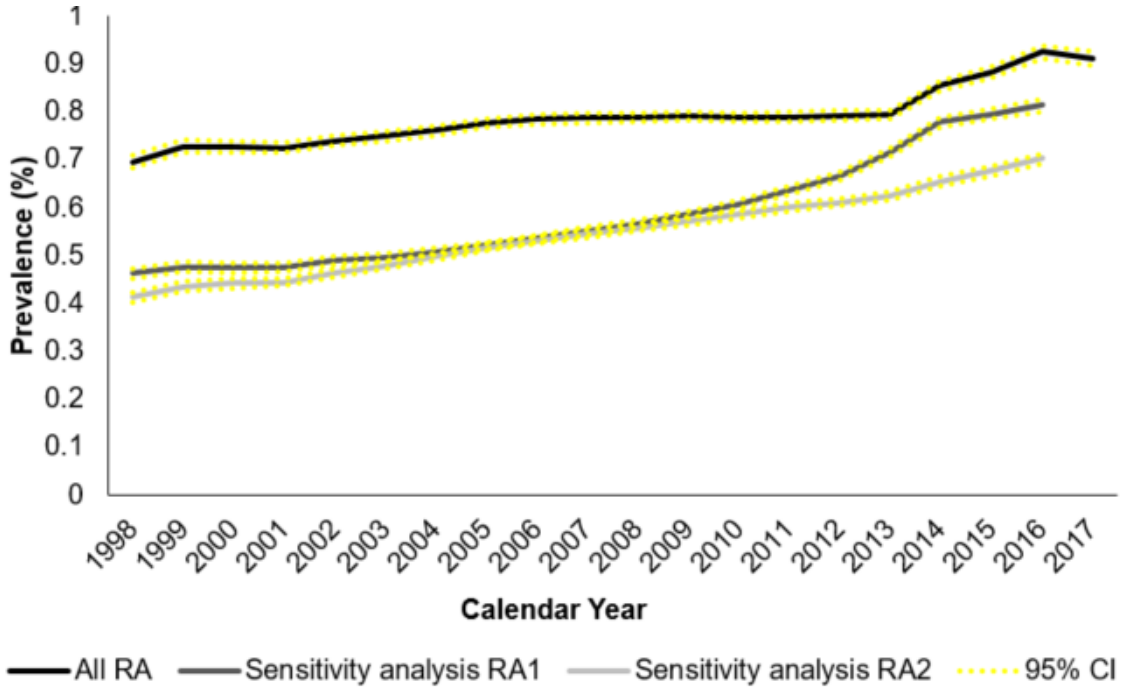
	Percentage prevalence (±95% CI)		
	Full RA cohort	**Sensitivity Analysis RA1**	**Sensitivity Analysis RA2**
Overall	0.89 (0.01)	0.58 (0.01)	0.58 (0.0)

Year			
1998	0.70 (0.01)	0.46 (0.01)	0.41 (0.0)
1999	0.73 (0.01)	0.48 (0.01)	0.44 (0.0)
2000	0.73 (0.01)	0.47 (0.01)	0.44 (0.0)
2001	0.73 (0.01)	0.47 (0.01)	0.44 (0.0)
2002	0.74 (0.01)	0.49 (0.01)	0.46 (0.0)
2003	0.75 (0.01)	0.50 (0.01)	0.48 (0.0)
2004	0.76 (0.01)	0.51 (0.01)	0.50 (0.0)
2005	0.78 (0.01)	0.52 (0.01)	0.52 (0.0)
2006	0.78 (0.01)	0.54 (0.01)	0.53 (0.0)
2007	0.79 (0.01)	0.55 (0.01)	0.55 (0.0)
2008	0.79 (0.01)	0.57 (0.01)	0.56 (0.0)
2009	0.79 (0.01)	0.59 (0.01)	0.57 (0.0)
2010	0.79 (0.01)	0.61 (0.01)	0.59 (0.0)
2011	0.79 (0.01)	0.64 (0.01)	0.60 (0.0)
2012	0.79 (0.01)	0.67 (0.01)	0.61 (0.0)
2013	0.79 (0.01)	0.72 (0.01)	0.62 (0.0)
2014	0.86 (0.01)	0.78 (0.01)	0.65 (0.0)
2015	0.88 (0.01)	0.80 (0.01)	0.68 (0.0)
2016	0.93 (0.01)	0.81 (0.01)	0.70 (0.0)
2017	0.91 (0.01)		
Sex			
Female	1.22 (0.01)	0.80 (0.0	0.79 (0.01)
Male	0.54 (0.01)	0.34 (0.0	0.36 (0.01)
Age-group			
18-29	0.05 (0.003)	0.03 (0.0	0.04 (0.002)
30-39	0.20 (0.01)	0.14 (0.0	0.15 (0.005)
40-49	0.49 (0.01)	0.35 (0.0	0.36 (0.01)
50-59	1.00 (0.01)	0.75 (0.0	0.74 (0.01)
60-69	1.67 (0.02)	1.25 (0.0	1.22 (1.02)
70-79	2.26 (0.03)	1.57 (0.0	1.50 (1.02)
80-89	2.33 (0.04)	1.36 (0.0	1.20 (1.03)
90-99	1.72 (0.06)	0.77 (0.0	0.50 (0.04)
Geographic Region			
North East	0.93 (0.06)	0.61 (0.0	0.60 (0.05)
North West	0.94 (0.02)	0.62 (0.0	0.62 (0.02)

Yorkshire & The Humber	1.04 (0.034)	0.52 (0.0	0.58 (0.03)
East Midlands	0.90 (0.04)	0.40 (0.0	0.55 (0.02)
West Midlands	0.92 (0.02)	0.62 (0.0	0.60 (0.02)
East of England	0.92 (0.02)	0.53 (0.0	0.55 (0.02)
South West	0.94 (0.02)	0.57 (0.0	0.57 (0.02)
South Central	0.78 (0.02)	0.53 (0.0	0.53 (0.02)
London	0.66 (0.02)	0.42 (0.0	0.39 (0.01)
South East Coast	0.88 (0.02)	0.63 (0.0	0.58 (0.02)
Northern Ireland	1.06 (0.05)	0.78 (0.0	0.72 (0.04)
Scotland	0.92 (0.02)	0.60 (0.0	0.67 (0.02)
Wales	1.00 (0.02)	0.72 (0.0	0.67 (0.02)

The crude annual prevalence rose significantly from 0.70% (±0.013) in 1998 to 0.91% (±0.014) in 2017, compared with 0.46 (±0.01) to 0.81 (±0.01) in 2016 and 0.41 (±0.01) to 0.70 (±0.01) in 2016 for sensitivity analyses RA1 and RA2 respectively (Figure 44). The APC rose by mean +1.61 until 2006, before plateauing (mean +0.27) until 2013. The greatest change in APC was in 2013-14 (+7.70; +8.85 and +4.83 in sensitivity analyses RA1 and RA2), with a slowing increase thereafter. Sensitivity analyses RA1 and RA2 showed a greater upward trend in annual prevalence (overall mean APC +3.21 and +3.00).

Figure 44. Annual percentage prevalence of RA in 1997-2017 (N = 7,532,147) and in sensitivity analyses in 1998-2016 (N = 7,412,859)



Note: RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug

Although rising, the prevalence was lower in men, with the difference between women and men remaining stable over time; 0.97% and 0.41% in 1998 and 1.26% and 0.56% in 2017 respectively (Figure 45). By age-group, the largest temporal change in prevalence was among patients aged 80-89: 1.50% (±0.09) in 1998 and 2.30% (±0.10) in 2017. The APC was small in patients aged <70 while prevalence increased with age, reaching 2.21% (±0.05) in 2017 for patients aged ≥70, although prevalence declined from 2008 for the 90-99 age-group (Figure 46). Sensitivity analyses reported a greater increase in prevalence for all patients aged ≥70 (Figure 47). Regional prevalence varied from 0.58% in London and the North East to 0.80% in Wales in 1998 and from 0.72% in East of England to 1.2% in South West England in 2017 (Figure 48). In sensitivity analyses the peak in South West England was absent and the prevalence remained lowest in London and South Central England, though prevalence did not rise in the East Midlands in sensitivity analysis RA1 (Figure 49).

Figure 45. Annual percentage prevalence of RA in women and men, 1997-2017 (N = 7,532,147) and in sensitivity analyses, 1998-2016 (N = 7,412,859)

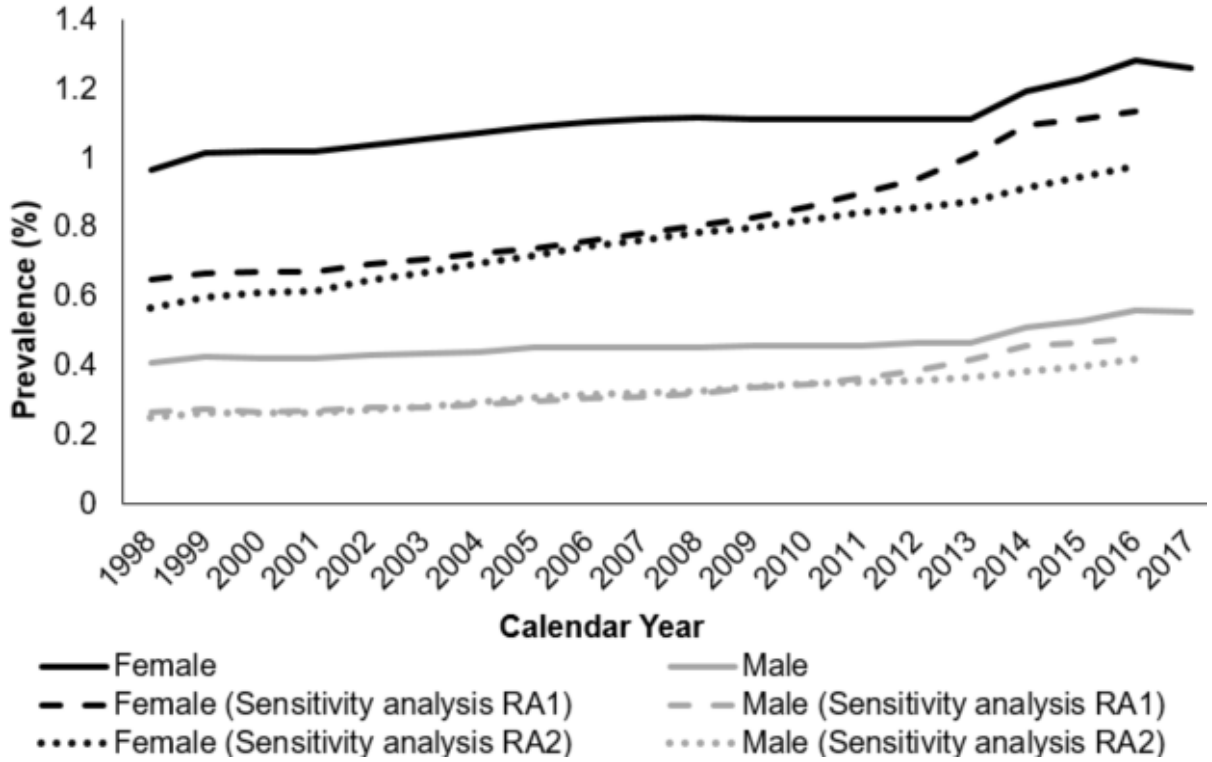


Figure 46. Annual percentage prevalence of RA per age-group among patients aged 18-99, 1997-2017 (N = 7,531,867)

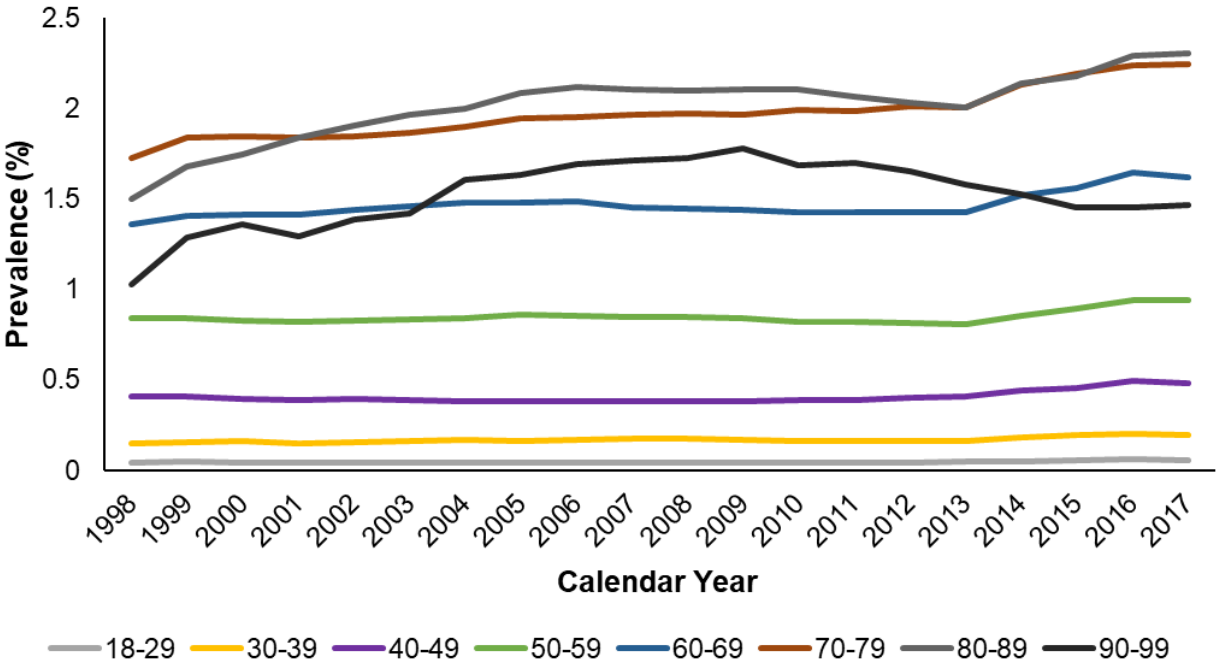
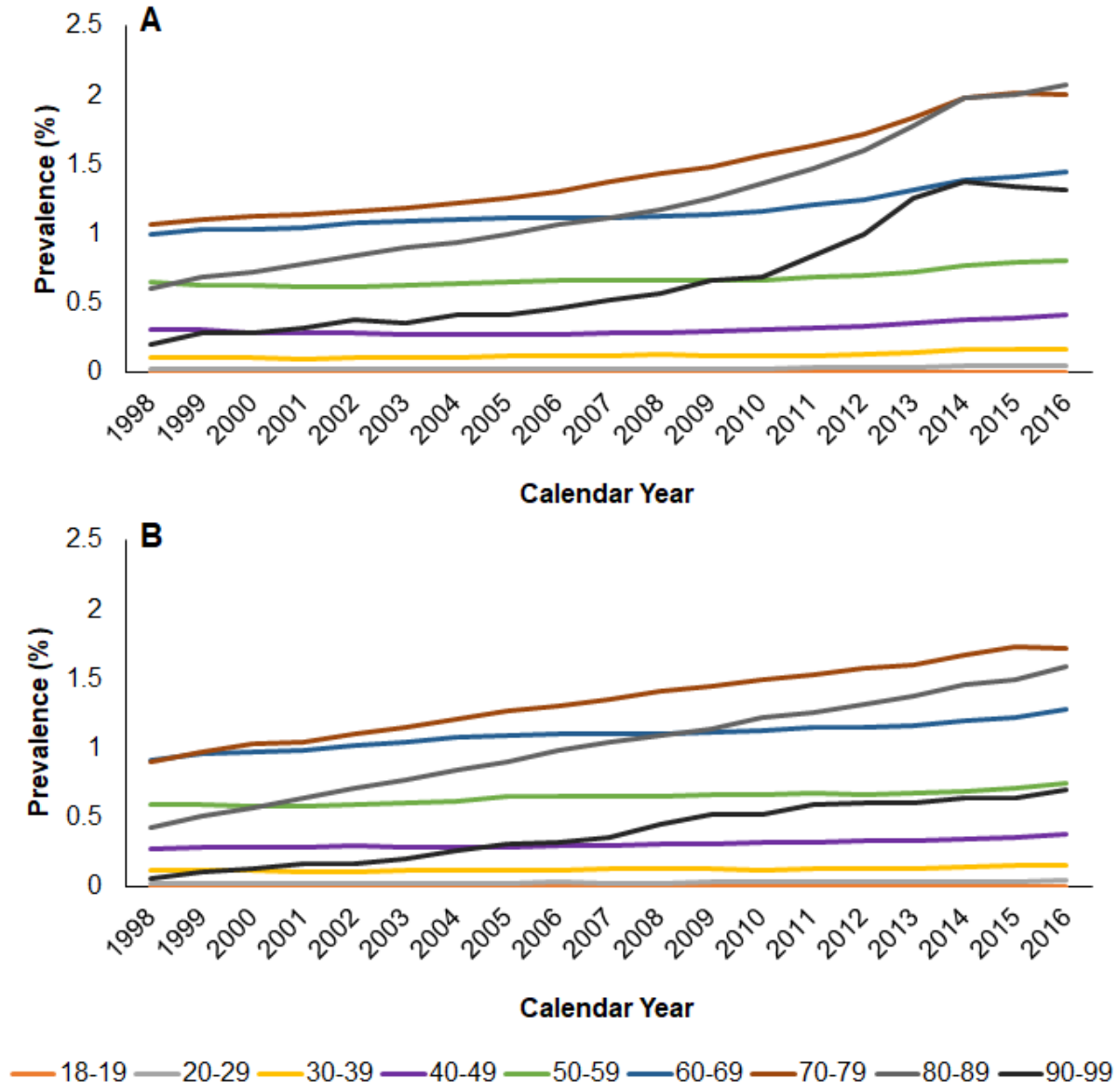


Figure 47. Annual percentage prevalence of RA per age-group among patients aged 18-99 in sensitivity analyses, 1997-2016 (N = 7,412,859): A) RA1; B) RA2



Note: RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug

Figure 48. Annual percentage prevalence of RA per geographic region, 1997-2017 (N = 7,521,506)

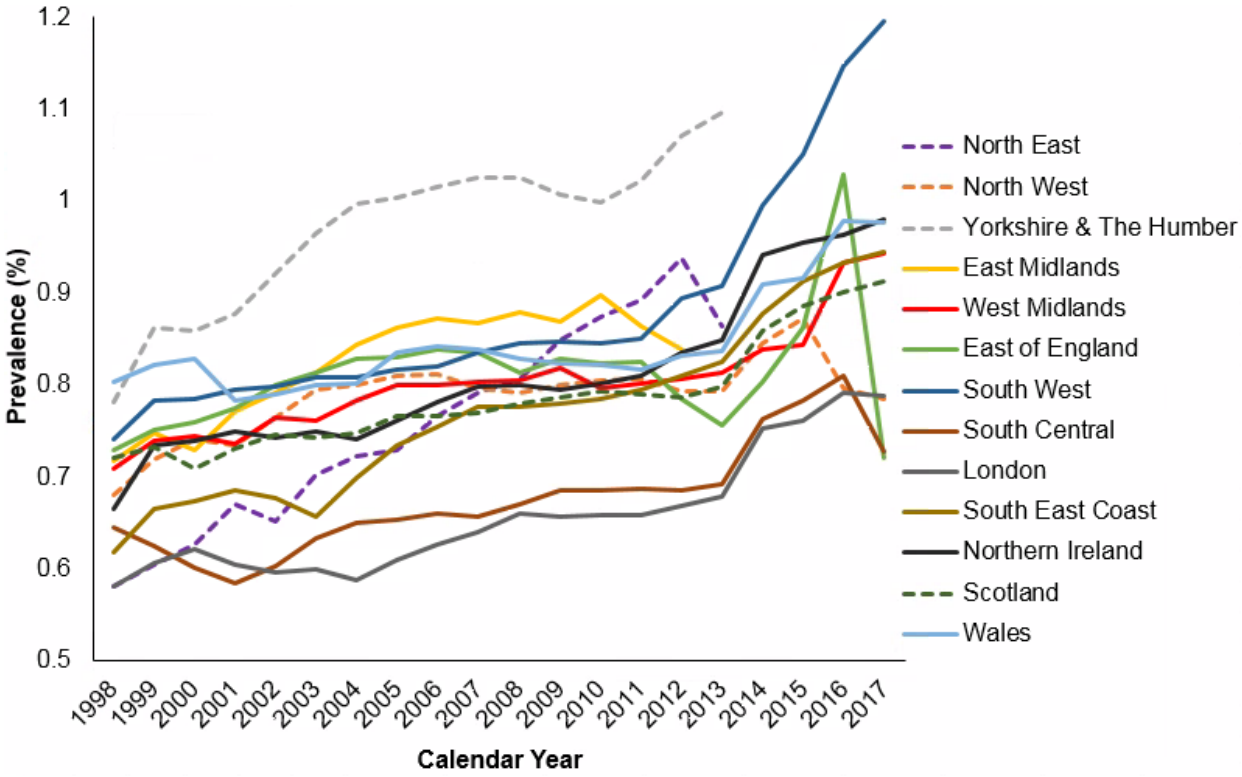
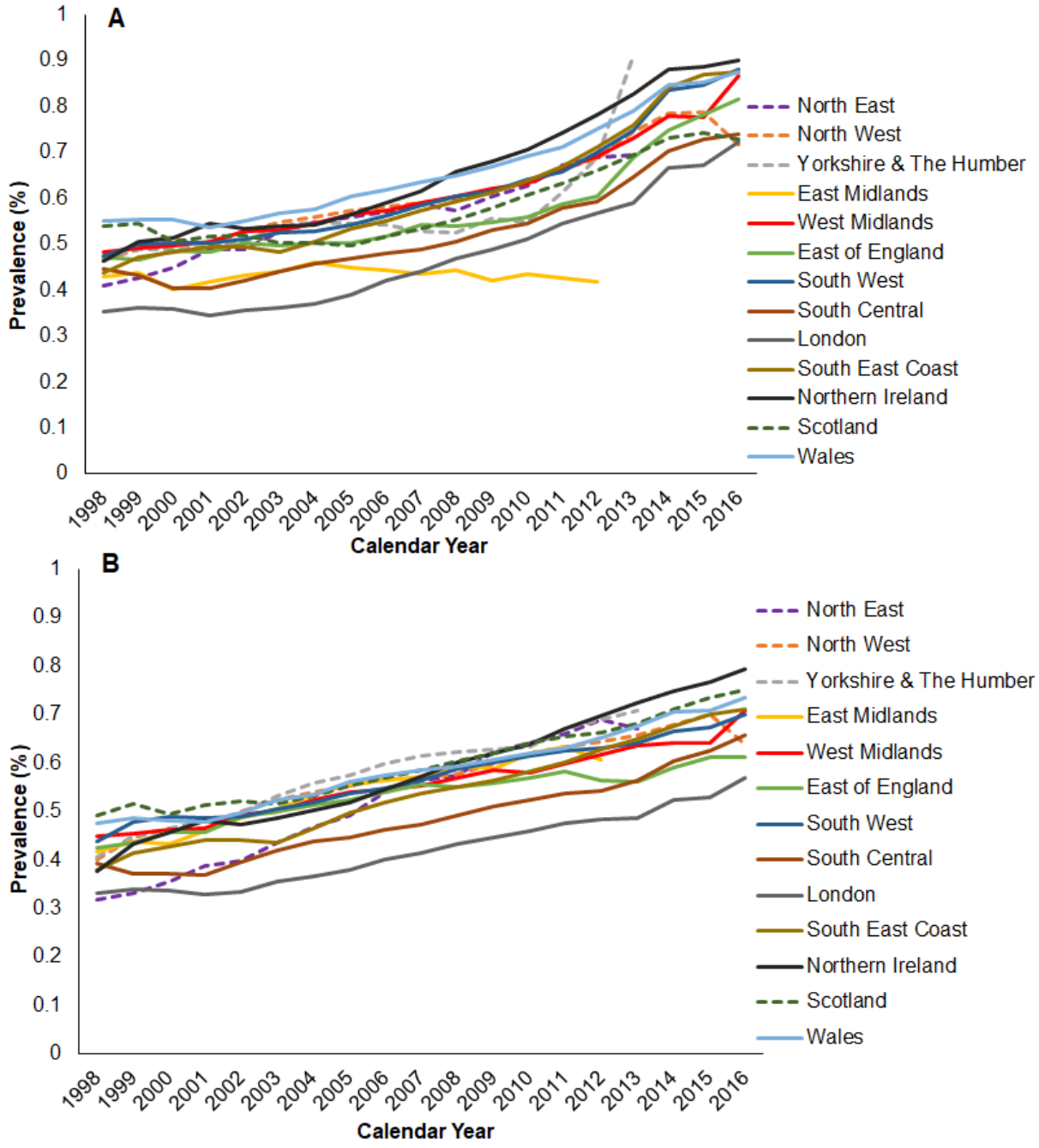
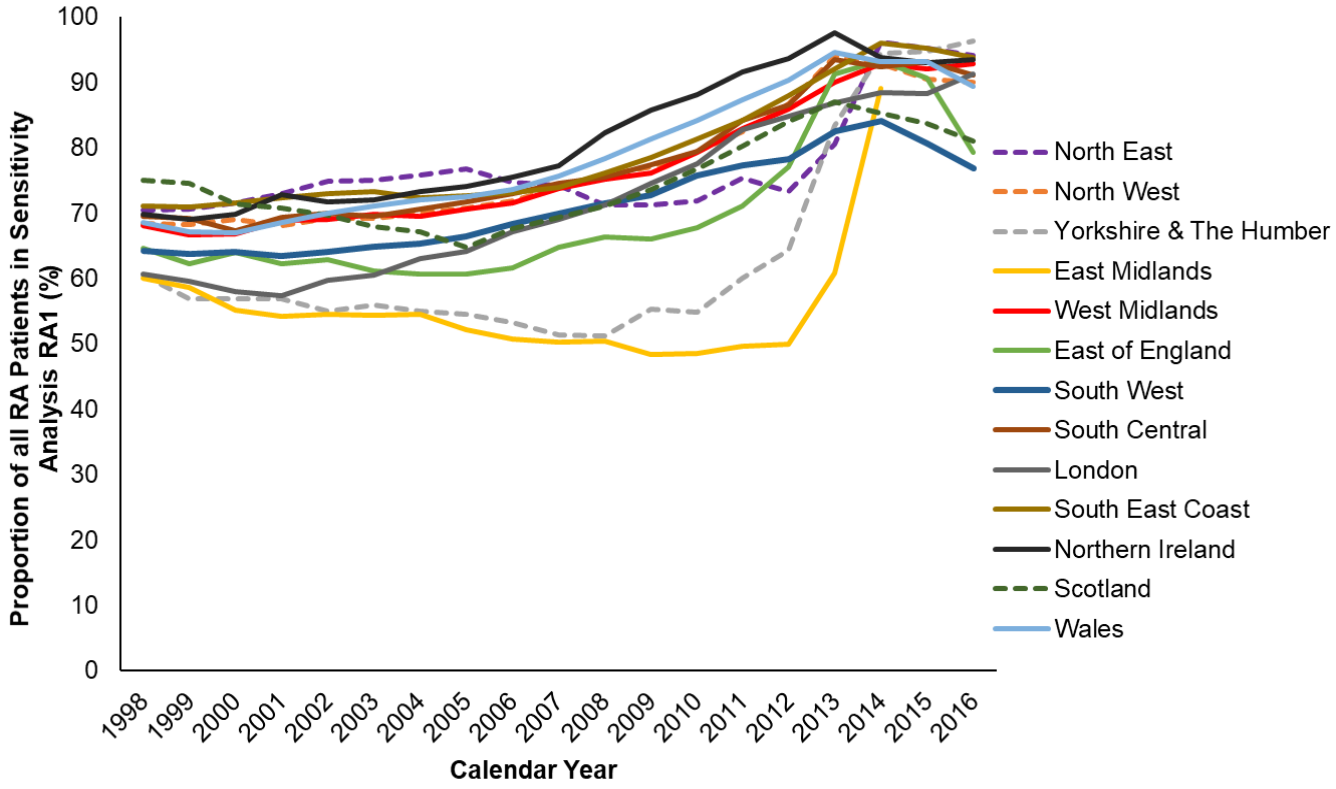


Figure 49. Annual percentage prevalence of RA per geographic region in sensitivity analyses, 1997-2016 (N = 7,403,920): A) RA1; B) RA2



RA coding in RA patients increased over time as the proportion of all RA patients present in sensitivity analysis RA1 (i.e. the proportion of all RA patients having a subsequent RA code ≥6 months later) rose from 66.5% in 1998 to 87.9% in 2016, though the East Midlands and Yorkshire and the Humber initially lagged behind (Figure 50).

Figure 50. The annual percentage of RA patients that have a subsequent RA code ≥6 months after their first RA code (sub-analysis), per region, 1998-2016 (N = 7,403,920)



In the sub-analysis of sensitivity analysis RA1 where the RA code ≥6 months after the first was used to assign the date of RA diagnosis, prevalence was low until it peaked sharply in 2013 (Figure 51, Figure 52, Figure 53). The proportion of all RA patients with a DMARD prescription after their first RA code also increased over time until 2013 (59.4% in 1998; 78.7% in 2013; 75.9% in 2016), with little regional variation (Figure 54).

Figure 51. Annual percentage prevalence of RA in sensitivity analysis RA1 where the RA code ≥6 months after the first was used to assign the date of RA diagnosis (sub-analysis), 1998-2016 (N = 7,412,859)

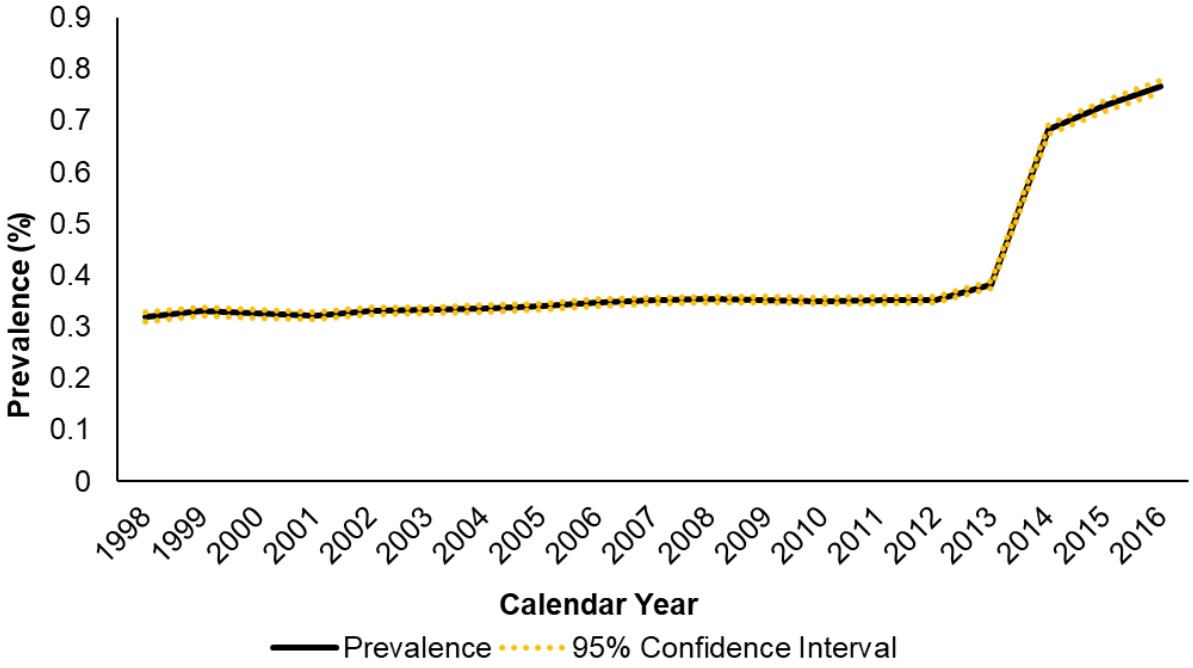


Figure 52. Annual percentage prevalence of RA in women and men, in sensitivity analysis RA1 where the RA code ≥6 months after the first was used to assign the date of RA diagnosis (sub-analysis), 1998-2016 (N = 7,412,859)

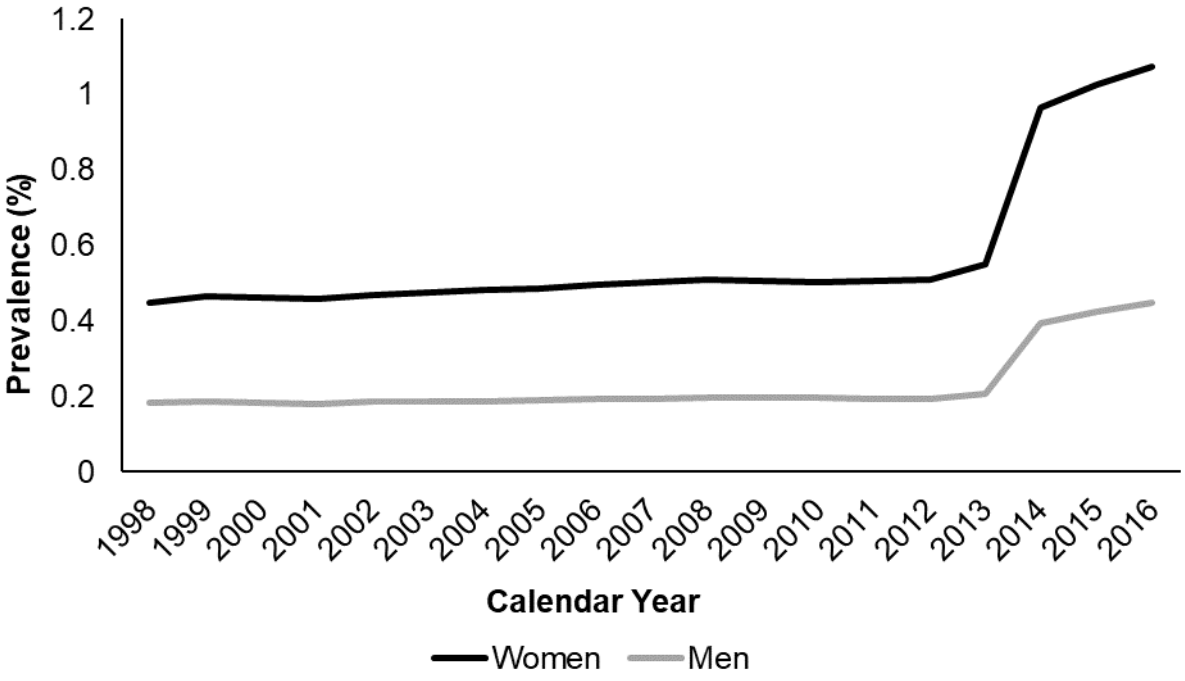


Figure 53. Annual percentage prevalence of RA in sensitivity analysis RA1 where the RA code ≥6 months after the first was used to assign the date of RA diagnosis (sub-analysis), 1997-2016: A) per age-group, (N = 7,412,859); B) per geographic region (N = 7,521,506)

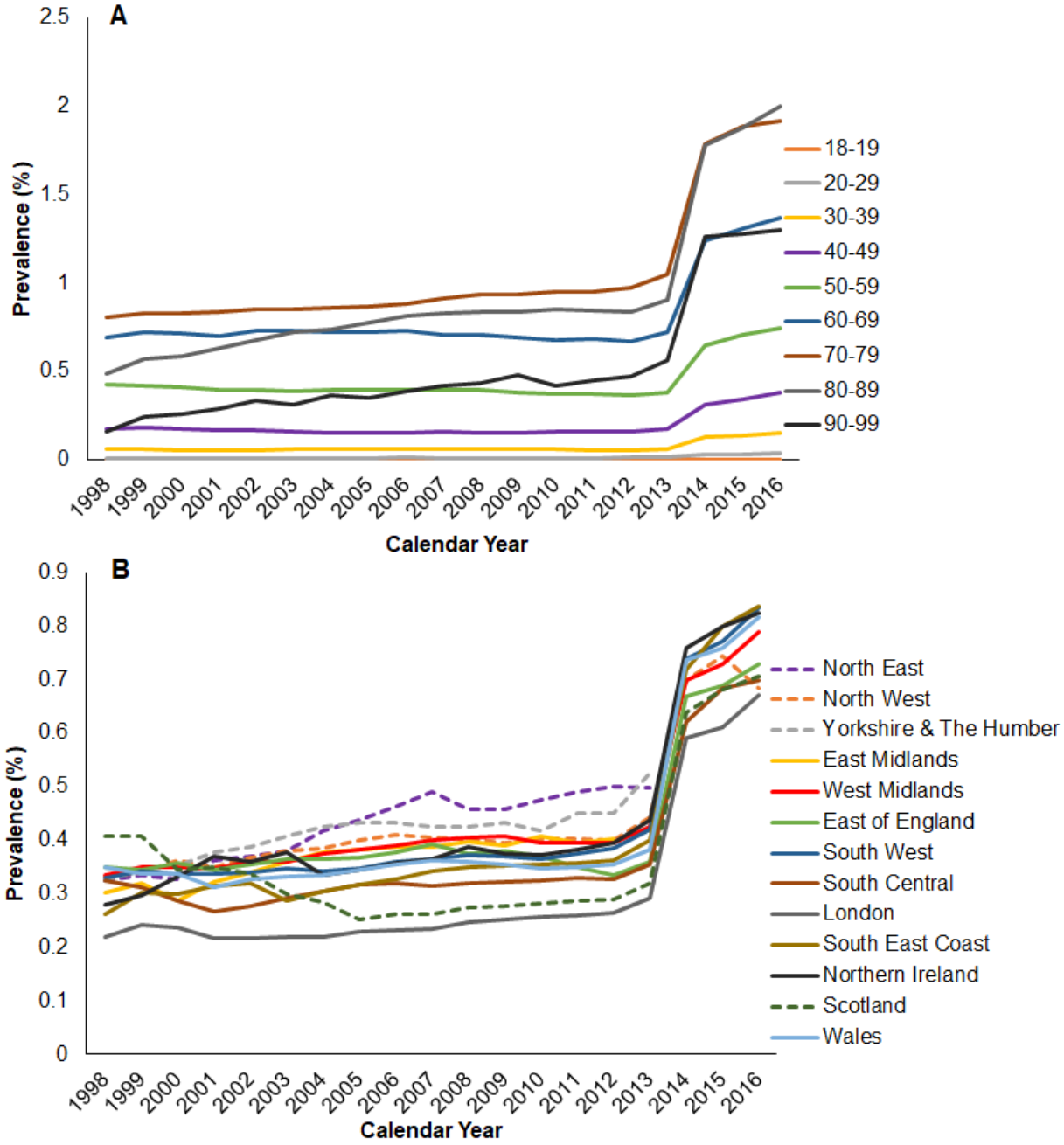
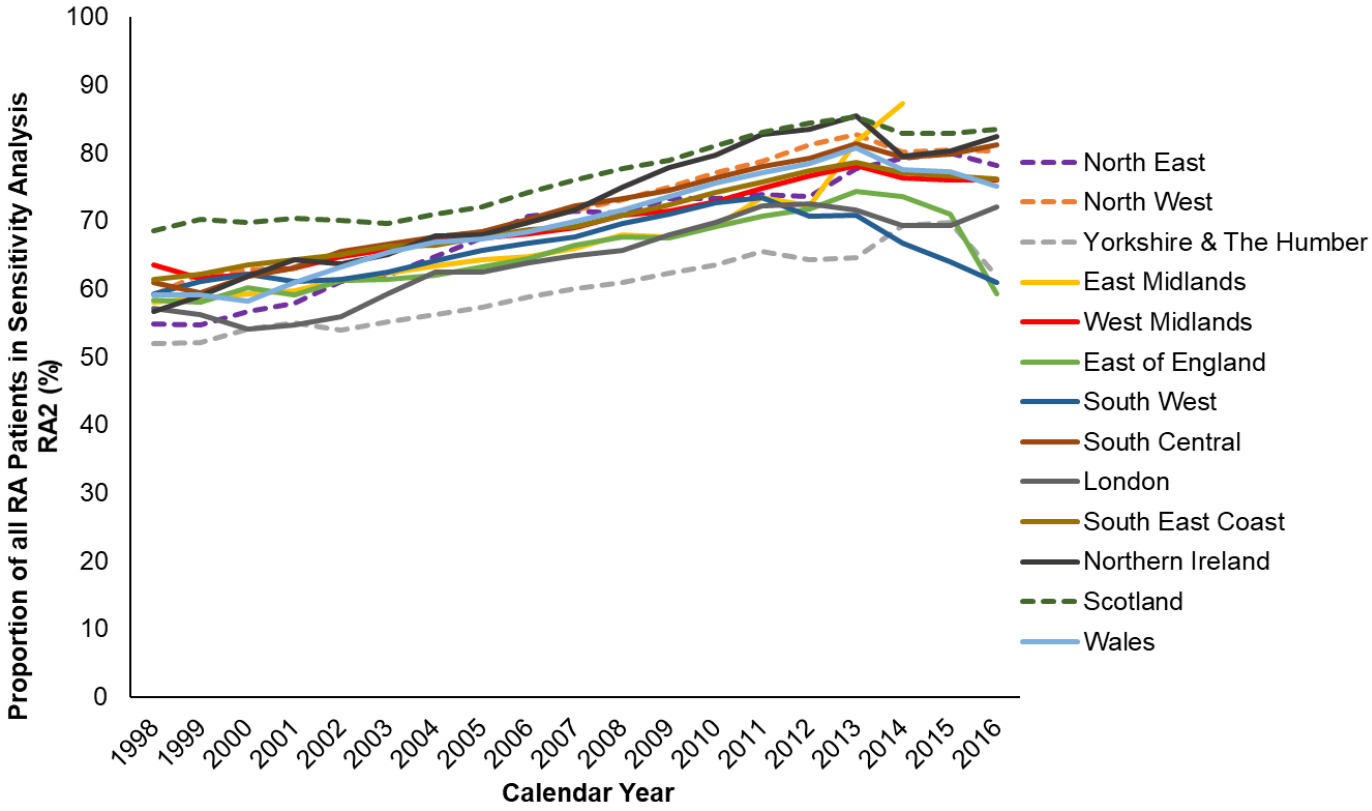


Figure 54. The annual percentage of all RA patients that have a DMARD prescription after their first RA code, per region, 1998-2016 (N = 7,403,920)



7.4 Discussion

The study addressed the objective of describing the epidemiology of disease using EHR data, specifically evaluating regional and demographic variation in the incidence and prevalence of AS and RA over two decades (1998-2017). In this first population-based study of AS incidence and prevalence in the UK, incidence appeared to decline pre-2007 before stabilising. The prevalence of AS rose among patients aged ≥60 and among women. In contrast, the pattern of incidence in RA was less clear although the prevalence of RA rose significantly, especially among the elderly.

7.4.1 Ankylosing Spondylitis

The patient characteristics and the observed pattern of incidence and prevalence of AS differed between the main and sensitivity analyses. Fewer women with an AS code (i.e. women in the main analysis) were included in sensitivity analyses than men (e.g. 34.1% [n = 1,095] and 41.5% [n = 3,787] respectively in sensitivity analysis AS1). This suggests that suspected cases of AS were less likely to be confirmed in women. The age at diagnosis was younger in sensitivity analyses, especially in women, suggesting

that suspected cases of AS were less likely to be confirmed in older patients. The rise in incidence from 2007 was not seen in sensitivity analyses, nor was a stable annual prevalence in patients aged <60; this is discussed further below. A more specific definition of AS, requiring additional coding at least 7 days later, is therefore important in EHR-based studies, particularly when assessing AS in women. Findings from the two sensitivity analyses, performed using a different time window for confirmation of AS diagnosis (i.e. >7 and ≥180 days), yielded comparable results, suggesting that the use of a 7 day window for diagnosis confirmation is sufficient for AS GP EHR studies.

Increasing use of MRI in screening of suspected AS and confidence in distinguishing undifferentiated axial spondyloarthritis from AS may have contributed to the initial decline in incidence. From 2007 an increasing proportion of patients had a single (rather than ≥2 instances of) AS code during follow-up (i.e. that were not included in the sensitivity analyses), suggesting that suspicion of AS may have been raised for an increasing number of patients over time. British Society for Rheumatology guidelines published in 2005 providing recommendations on treating AS with biologics may have raised awareness and prompted screening amongst GPs (240). The increase in single coding was more marked in women and amongst patients aged <40, which suggests that understanding of inflammatory back pain in women and young people might have particularly improved. Similarly, in Canada an increase in single AS coding in women was reported following the introduction of biologics (249). Increasing rates of suspected AS might suggest improved identification of confirmed AS so that measures of AS in the sensitivity analyses post-2007 may be most accurate. In addition, the higher proportion of women in the incident AS cohort than in the full AS cohort, including both incident and prevalent AS patients, (31% and 26% respectively) may also be more representative of the burden of AS in women if consideration of the possibility of an AS diagnosis in women improved over time. Alternatively, recording of AS diagnoses through single but not multiple coding events may have improved over time, which would again suggest that more recent estimates of incidence and prevalence are more accurate.

The period prevalence reported in this thesis was comparable to the weighted mean prevalence of 0.19% estimated using data from other European countries (508). However, findings from the sensitivity analyses performed as part of this thesis (i.e. based on confirmed AS cases) suggests that the 'real' prevalence of AS may be lower, especially in patients aged >60 where the calculated period prevalence showed greatest divergence between the primary and sensitivity analyses. While mortality is higher in AS (particularly cardiovascular mortality, as highlighted in Chapter 2), the finding of significantly rising prevalence in older cohorts (e.g. age 70-79: 0.12% in 1997; 0.30% in 2017) may suggest improved patient survival.

There may have been increasing use of the axial spondyloarthritis diagnosis following the publication of spondyloarthritis classification criteria in 2009 (509). However, this study showed no change in the incidence and prevalence of AS from 2009, suggesting that the evolving nomenclature may have had limited effect in primary care coding of AS.

7.4.2 Rheumatoid Arthritis

The definition of RA did not affect the patient characteristics or the observed pattern of incidence and prevalence although incidence and prevalence was lower in sensitivity analyses. The definition of RA did not affect the patterns in regional or demographic variation in incidence and prevalence, although temporal variation was generally lower in sensitivity analyses. Median age at diagnosis was lower in sensitivity analyses, suggesting that younger patients may be more likely to receive additional coding and prescribed DMARDs. The scale of increase in incidence and prevalence in 2013 was lowest in sensitivity analysis RA2 that required prescribed DMARDs. This suggests that the QOF payment incentives introduced in 2013 for the coding of RA (highlighted in Chapter 2) may have affected RA coding but not DMARD prescribing. Tate et al. (2017) reported a similar increase in coding following changes in payment relating to diabetes (510). This highlights the importance of considering any changes in policy that may affect coding and in turn the sensitivity or specificity of a diagnostic definition across a study period.

Incidence varied over time with no clear trend although with a peak in 2013, especially where RA was defined using RA codes alone. This suggests that coding practice was affected by the change in coding incentives from April 2013. Payment was introduced for using specific RA codes to maintain a registry of RA patients and perform annual review and risk assessments (20). This may have prompted retrospective review of EHRs to identify RA patients that did not have a listed code but instead had another RA diagnostic code, an unspecified arthritis code or an RA free text reference recorded. During this review of extant cases, the payment-related RA codes may either have been retrospectively assigned to the date at which the RA diagnosis was made, or it could be added to the record at the date of the review (i.e. in or after 2013). Occurrences of the former would retrospectively increase estimates of incidence pre-2013 and the latter would cause an apparent spike in incidence. Those RA codes that were associated with payment may have preferentially been selected. This is corroborated by the lower and declining annual incidence reported by Abhishek et al.

between 1990 and 2014 as this study did not use all of the codes used in payment for their definition of RA (265). This included the Read code ‘Nyu10’ which was the only code accepted for one of the three payments that were introduced relating to RA. In corroboration, the sub-analysis of sensitivity analysis RA1 showed that for over half of the patients diagnosed pre-2013, the subsequent RA code ≥6 months after the first was recorded in 2013 (Figure 51). This suggests that RA coding increased substantially in this year. Given the apparent change in coding practice, the study in this chapter, by including in the diagnostic definition of RA all of those codes to which payment was attributed, may have enabled a more accurate estimation of incidence.

The increase in prevalence was greatest in patients aged ≥70, highlighting the importance of RA management in an ageing population. In addition, the median age at diagnosis was higher in the incident cohort than the full RA cohort (incident and prevalent cases), suggesting that the age at diagnosis might be increasing. The age-group with highest prevalence was 80-89 but this was 70-79 in sensitivity analyses, suggesting that follow-up RA coding and DMARD prescribing are less common in older patients despite the payment incentives for performing annual review in RA patients. While RA prevalence plateaued between 2006 and 2013 (mean APC +0.27), in this period it rose in sensitivity analysis RA2 (+2.40) and the proportion of RA patients with a subsequent DMARD rose from 67.9% in 2006 to 78.7% in 2013 (+2.14). This suggests improved DMARD prescribing following the publication of British Society for Rheumatology guidelines in 2006 (479). The sharp increase in the proportion of RA patients in sensitivity analysis RA1 and in the sub-analysis in 2013 show that RA coding increased in 2013, which may follow the above-mentioned change in RA coding incentives. The higher prevalence estimates observed post-2013 may therefore be most accurate.

7.4.3 Strengths and Limitations

The methods appraisal is largely presented in Chapter 10, with chapter-specific strengths and limitations reported here.

Study strengths included the large cohort size with long follow-up. Incidence was reported with the denominator in person-years from an at-risk cohort, which in open observational cohorts (i.e. with changing population and varying lengths of follow-up) may be more appropriate than using a mid-term population and considers the date of diagnosis and duration of disease (511). Similarly, point prevalence calculations accounted for loss from follow-up and was more appropriate than calculating annual

period prevalence proportions, which might under-estimate prevalence for chronic diseases such as AS and RA (512).

Performing two sensitivity analyses for AS and RA helped not only to confirm the study findings but to identify potential change in cording practice in RA from 2013. The proportion of RA patients that were included in sensitivity analysis RA1 (requiring ≥2 RA codes) showed more geographic variation pre-2013, suggesting that payment changes introduced consistency in coding practice. The proportion in sensitivity analysis RA2 showed more geographic variation post-2013, suggesting continued variation in DMARD prescribing. The apparent change in coding practice highlighted the importance of performing sensitivity analyses, and of considering the factors influencing recording of elements used in case definition when interpreting results and comparing studies. For example, two studies assessing the incidence of RA in 1996 in the CPRD derived different results from different definitions: Abhishek et al. defined RA by diagnostic codes and reported this as 3.9 per 10,000 person-years (265); Rodriguez et al. required an additional specialist referral, diagnostic test, specific treatment or confirmation from the GP and reported the incidence as 1.5 per 10,000 person-years (460).

Study limitations include that RA definitions were affected by changing coding practices and rising DMARD prescribing. The declining coverage of CPRD GOLD coincided with the timing of the change in RA incidence and prevalence. However, there was no corresponding change in AS incidence and prevalence, suggesting that this is not a key contributing factor. Further, regional variation in outcomes did not reflect coverage and regional analyses excluded regions with ≤5 contributing GP practices. The relevance of results to patients aged >99 is limited in that age stratifications excluded this cohort, however this is a small cohort and the difference in period prevalence when excluding these patients differed by <0.1%. Spondyloarthritis nomenclature also evolved during the study time-frame, with increasing use of the concept of axial spondyloarthritis among rheumatologists. The incidence of axial spondyloarthritis coding across the study period was not examined, though could have been informative. However, this change in nomenclature may predominantly affect secondary care, as incidence and prevalence of AS did not seem to change in this study following the publication of spondyloarthritis classification criteria in 2009 (509). In EHR-based studies, it is important to consider any such external factors that influence coding practice over time. While the study used UK data, the results may be applicable to countries with similar demographic profiles.

7.4.4 Conclusion

The analyses in this chapter confirm that disease epidemiology can be investigated using EHR data. The incidence of AS in the UK seemed to decline before stabilising from 2007 and rates of suspected AS may have improved, while the pattern of RA incidence was less clear. Prevalence of AS and RA has risen particularly among older patients, highlighting the importance of appropriate disease management in an ageing population such as the UK. RA coding increased in 2013 and led to an peak in the observed incidence (4.91/10,000 person-years in 2012, 8.48 in 2013). This highlights the importance of considering changes in prescribing or coding practice that affect disease definitions, in spatio-temporal comparisons of incidence and prevalence.

Chapter 8 Trends in the Time to Diagnosis in Ankylosing Spondylitis

8.1 Introduction

This chapter addresses the third thesis objective of describing the trends in the timeliness of diagnosis using EHR data. The importance of early diagnosis and treatment for reducing disease activity, increasing the likelihood of remission and extending quality of life in AS was highlighted in Chapter 2 (178, 179, 229). In the same chapter, it was noted that significant diagnostic delay, especially in women, has been reported in studies using hospital data or surveys from patients and rheumatologists (230, 513). Patients complaining of symptoms of IBP typically first present in primary care services. Therefore, GPs play an important role in the timeliness of referral to rheumatology and diagnosis of AS. Chapter 2 also highlighted that modern diagnostic criteria and use of MRI have increased the likelihood of diagnosing AS at early stages of disease development (226, 514). In addition, in Chapter 7 it was suggested that AS screening in UK primary care may have improved since the publication of BSR guidelines on biologic therapy in 2005 (240). However, it is uncertain whether diagnostic delay has subsequently reduced. GP data facilitates deriving information on diagnostic delay, in relation to first evidence of consultation for suggestive AS symptoms over recent calendar years.

Using the large primary care dataset described in Chapter 3, this chapter aims to investigate trends in the time to rheumatology referral and diagnosis in AS, in men and women, over two decades. This study provides information on the diagnostic delay in AS in the UK and the impact of modern diagnostic practices. It also contributes information on the importance of efforts to promote early referral and diagnosis.

8.2 Methods

This retrospective observational study was reported using GP EHR data, following the STROBE guidelines (322). The study protocol approval, data source, study population, study timeframe and criteria for the start and end of study follow-up were described in Chapter 3.

8.2.1 Ankylosing Spondylitis Cohort

Patients diagnosed with AS during follow-up between 1 January 1998 and 31 December 2017 were included in the AS cohort. Given the patient follow-up criteria, the AS cohort had ≥1 year of UTS GP data prior to diagnosis. As described in Chapter 7, the previously validated RCV2 code 'N100.' was used to define AS diagnosis. In sensitivity analyses AS1 and AS2, additional diagnostic or measurement codes were required >7 and ≥180 days later, respectively, in order to confirm the diagnosis of AS (as described in Chapter 7).

8.2.2 Symptoms and Rheumatology Referral

Read codes were used to define symptoms of back pain and determine rheumatology referral (Table 32, Table 33).

Table 32. Read Codes used to determine back pain

Code	Term Description
N142.	Low back pain
16C9.	Chronic low back pain
16CA.	Mechanical low back pain
16C6.	Back pain without radiation NOS
N149.	Back stiffness
N10z.	Spondylitis NOS
16C7.	C/O - upper back ache
16C..	Backache symptom
16C2.	Backache
16C3.	Backache with radiation
16C8.	Exacerbation of backache
16CZ.	Backache symptom NOS
14G4.	H/O: back problem
N145.	Backache; unspecified
N141.	Pain in thoracic spine
16C5.	C/O - low back pain
N1460	Lumbosacral ankylosis
N1461	Sacroiliac ankylosis
N1462	Sacral ankylosis NOS

N1466	Sacroiliac disorder
N148.	Ankylosis/instability of cervical;thoracic or lumbar spine
N1486	Lumbar spine ankylosis
N14z.	Ankylosis of spine NOS

Table 33. Read Codes used to determine rheumatology referral

Code	Term Description
66H9.	Rheumatology management plan given
67Ih.	Advice to GP from rheumatology service
8H2C.	Admit rheumatology emergency
8H3H.	Non-urgent rheumatology admisn
8H4B.	Referred to rheumatologist
8HJC.	Rheumatology self-referral
8HKA.	Rheumatology D.V. requested
8HLA.	Rheumatology D.V. done
8HMA.	Listed for Rheumatology admiss
8HTd.	Referral to rheumatology clinic
8HTP.	Referral to musculoskeletal clinic
8HVQ.	Private referral to rheumatologist
99HB.	Rheumatology disorder annual review
9N0w.	Seen in musculoskeletal clinic
9N1C0	Rheumatology service home visit
9N1O.	Seen in rheumatology clinic
9NIR.	Seen by rheumatology nurse specialist
9NNT.	Under care of rheumatologist
ZL18T	Under care of rheumatologist
ZL22G	Under care of rheumatology nurse specialist
ZL5AR	Referral to rheumatologist
ZL62G	Referral to rheumatology nurse specialist
ZL9AT	Seen by rheumatologist
ZLA2G	Seen by rheumatology nurse specialist
ZLD3T	Discharge by rheumatologist
ZLD7E	Discharge by rheumatology nurse specialist
ZLE6Q	Discharge from rheumatology service

8.2.3 Outcomes

The primary outcome was the annual time to diagnosis among men and women, defined as the number of years between first coded non-specific back pain symptom and the first recorded diagnosis of AS (515). Time from first symptom to rheumatology referral, and from rheumatology referral to diagnosis, were secondary outcomes.

8.2.4 Statistical Analyses

Baseline cohort characteristics were described for the AS cohorts. Annual trends in outcome measures were reported per calendar year between 1 January 1997 and 31 December 2017, stratified by sex. Sensitivity analyses, using alternative definitions of AS, ran until 31 December 2016 to enable >16 months of follow-up for the additional coding to occur. Data were suppressed where there were ≤5 cases (e.g. if ≤5 women or men were diagnosed with AS in a calendar year, the calculation is not reported). Sensitivity analysis AS2 (additional code ≥180 days later) was not performed for measures of the secondary outcomes given the finding in Chapter 7 that a 7 day window for diagnosis confirmation sufficed for AS GP EHR studies.

The earliest recorded back pain symptom, and the first subsequent rheumatology referral, prior to AS diagnosis were identified. The median time between these (symptom to referral, referral to diagnosis, symptom to diagnosis) were calculated overall and for patients diagnosed with AS each year. In 'UTS-related sensitivity analyses' of time from symptom to diagnosis (primary outcome), patients with ≥2 and ≥3 years of quality (UTS) registration prior to AS diagnosis were included.

8.3 Results

Between 1998 and 2017, 3,101 patients were diagnosed with AS during follow-up; 1,071 and 837 having a subsequent diagnostic or measurement code >7 and ≥180 days later (sensitivity analyses AS1 and AS2). The median duration of follow-up was 12.61 (IQR = 7.69-16.06) years. The proportions of women were 30.7% (n = 953), 27.9% (n = 299) and 29.4% (n = 246) respectively. The median ages at diagnosis were 43 (IQR = 33-56), 40 (32-51) and 40 (32-50) years.

8.3.1 Time to Diagnosis

At AS diagnosis, 2,120 patients (68.4%; 673 women, 1,447 men) had a prior-recorded back-pain symptom (757 [70.7%] and 592 [70.7%] in sensitivity analyses AS1 and AS2). The proportions rose over time (60.4%, 72.0% and 75.0% in 1998, 77.6%, 76.2% and 75.6% in 2017/16) (Figure 55, Figure 56). More women than men had a prior-recorded symptom (70.6%, 73.6% and 73.2% in women compared with 67.4%, 69.6% and 69.7% in men) (Figure 57). In sensitivity analyses with ≥2 and ≥3 prior UTS years, the proportions of patients with prior-recorded back-pain symptoms were 70.6% and 72.4%, higher than in the primary analysis, and rose from 60.6% and 62.5% in 1998 to 82.8% and 82.5% in 2017. As in the primary analysis, the proportion with a prior-recorded symptom was higher in women (69.6% and 81.3% in 1998, 92.0% and 91.7% in 2017) than in men (57.7% and 62.5% in 1998, 79.0% and 78.6% in 2017). Sensitivity analyses AS1 and AS2 showed consistent patterns in UTS-related sensitivity analyses.

Figure 55. Annual percentage of patients diagnosed with AS (having ≥1, ≥2 and ≥3 years of prior UTS registration) who had a prior back-pain symptom code, 1998-2017 (N = 3,101; N = 2,734; N = 2,417)

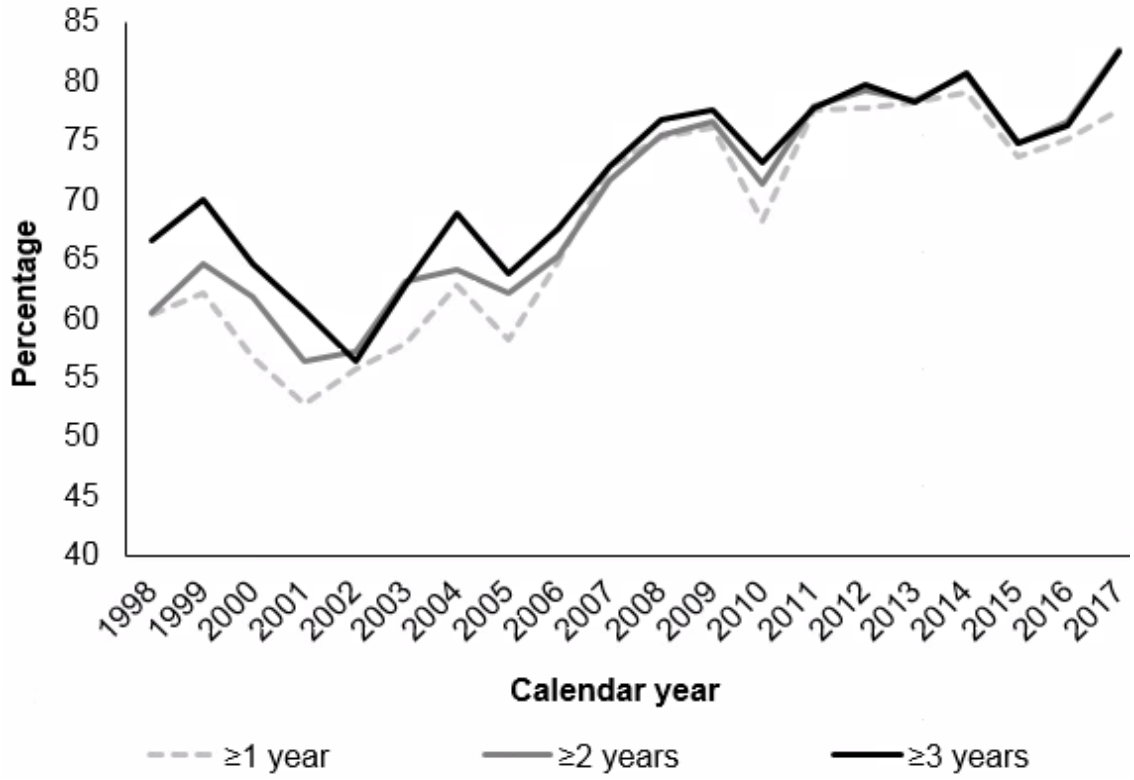
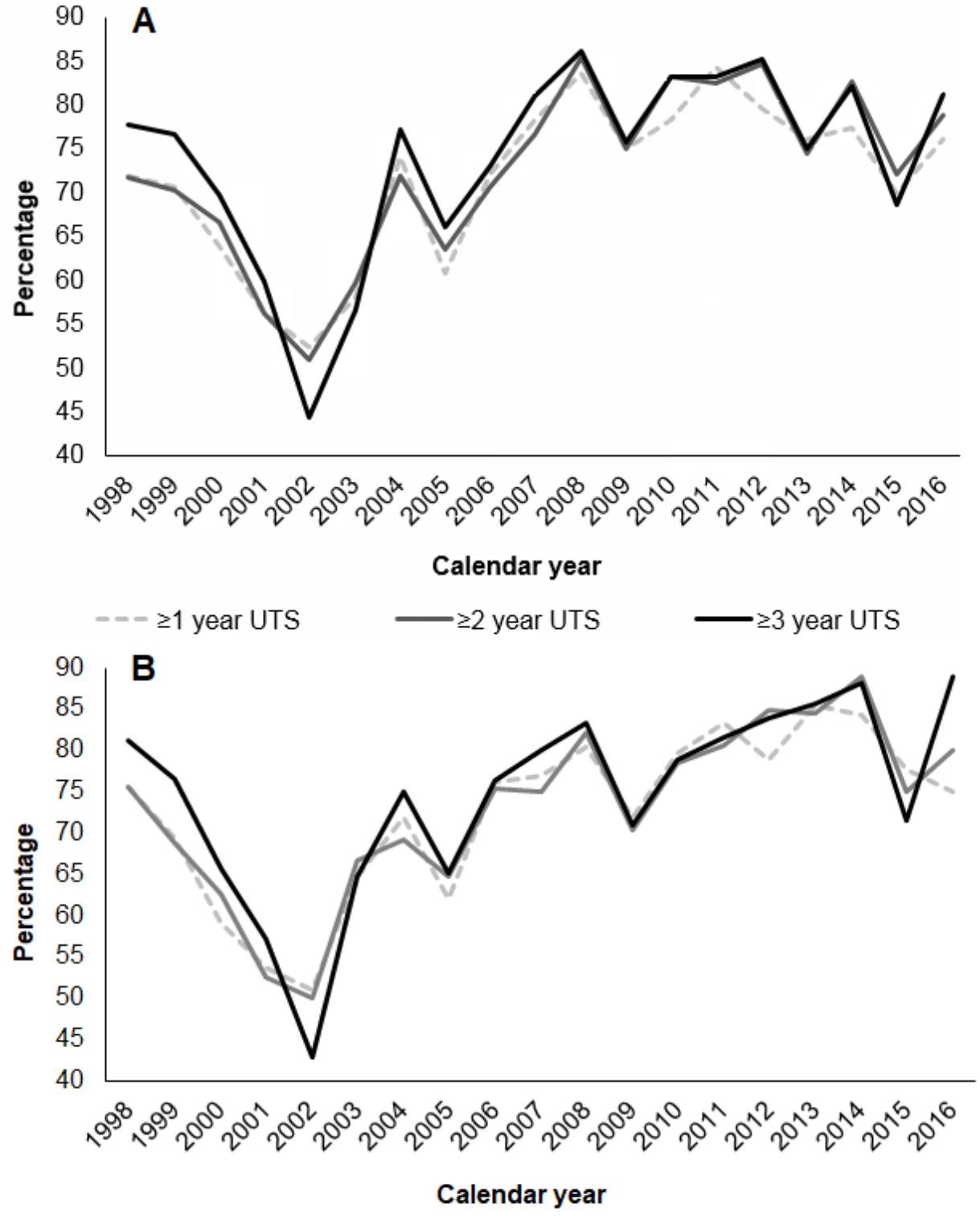
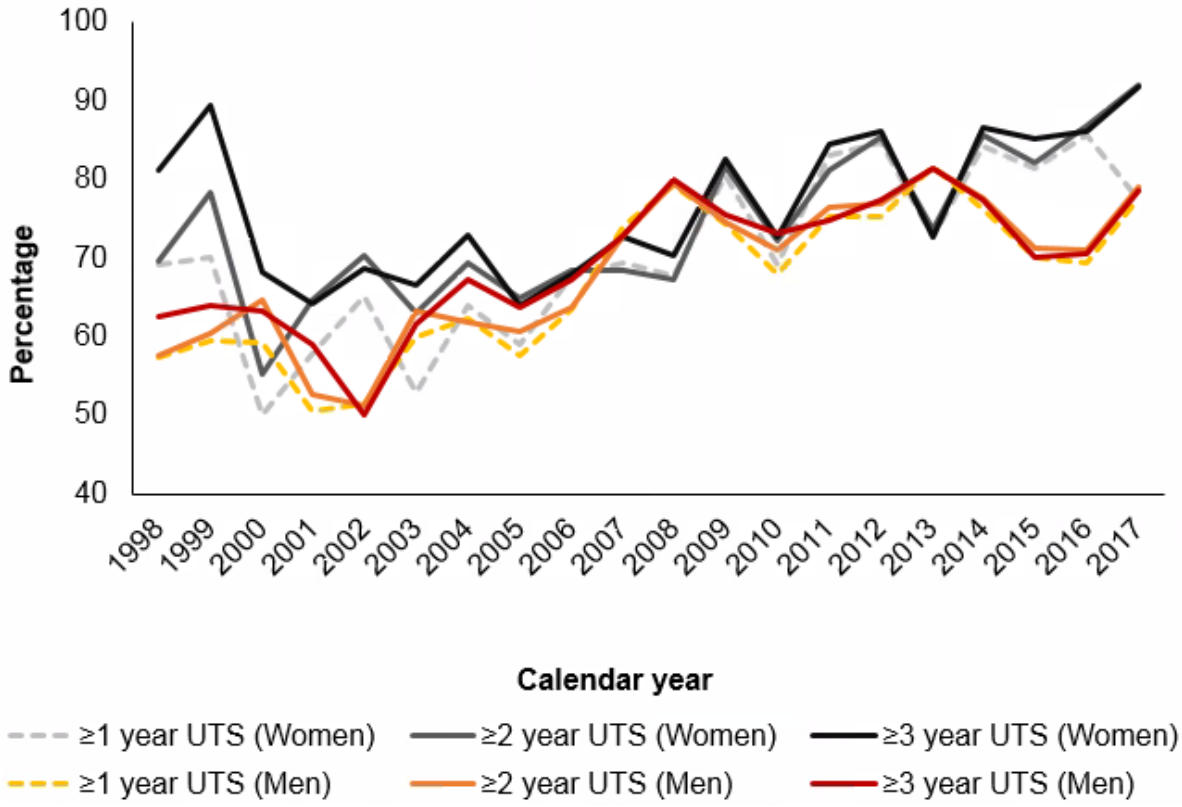


Figure 56. Annual percentage of patients diagnosed with AS (having ≥1, ≥2 and ≥3 prior years of UTS) who had a prior back-pain symptom code in sensitivity analyses, 1998-2016: A) AS1 (N = 1,071; N = 957; N = 821); B) AS2 (N = 843; N = 751; N = 634)



Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later

Figure 57. Annual percentage of women and men diagnosed with AS (having ≥1, ≥2 and ≥3 prior years of UTS) who had a prior back-pain symptom code, 1998-2017 (N = 3,101; N = 2,734; N = 2,417)

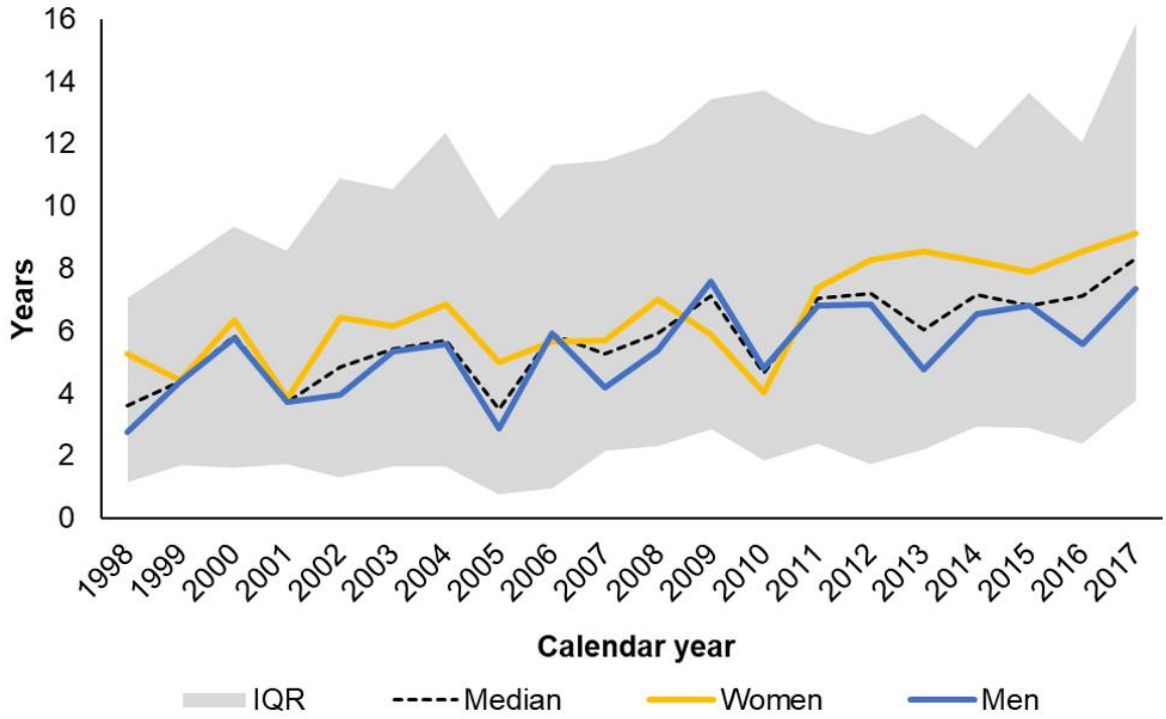


The median time from symptom to diagnosis was 5.97 years (IQR = 1.88-11.56), higher in women (6.71 years [2.30-12.36]) than men (5.65 years [1.66-11.20]) and in sensitivity analyses with ≥2 and ≥3 prior UTS years (6.20 years [IQR = 2.17-11.82]; 6.60 ([2.41-12.02] respectively). In sensitivity analysis AS1 (additional AS-related code >7 days later) the time to diagnosis was lower than in the primary analysis (5.21 years [IQR = 1.70-10.86]), again higher in women (6.03 [2.38-11.99]) than men (4.89 [1.42-10.40]) and higher in UTS-related sensitivity analyses (5.49 years [1.89-11.20]; 6.01 [2.20-11.30]). Similarly, in sensitivity analysis AS2 (additional AS-related code ≥180 days later) the time to diagnosis was lower than in the primary analysis (5.15 years [IQR = 9.16]), higher in women (6.20 [9.49]) than men (4.69 [9.01]) and higher in UTS-related sensitivity analyses (5.39 years [9.07]; 5.96 [8.91]).

During the study period, time from first back-pain symptom to diagnosis more than doubled from 3.62 (IQR = 1.14-7.07) years (5.26 [1.73-6.89] in women, 2.74 [1.12-7.09] in men) in 1998 to 8.31 (3.77-15.89) years (9.12 [5.59-12.29] in women, 7.33 [1.99-16.43] in men) in 2017 (Figure 58). The increase over time was less clear in sensitivity analyses although again time to diagnosis remained generally higher in women (Figure

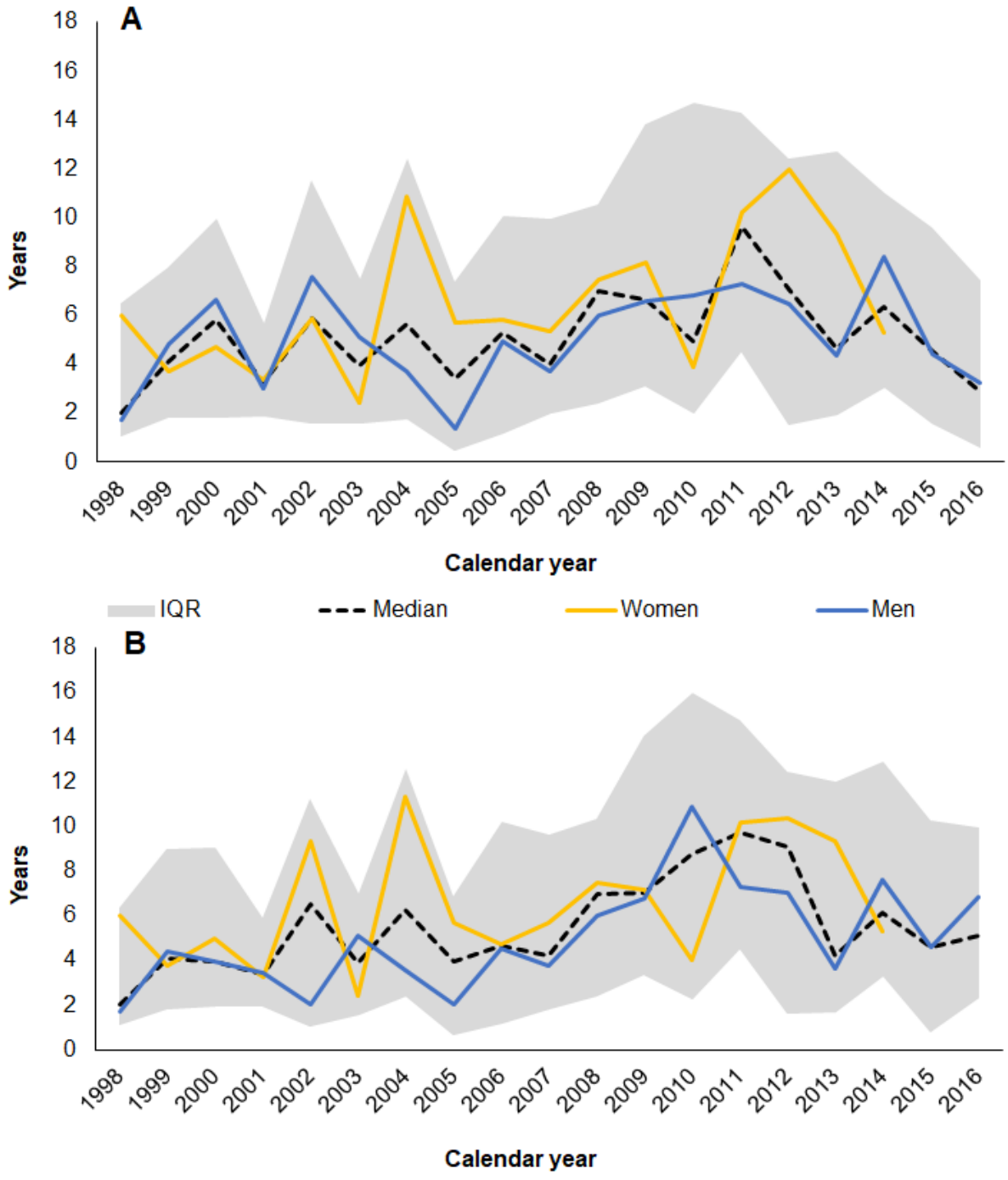
59). UTS-related sensitivity analyses showed consistent trends although the time to diagnosis was higher (Figure 60, Figure 61).

Figure 58. Annual median time in years from first recorded back-pain symptom to diagnosis, with interquartile range, 1998-2017 (N = 2,120)



Note: Dashed line represents the overall annual diagnostic delay, and the shaded area the interquartile range (IQR)

Figure 59. Annual median time in years from first recorded back-pain symptom to diagnosis in sensitivity analyses, with interquartile range, 1998-2016: A) AS1 (N = 757); B) AS2 (N = 592)



Note: Dashed line represents the overall annual diagnostic delay, and the shaded area the interquartile range (IQR); AS1 and AS2 = additional AS-related code >7 and ≥180 days later, respectively

Figure 60. Median time in years from back-pain symptom to diagnosis, for patients with ≥1, ≥2 and ≥3 years prior UTS registration, 1998-2017 (N = 2,120; N = 1,929; N = 1,750)

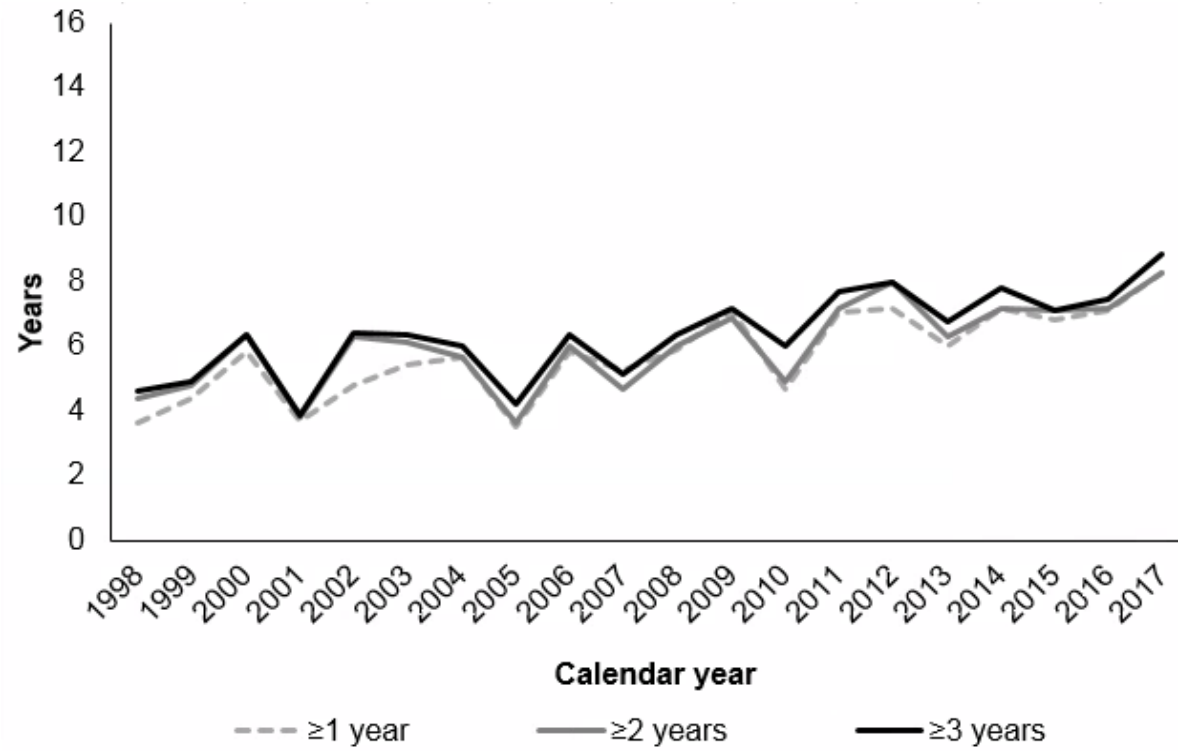
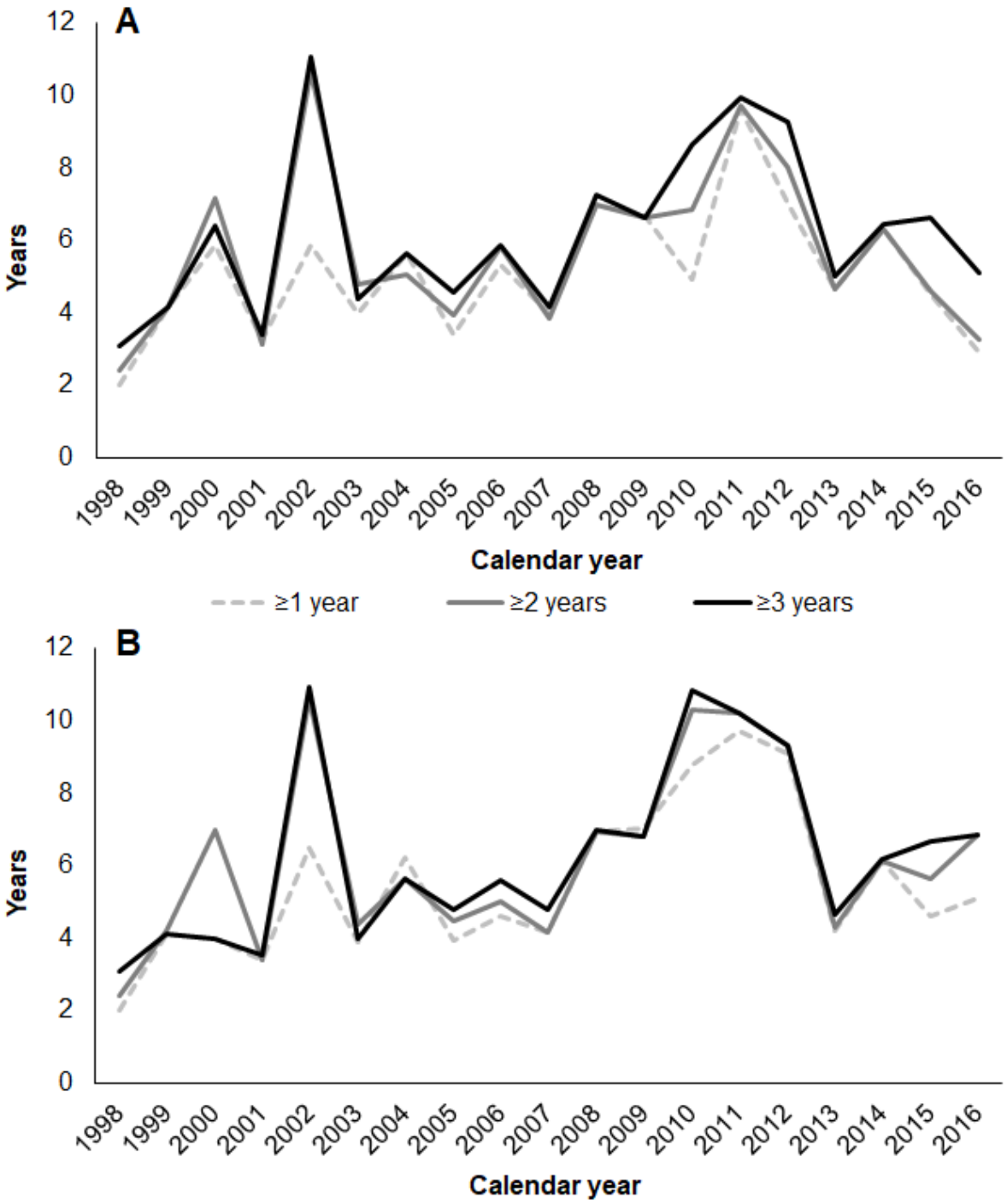


Figure 61. Median time in years from back-pain symptom to diagnosis, for patients with ≥1, ≥2 and ≥3 years prior UTS registration in the sensitivity analyses, 1998-2016: A) AS1 (N = 757; N = 688; N = 606); B) AS2 (N = 592; N = 533; N = 463)

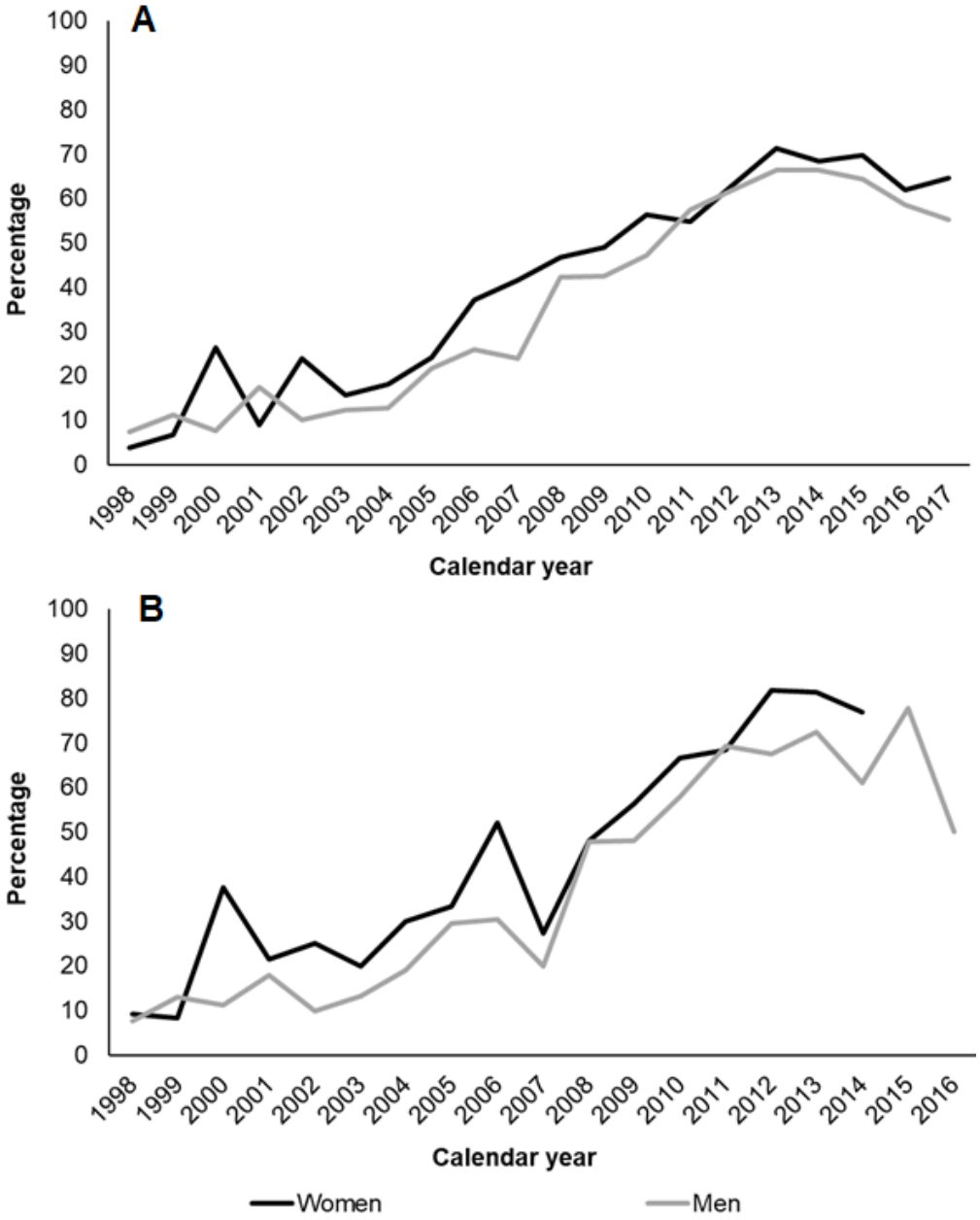


Note: AS1 = additional AS-related code >7 days later; AS2 = additional AS-related code ≥180 days later

8.3.2 Time from Rheumatology Referral to Diagnosis

The proportion of patients with a recorded prior rheumatology referral was 37.6% (n = 1,167), higher in women (42.1%) than men (35.7%). Recording of rheumatology referrals increased over time; 6.6% of patients diagnosed with AS had a rheumatology referral recorded in 1998, rising to 58.2% in 2017 (Figure 62). The proportion with a recorded referral was comparable in sensitivity analysis AS1 (37.3%, n = 399; 43.5% in women, 34.7% in men), rising from 8.0% in 1998 to 52.4% in 2016.

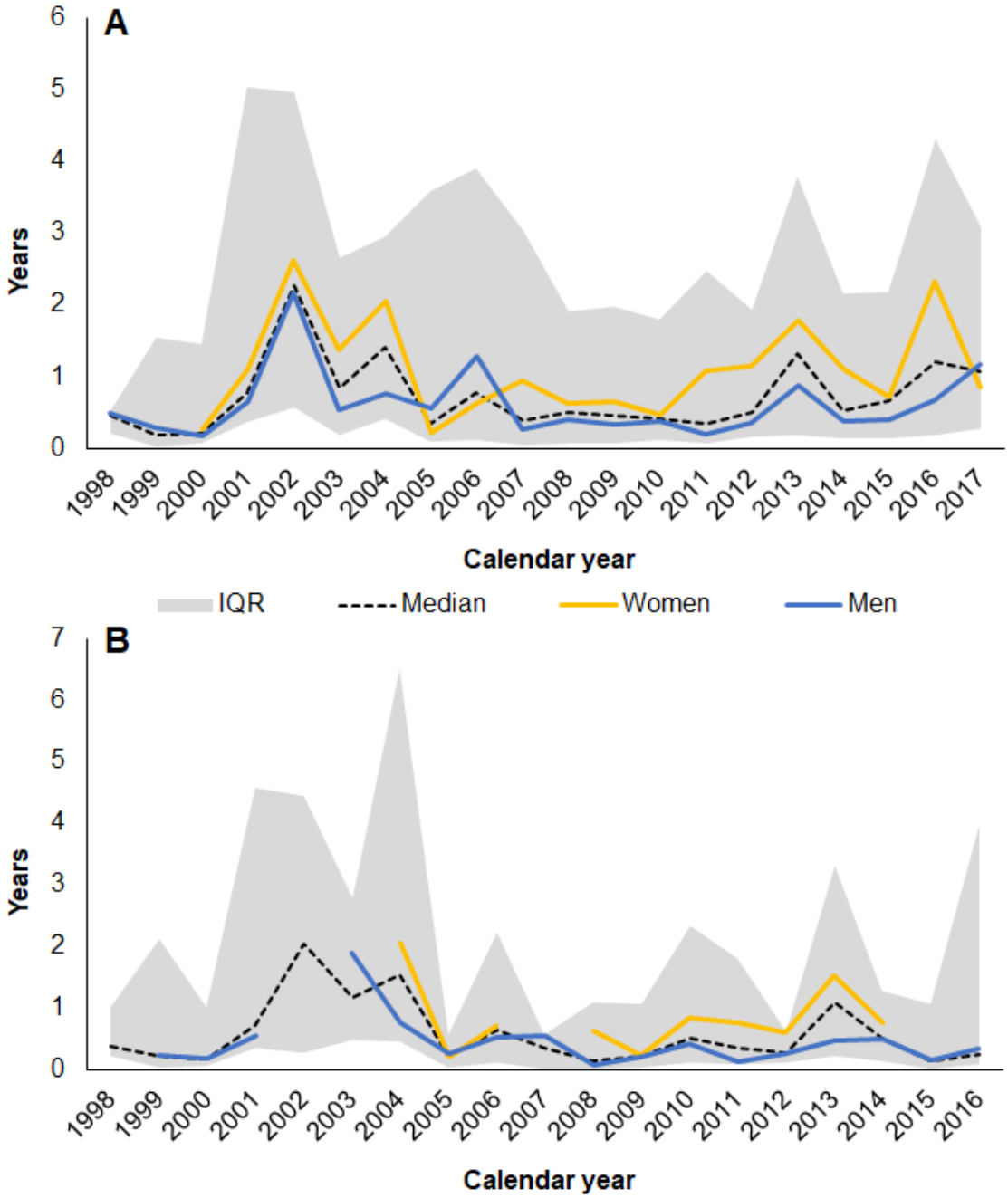
Figure 62. Annual percentage of women and men diagnosed with AS that had a prior rheumatology referral recorded: A) primary analysis, 1998-2017 (N = 3,101); B) sensitivity analysis AS1, 1998-2016 (N = 1,071)



Note: Data suppressed where there are ≤5 cases; AS1 = additional AS-related code >7 days later

The median time from rheumatology referral to diagnosis was 0.63 years (IQR = 0.13-2.74; n = 1,167), two times higher in women (1.00 [0.23-3.21]) than men (0.48 [0.08-2.63]) and with no clear change over time (Figure 63). The median time from referral to diagnosis was lower in sensitivity analysis AS1 (0.39 years, IQR = 0.06-1.92, n = 399), again higher in women (0.63 [0.13-2.04] than men (0.31 [0.06-1.57]) and with no clear temporal change.

Figure 63. Annual median time in years from first rheumatology referral to diagnosis, for women and men: A) primary analysis, 1998-2017 (N = 1,167); B) sensitivity analysis AS1, 1998-2016 (N = 399)



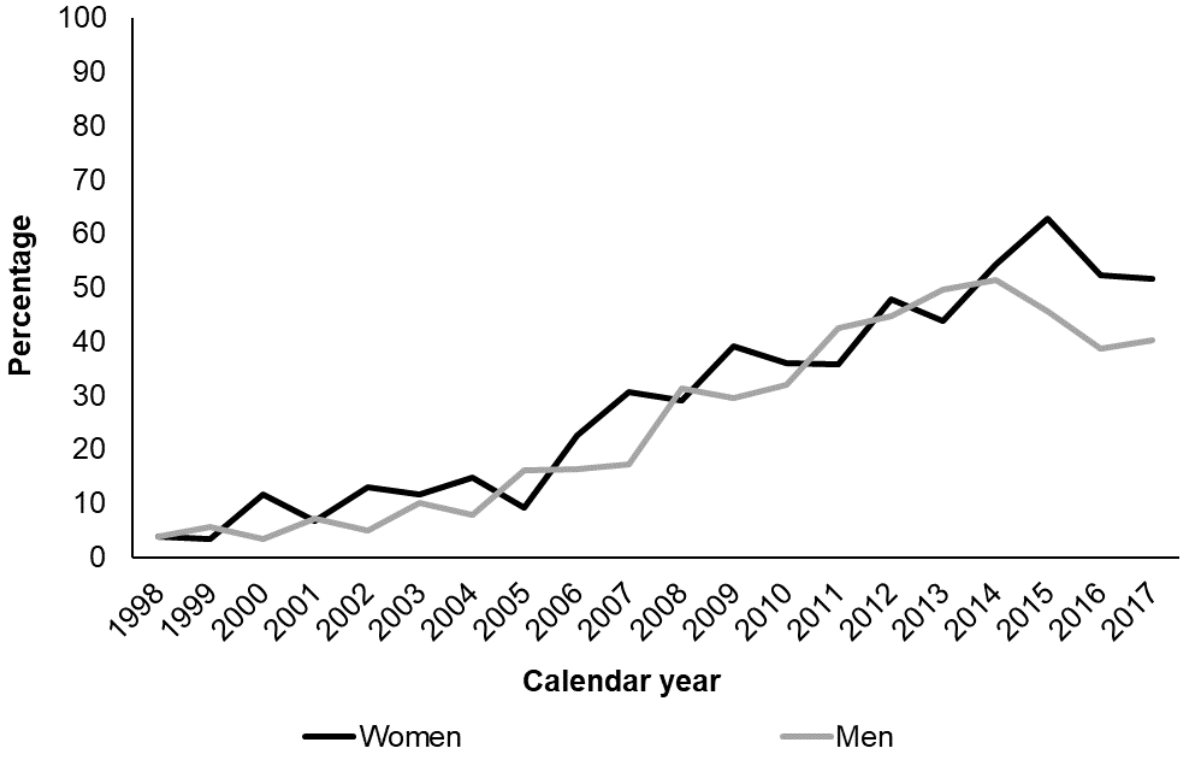
Note: Data suppressed where there are ≤5 cases; AS1 = additional AS-related code >7 days later

8.3.3 Time from Symptom to Rheumatology Referral

The proportion of AS patients with both a prior recorded back-pain symptom and rheumatology referral was 26.4% (n = 819), 29.6% in women and 25.0% in men, and rose from 3.8% in 1998 to 43.8% in 2017 (Figure 64). Similarly, in sensitivity analysis

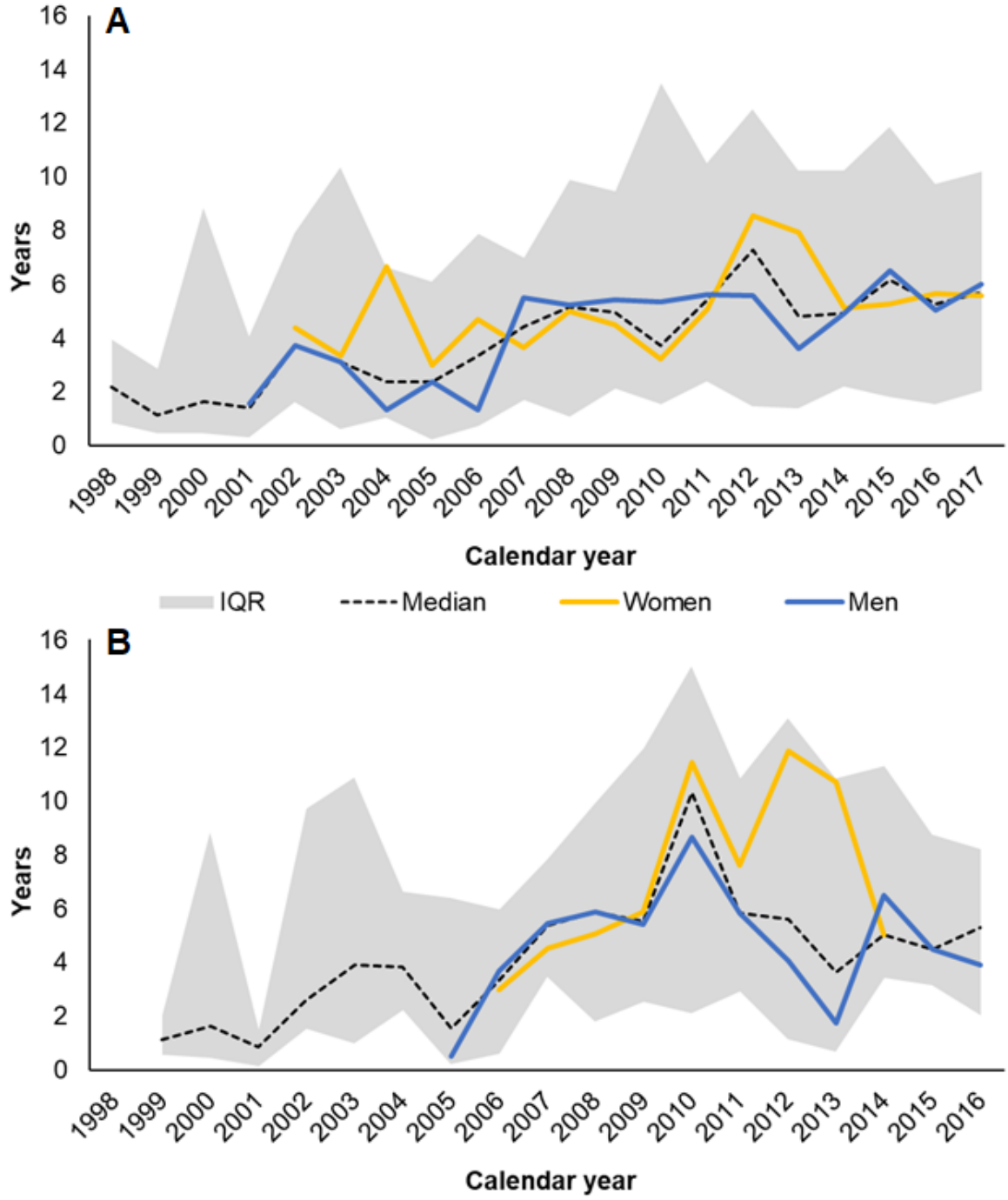
AS1, 26.1% (n = 279) had a prior recorded symptom and referral, although there were ≤5 cases per sex in many years (data not shown).

Figure 64. Annual percentage of women and men diagnosed with AS who had prior back pain and rheumatology referral recorded, 1998-2017 (N = 3,101)



The median time from first recorded back-pain symptom to rheumatology referral was 4.87 years (IQR = 1.42-10.23; n = 819), higher in women (5.16 [1.80-11.09]) than men (4.51 [1.22-9.72]). This median time more than doubled between 1998 (2.16 [IQR = 0.83-3.95]) and 2017 (5.70 [2.03-10.21]) (Figure 65). In sensitivity analysis AS1, the median time from symptom to rheumatology referral was 4.88 years (IQR = 1.20-10.47; n = 279), almost two years higher in women (5.86 [1.80-13.22]) than men (4.02 [0.94-9.42]) and also rose over time.

Figure 65. Annual median time in years from first recorded back pain to rheumatology referral, for women and men: A) primary analysis, 1998-2017 (N = 819); B) sensitivity analysis AS1, 1998-2016 (N = 279)



Note: Data suppressed where there are ≤5 cases; AS1 = additional AS-related code >7 days later

8.4 Discussion

This study showed a trend over the last two decades to worsening in delay to rheumatology referral and AS diagnosis, despite the diagnostic delay first being

highlighted in 1999 (229). The reported rise in diagnostic delay corresponds with survey reports of 6-year and 8.5-year delay in 2010 and 2016 respectively (228, 230). Diagnostic delay was a year longer for women, and following rheumatology referral, the time to diagnosis was twice as long as in men, with no apparent reduction in the sex difference over time. The sex difference in diagnostic delay of a year was greater than the 7-month difference reported in a systematic review of delay in all spondyloarthritides (516). In addition to a worsening trend and persistent sex difference, UTS-related sensitivity analyses suggest even greater delay in diagnosis. Further, there may be an unmeasured additional period of diagnostic delay, as a survey by Hamilton et al. highlighted a delay in presentation of AS symptoms to primary care (228).

The recording of back pain symptoms and referral to rheumatology increased over time. The publication of BSR guidelines on therapy in AS in 2005 may have contributed to increasing awareness among GPs of AS symptoms and the importance of referral for rheumatology-led biologic therapy (240). This is in accordance with a similar increase in screening for AS that was suggested in Chapter 7. The apparent rise in referral is corroborated by survey reports that show the proportion of patients with AS currently attending a rheumatology clinic was 68% in 2010 and 82% in 2016 (228, 230). Increasing referral to rheumatology prior to diagnosis may have increased the proportion of the diagnoses being made by rheumatologists, as the surveys also suggest this rose from 70% in 2010 to 85% in 2016 (228, 230).

The apparent delay to referral suggests a necessity to raise awareness of IBP and the associated features of AS among non-rheumatologists. A survey of GPs reported that only 13% and 50% respectively recognised alternating buttock pain and pain improving with exercise as symptoms of IBP, and 60% recognised uveitis as an associated AS feature (517). A pathway to aid recognition and referral of IBP may help in reducing diagnostic delay, as has been reported in a UK military setting (518). Recognition of associated features should also be encouraged. Where MRI scans are performed to investigate persistent back pain, a study has suggested that inflammation-detecting sequences should be included (513). Given the sex difference in the apparent delay following rheumatology referral, inclusion of inflammation-detecting sequences should especially be considered in women. Encouragingly, UK guidelines published subsequently in 2019 have highlighted the importance of early referral to rheumatology in AS (223). Future studies could examine the impact of this revised guideline on diagnostic delay in the forthcoming years.

8.4.1 Strengths and Limitations

The methods appraisal is largely presented in Chapter 10, with chapter-specific strengths and limitations reported here.

Study strengths included using data from primary care for a large population-based cohort, potentially capturing early presentations with symptoms of AS, and ensuring that GP-diagnosed cases are not missed (228). Sensitivity analyses AS1 and AS2 (requiring additional codes in >7 and ≥180 days) improved the specificity of the definition of AS and confirmed the robustness of the study findings. The results for the primary outcome were comparable between sensitivity analyses AS1 and AS2, suggesting that it was sufficient to just perform the one sensitivity analysis (AS1) for the secondary outcomes. The long-term follow-up of patients (median 12.61 years) was important for the identification of early symptoms and rheumatology referrals recorded in the years prior to the diagnosis of AS. The longer diagnostic delays revealed in UTS-related sensitivity analyses, restricted to patients with longer quality registration periods prior to the index date, highlighted the importance of long-term quality follow-up for capturing earlier symptom presentations.

The study limitations include those common to EHR-based studies as described in Chapters 2 and 4, e.g. incomplete and changing coding practices. Increased symptom recording over time will have improved the accuracy of measures of diagnostic delay. Diagnostic delay due to delayed presentation of symptoms to primary care was not examined, though survey data suggest this to be ≤12 months in the majority of cases (228). Recording of rheumatology referrals increased over time but it was not possible to determine whether referrals occurred without this being recorded using RCV2. However, a comparison between the proportion of patients with a rheumatology referral recorded prior to AS diagnosis (49% in 2010, 60% in 2016), and the proportion of patients under the care of a rheumatologist (68% in 2010 and 82% in 2016), suggests a high level of recording given the ongoing role of rheumatology in AS biologic therapy post-diagnosis (228, 230). A survey by Hamilton et al. reported that, over time in the UK, diagnoses of AS have increasingly been made in rheumatology rather than primary care (519); this may have increased the calculated diagnostic delay in this study over time as a temporal lag in the recording of rheumatology-led diagnoses in GP records is expected. However, the time from rheumatology referral to AS diagnosis contributed only marginally to the overall diagnostic delay. While the study used UK data, the considerations for improving the education and recognition of IBP and development of a care pathway may be more universal. In the USA, a recent study also highlighted in

healthcare professionals only partial recognition of the features of IBP and inconsistent approaches to diagnostic investigation and management (520).

There were issues in defining first symptom presentation based on recorded back pain symptoms. The specificity of this approach was constrained by the commonality of back pain recording in GP EHRs for numerous other acute and chronic reasons. Other causes of back pain include poor posture, sciatica, and trauma such as a fall or accident. In addition, other symptoms of AS besides back pain may have been presented, particularly peripheral joint and tendon pain, and other extra-articular features such as uveitis, and these were not measured. A more nuanced approach may be required to assess the validity of the approach used in defining the date of the first presentation with AS. The robustness could be assessed through a sensitivity analysis that excluded back pain symptoms recorded in the same consultation as an observation or diagnosis to which the back pain could be attributed, and considered additional AS-related symptoms, NSAID prescriptions and EAM diagnoses.

The diagnostic delay in this study related to the time from first symptom presentation (to a GP), rather than the time from disease onset, which is more meaningful from a disease management perspective. As discussed above, there may be a period of disease activity prior to such symptoms being presented to a clinician. Symptoms of back pain have many more commonly known causes, and patients may attempt to self-manage these or consult with another healthcare service such as physiotherapy or osteopathy. While a patient-reported account of symptom duration may be recorded by GPs in EHRs, this would only be recorded in free text. However, as described in the RA scoping review (Chapter 6), EHR data can be used in efforts to define disease onset. In some diseases, phenotyping algorithms are reported that predict disease onset based on factors including symptoms, referrals, tests and prescriptions. A recent study has identified several predictors of AS present in EHRs prior to diagnosis, including episodes of axial pain separated by >6 months and co-occurrence of axial pain with NSAID prescriptions (521). The study also reported that coded episodes of axial pain increased in frequency in the three years pre-diagnosis. Further research could investigate the time to diagnosis from such early indicators of disease onset, and whether the time to diagnosis is influenced by the frequency of consultations regarding AS-related symptoms and EAMs.

8.4.2 Conclusion

This study suggests that the delay to diagnosis has persisted over twenty years and appears largely driven by delay in referral to rheumatology. The study also highlighted the importance of long-term quality data follow-up in determining diagnostic delay. The worsening trend in time to diagnosis and the worse delay in women, even following rheumatology referral, is of concern given the importance of early therapy initiation for treatment success. Much effort is required to promote the education and recognition of IBP and associated AS features among non-rheumatology health practitioners, and to prompt early rheumatology referral.

Chapter 9 Trends in the Pharmacologic Management of Rheumatoid Arthritis in Primary Care

9.1 Introduction

This chapter addresses the fourth thesis objective of describing trends in real-world management using EHR data. As described in Chapter 2, RA is a common inflammatory arthritis of increasing relevance in our ageing population. As also noted, it usually first presents in primary care as joint pain and swelling and is managed through prescription of medication that is initiated by rheumatologists prior to co-management with GPs (261, 289). It was described how the principles of RA management have shifted over recent decades to immediate initiation and more effective escalation of DMARDs, following mounting evidence for the efficacy of early aggressive therapy for inflammation suppression (255, 290, 292, 522). As noted in Chapter 6, increased prescribing of DMARDs in RA management has been reported in the year post-diagnosis, especially following publication of BSR guidelines in 2006 up to 2010 (477-480). However, estimates could be extended to more recent calendar years to examine diversion from the guidance, especially following the publication of national guidelines in 2009, recommending immediate commencement of DMARDs upon diagnosis (280). In addition, it is important to understand prescribing in the prevalent RA population and across the life-course.

Chapter 2 also described modern guideline recommendations in the UK for prophylaxis, where long-term corticosteroids or NSAIDs are prescribed, given the potential for toxicity even at low doses (216, 238, 280, 295, 302). Although corticosteroids can mask uncontrolled disease activity, a short-term course is recommended when initiating or changing DMARDs (32). While there are varying approaches regarding corticosteroid dose and duration, tapering is recommended “as rapidly as clinically feasible”, guided by response and risk factors (280, 298, 523). However, in the UK, between 1992 and 2009 the reported median duration for GP-prescribed corticosteroids among RA patients was 0.8 years (IQR = 0.15–2.56) (524). It is unclear whether the prescribing duration has reduced following the publication of national guidelines in 2009 recommending a prescribing duration of <3 months for prednisolone (302). The publication of national guidelines in 2008 for GI prophylaxis co-prescribing alongside NSAIDs was also reviewed in Chapter 2 (238, 302). However, the pattern of corticosteroids and NSAID prescribing and prophylaxis is uncertain. An

evaluation of the trends in corticosteroid and NSAID prescribing and prophylactic co-prescribing in RA patients would provide information on the impact of guideline recommendations and inform future policymaking.

This chapter aims to explore trends in prescribing for RA and prophylactic therapy between 1998 and 2017, following the shift in modern management. The temporal pattern of DMARD, corticosteroid and NSAID prescribing will be investigated, for all RA patients and in the year post-diagnosis, and across the patient life-course, as well as concomitant prescribing of pharmacologic prophylaxis. RA management with corticosteroids and NSAIDs will be described in comparison to a non-RA population. This study will establish whether modern use of DMARDs and tighter control of inflammation has facilitated reduced long-term use of potentially toxic corticosteroids and NSAIDs, and evaluate any changes in prophylaxis prescribing. This may inform and update understanding of the management of RA and guideline compliance in the UK, with relevance for policy-making.

9.2 Methods

A population-based retrospective longitudinal observational study using GP EHR data was reported, following the STROBE guidelines (322). The study protocol approval, data source, study timeframe and study period were described in Chapter 3. In addition to the patient eligibility criteria defined in Chapter 3, the study population for the analysis in this chapter also excluded patients with RA diagnosed while aged below 18 years, or juvenile RA diagnosis. The criteria for the start and end of patient follow-up are as defined in Chapter 3.

9.2.1 Rheumatoid Arthritis Cohort

Rheumatoid arthritis was defined using Read Codes as described in Chapter 7. Two cohorts of patients with RA were identified using the code-list: a cohort of all patients with RA, and a cohort of patients with incident RA. All RA patients were identified by having ≥1 instance of these codes in their clinical data. Follow-up for the RA cohort commenced on the date of RA diagnosis or study follow-up start, whichever was latest. Incident RA cases had no diagnosis prior to follow-up commencement for a given analysis, meaning that they had ≥1 year of data prior to incident diagnosis. This excluded from the incidence cohort any prevalent cases that were incorrectly recorded

as an incident diagnosis instead of as medical history of RA during GP registration (396).

Sensitivity analyses RA1 and RA2 were performed, as defined in Chapter 7, to assess any impact of cohort specificity on the results. These excluded patients in which RA diagnosis was not confirmed (RA1: ≥2 RA diagnoses at least 6 months apart) or followed by an RA-specific prescription (RA2: ≥1 RA diagnosis and subsequent DMARD). The sensitivity analyses were based on approaches used in previous studies, which showed that DMARD medication or multiple diagnosis codes increased the validity of RA diagnosis by GPs (458, 504).

To explore the impact of an RA definition on prescribing assessment, sub-analysis A was performed in sensitivity analysis RA1: in this, the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis.

9.2.2 Non-Rheumatoid Arthritis Cohort

To assess long-term medication use, for each patient with a diagnosis of RA during the study period, five non-RA patients were randomly selected and matched by sex and date of birth ±5 years from amongst patients alive and registered at the same practice on the index date (latest of the RA patient's study follow-up start or first recorded RA diagnosis during the study period). Non-RA patients had no RA diagnosis prior to or in the 6 months following the index date, to enable comparison of long-term management practices. A random matching process was used without replacement. Patient follow-up (and that of the matched RA cohort) commenced on the index date and ended when one of the following occurred: end of study follow-up, end of study follow-up for the matched RA patient, or if the patient was diagnosed with RA.

9.2.3 Medication Definition

The medication prescribed during the study follow-up was identified using BNF terms. Prescriptions in CPRD have a drug product name, drug substance name and a BNF ID that links to a BNF chapter, as described in Chapter 3. For NSAIDs, the BNF chapter relevant to RA was identified with clinical guidance (PGC) as '10010100', 'Non-Steroidal Anti-Inflammatory Drugs'. The distinct product and drug substance names were determined from all prescriptions for drugs in this chapter. Corticosteroids and DMARDs belonged to less specific BNF chapters such as 'Corticosteroids and Other

Immunosuppressants' or 'Drugs Affecting the Immune Response'. Therefore, selection of these drugs was informed by consideration of drug lists on the NHS and Versus Arthritis websites and the Yorkshire DMARD guidelines (304, 505-507).

Relevant prescriptions were defined by having one of the listed terms in their product or drug substance name and being administered by the relevant route (e.g. oral for corticosteroids), regardless of the associated BNF chapter. Table 34 defines how the medications (NSAIDs, corticosteroids, DMARDs, prednisolone, PPIs, and bone protectants vitamin D, bisphosphonate and calcium) were selected.

Table 34. Drugs used to determine prescribed medication

Medication
Disease-modifying anti-rheumatic drugs Route: gastroenteral, intraarterial, intravenous, oral, subcutaneous Term: Abatacept; Adalimumab; Azathioprine; Baricitinib; Certolizumab; Ciclosporin / cyclosporine; Cyclophosphamid/e; Etanercept; Gold injections / injectable gold / sodium aurothiomalate; Golimumab; Hydroxychloroquine; Infliximab; Leflunomide; Methotrexate; Mycophenolate / mycophenalte mofetil; Penicillamine; Rituximab; Sarilumab; Sulfasalazine; Tocilizumab; Tofacitinib; Ustekinumab
Corticosteroids Route: oral Term: Betamethasolone; Betamethasone; Bethamethasone; Budesonide; lobetasone; Cortisone; Deflazacort; Dexamethasone; Fluorometholone; Hydrocortisone; Loteprednol; Methylprednisolone; Prednisolone; Prednisone; Rimexolone; Triamcinolone
Non-steroidal anti-inflammatory drugs Route: cutaneous, oral, rectal, topical, transdermal Term: Aceclofenac; Acemetacin; Celecoxib; Dexibuprofen; Dexketoprofen Trometamol; Diclofenac potassium; Diclofenac sodium; Misoprostol; Diflunisal; Etodolac; Etoricoxib; Fenbufen; Fenoprofen calcium; Flurbiprofen; Ibuprofen; Indometacin; Ketoprofen; Lornoxicam; Magnesium trisilicate; Mefenamic acid; Meloxicam; Nabumetone; Naproxen; Naproxen sodium; Phenylbutazone; Piroxicam; Piroxicam betadex; Salsalate; Sulindac; Tenoxicam; Tiaprofenic acid; Tolmetin sodium
Prednisolone Route: oral Term: Methylprednisolone; Prednisolone; Prednisone
Proton pump inhibitors Route: oral Term: Esomeprazole; Lansoprazole; Omeprazole; Pantoprazole; Rabeprazole
Bisphosphonates Route: oral Term: Alendronate sodium; Alendronic acid; Ibandronic sodium monohydrate; Risedronate sodium
Vitamin D Route: oral Term: Alfacalcidol; Calcitriol; Colecalciferol; Dihydrotechysterol; Ergocalciferol; Paricalcitrol
Calcium

Route: oral Term: Calcium; Calcium carbonate; Calcium chloride dihydrate; Calcium gluconate; Calcium lactate; Tricalcium phosphate; and excluding the phrase 'indigestion' BNF chapter: contains the term vitamin, supplement or not stated

Chapter 6 highlighted that it is not clear from other studies how medications are selected. To investigate this, an alternative approach was applied in sub-analysis B, to define DMARDs, corticosteroids and NSAIDs. From the prescriptions identified as described above, medications were selected only if their BNF ID were in RA-relevant BNF chapters (Table 35). Further, only the oral corticosteroid terms used in a CPRD GOLD study by Black et al. 2015 were used (486). This study reported on the proportion of RA patients with an oral corticosteroid prescription between 1992 and 2009.

Table 35. Drugs used to determine prescribed medication in sub-analysis B

Medication
Disease-modifying anti-rheumatic drugs Route: gastroenteral, intraarterial, intravenous, oral, subcutaneous Term: Abatacept; Adalimumab; Azathioprine; Baricitinib; Certolizumab; Ciclosporin / cyclosporine; Cyclophosphamid/e; Etanercept; Gold injections / injectable gold / sodium aurothiomalate; Golimumab; Hydroxychloroquine; Infliximab; Leflunomide; Methotrexate; Mycophenolate / mycophenalte mofetil; Penicillamine; Rituximab; Sarilumab; Sulfasalazine; Tocilizumab; Tofacitinib; Ustekinumab BNF chapter: 08020300 Drugs Affecting The Immune Response; 08020400 Drugs Affecting The Immune Response; 10010302 Drugs Affecting The Immune Response In Rheumatic Disease; 10010300 Rheumatic Disease Suppressant Drugs; 10010301 Cytokine Modulators in Rheumatic Disease
Corticosteroids Route: oral Term: Bethamethasone; Budesonide; Cortisone; Deflazacort; Dexamethasone; Methylprednisolone; Prednisolone; Prednisone; Triamcinolone BNF chapter: 06030200 Glucocorticoid Therapy; 08020200 Corticosteroids and Other Immunosuppressants; 10010200 Corticosteroids; 10010201 Systemic Corticosteroids (in Musculoskeletal and Joint Conditions); 11040100 Corticosteroids (in Eye preparations)
Non-steroidal anti-inflammatory drugs Route: cutaneous, oral, rectal, topical, transdermal Term: Aceclofenac; Acemetacin; Celecoxib; Dexibuprofen; Dexketoprofen Trometamol; Diclofenac potassium; Diclofenac sodium; Misoprostol; Diflunisal; Etodolac; Etoricoxib; Fenbufen; Fenoprofen calcium; Flurbiprofen; Ibuprofen; Indometacin; Ketoprofen; Lornoxicam; Magnesium trisilicate; Mefenamic acid; Meloxicam; Nabumetone; Naproxen; Naproxen sodium; Phenylbutazone; Piroxicam; Piroxicam betadex; Salsalate; Sulindac; Tenoxicam; Tiaprofenic acid; Tolmetin sodium BNF chapter: 10010100 Non-Steroidal Anti-Inflammatory Drugs

Note: Sub-analysis B = with BNF chapter constraint so that drugs must be assigned a BNF code in a specific chapter, and using only corticosteroid terms listed by Black et al (2015) (486)

9.2.4 Outcomes

The primary outcomes were the annual mean prescription count and proportion of patients with ≥1 and ≥6 DMARD, oral corticosteroid or NSAID prescriptions, and the annual proportion with long-term (≥90 days) prescribing. For oral corticosteroids and NSAIDs, these proportions were compared with a non-RA cohort. A secondary outcome was the proportion with a long-term (≥90 days) duration of DMARD, oral corticosteroid or NSAID prescribing per year across the life-course. Another secondary outcome was the annual proportion with prophylaxis co-prescribing (for ≥90 days) alongside concomitant NSAID or high (≥7.5 mg) / low (<7.5 mg) dose oral prednisolone prescribing. Appropriate prophylaxis co-prescribing was defined as prescribed bone-protectants with oral prednisolone among women without prior osteoporosis and PPIs with NSAIDs, as described in Chapter 2.

9.2.5 Statistical Analysis

Baseline cohort characteristics were described for the all RA, incident RA and non-RA cohorts. Baseline prevalence of comorbidities asthma, chronic obstructive pulmonary disease (COPD) and osteoarthritis were assessed as these are commonly treated with corticosteroids or NSAIDs. The comorbidities were defined using code-lists (RCV2) based on previous CPRD validation studies (99, 525). Resolved cases of childhood asthma were discounted by excluding patients with an ‘asthma resolved’ code and no subsequent asthma code. Annual trends in patient outcomes were reported per calendar year between 1 January 1997 and 31 December 2017. Calculations of prescribing in the year post-diagnosis were reported for patients diagnosed between 1 January 1997 and 31 December 2016. Sensitivity analyses RA1 (≥2 RA diagnoses 6 months apart) and RA2 (subsequent DMARD) were performed for measures of the primary outcome, between 1 January 1997 and 31 December 2016. Sub-analyses A (subsequent RA code ≥6 months later used to assign the date of diagnosis) and B (BNF chapter and corticosteroid term constraints) were performed for measures of the primary outcome also. Difference between years was defined as significant where the 95% confidence intervals did not overlap.

The annual mean count of DMARD, oral corticosteroid and NSAID prescriptions per person-year and APC were calculated, and standardised for duration of follow-up. The annual mean count was also calculated in patients receiving ≥1 DMARD, oral corticosteroid or NSAID prescription in a given year. These were calculated for all RA patients, incident (diagnosed during follow-up) RA patients in the year post-diagnosis,

matched RA and non-RA patients. The proportion with ≥6 prescriptions in a given year was calculated for all RA patients as an indicator of long-term use. The annual mean count of oral corticosteroid and NSAID prescriptions and APC were also calculated for RA patients following their first DMARD prescription post-diagnosis.

Prednisolone prescription counts across England were also counted for the period from which this data are also available publically, April 2015 to March 2018, to provide context (526).

Long-term prescribing was defined as ≥90 days total prescription duration within 12 months (527). In secondary analyses the prescription durations of ≥1 day and ≥180 days were also assessed. Appropriate prophylaxis co-prescribing was defined as long-term PPI alongside long-term NSAIDs and long-term bone protectant alongside long-term high or low dose oral prednisolone. Low and high dose prednisolone were defined as <7.5 mg and ≥7.5 mg respectively (302). Prescription durations were calculated between 1998 and 2017, using an algorithm previously published by Partington et al. (2018) (504):

1. "If available, duration of each prescription recorded in CPRD was used.
2. If not, the duration of each prescription was the lowest of,
    a. the quantity of medication prescribed; or
    b. the gap until the next prescription [*of that drug group*] (if this was <90days); or
    c. the quantity of medication prescribed divided by the daily dose (if this was recorded).
3. If the duration was still missing, it was replaced with,
    a. the average of that patient's duration for other prescriptions of the same drug with the same strength (if present); or
    b. the average duration for all other patients' prescriptions of the same drug with the same strength.
4. If prescription duration was >90 days, it was replaced as 90 days."

The annual percentage of RA patients with long-term medication prescribing (DMARD, oral corticosteroid and NSAID) was calculated, from patients with ≥90 days of follow-up in that year. This was measured in the non-RA cohort also (oral corticosteroid and NSAID). Similarly, for incident patients with ≥90 days of follow-up in each year, the

annual percentage with long-term medication prescribing in the first year post-diagnosis was determined in the 1998-2016 period (to enable ≥1 year of follow-up). The period percentage with ≥90 days of prescribing in any one year or in their first year since diagnosis was also calculated. These were also calculated for oral corticosteroid and NSAID, in patients that had prescribed DMARDs (i.e. excluding DMARD-naïve patients). For incident patients in the 1997-2017 period, the percentage with long-term prescribing (DMARD, oral corticosteroid, NSAID or combination) in each year from diagnosis up to the 20th year or stopping at an earlier year if the cohort size fell below 1000. Patients with ≥90 days of follow-up in that year were included. In secondary analyses, these percentages were calculated for patients with ≥1 and ≥180 days of prescribing, amongst patients with ≥1 and ≥180 days of follow-up in that period.

The trend of long-term oral corticosteroid and NSAID prescribing among RA patients and in the first year post-diagnosis was assessed. Poisson regression was used with (log) person time as the offset and GP practice as a random intercept to analyse changes by calendar year, sex, age-band and GP practice (random intercept) while controlling for the other respective variables, with robust standard errors and 95% CI (

Appendix C: regression modelling code) (528). Poisson regression was selected as the observations were independent counts within an observed time interval; the offset accounted for difference in the denominator (person-time) per group (529, 530). Incidence rate ratios were calculated, with the Delta method used to define standard error. Quasi-Poisson regression was used to address over-dispersion where the dispersion parameter for the fixed model was >1. Comparison with a zero-inflation model was made where GP practice was included as a random intercept to test that the high count of zeroes were not generated by a separate process to count values (531, 532). The Laplace approximation was attempted in the mixed model but where this produced errors it was removed by setting the number of adaptive Gauss-Hermite quadrature points to zero, giving a less exact approximation of GP practice effect (533). The final coefficient inclusion was determined using the AIC, Hausman test and comparison of the coefficients and residual deviance (534).

The association of socio-economic deprivation with long-term prescribing trends was assessed. Patients with a recorded Index of Multiple Deprivation (IMD) quintile (derived from patient postcode) were selected for a subset analysis. The Poisson regression analysis was repeated and the effect of inclusion of the socio-economic deprivation coefficient was determined using the AIC and comparison of the coefficients and residual deviance.

For patients with ≥90 days of NSAID medication prescribing in a given year or incident year post-diagnosis, the percentage with ≥90 days of PPI prescribed was calculated. A similar analysis of the percentage with ≥90 days of bone protectant medication was performed among women with ≥90 days of low or high dose prednisolone prescribed and no prior osteoporosis diagnosis (defined using RCV2; Appendix B: Table B 2). The high dose cohort was determined using the highest prescribed dosage for each patient during the year. In a secondary analysis, the high dose cohort was determined as patients having ≥90 days of high dose prednisolone during the year. The bone protectants bisphosphonate, calcium and vitamin D were assessed separately and in the following combinations: bisphosphonate or calcium and vitamin D; calcium and vitamin D; calcium, vitamin D and bisphosphonate. A smaller cohort size was anticipated for this analysis and so 95% CIs were calculated. In period calculations, the proportions with ≥90 days of NSAIDs and PPI or prednisolone and bone protectant in any same year were calculated.

9.3 Results

From 8,077,644 patients eligible for follow-up in the study period, 1,164 juvenile RA patients were excluded and 71,411 RA patients were identified (44,426 (62.2%) with ≥2 diagnoses; 45,438 (63.6%) with diagnosis and prescribed DMARD) (Figure 66). The majority of RA patients in the primary analysis were also in a sensitivity analysis (76.57%, n = 54,685). The median age at diagnosis was 57 (IQR = 23), 70.0% (n = 49,974) were female and 58.1% (41,509) had IMD recorded (Table 29). Asthma was recorded as 'resolved' in 403 of the 10,486 RA patients with an asthma diagnosis. During the study period 31,768 patients were identified with incident RA (18,809 with ≥2 diagnoses; 21,880 with diagnosis and prescribed DMARD), with median age of 61 (IQR = 22) at diagnosis and of whom 67.58% (n = 21,464) were female. During the study period 41,198 patients received an RA diagnostic code and were matched to 205,990 non-RA patients; 68.05% were female, with median ages at follow-up of 61 (IQR = 22) and 60 (IQR = 22) respectively.

Figure 66. Study flow diagram of cohort selection

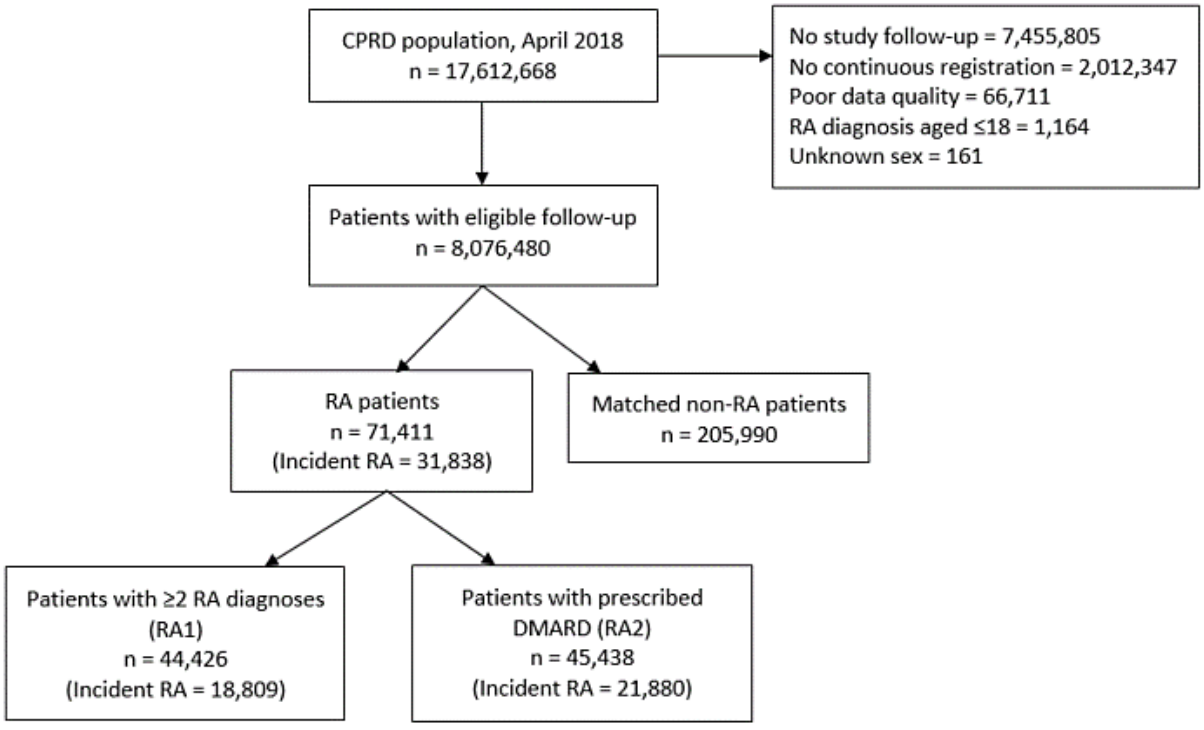


Table 36. Cohort baseline characteristics (at index date)

	All RA patients (N = 71,411)	Incident RA patients (N = 31,768)	Non-RA patients (N = 205,990)
Median age (years) [IQR]	57 [23]	61 [22]	60 [22]

Females (%)	49,974 (70.0)	21,464 (67.6)	140,685 (68.3)
Median follow-up duration (years) [IQR]	5.1 [7.6]	4.4 [6.1]	3.6 [5.4]
Asthma (%)	10,083 (14.1)	5,130 (16.1)	23,672 (11.5)
COPD (%)	4,302 (6.0)	2,101 (6.6)	7,431 (3.6)
Osteoarthritis (%)	18,551 (26.0)	9,632 (30.3)	31,121 (15.1)

Note: IQR = interquartile range; COPD = Chronic obstructive pulmonary disease

9.3.1 Trends in DMARD Prescribing

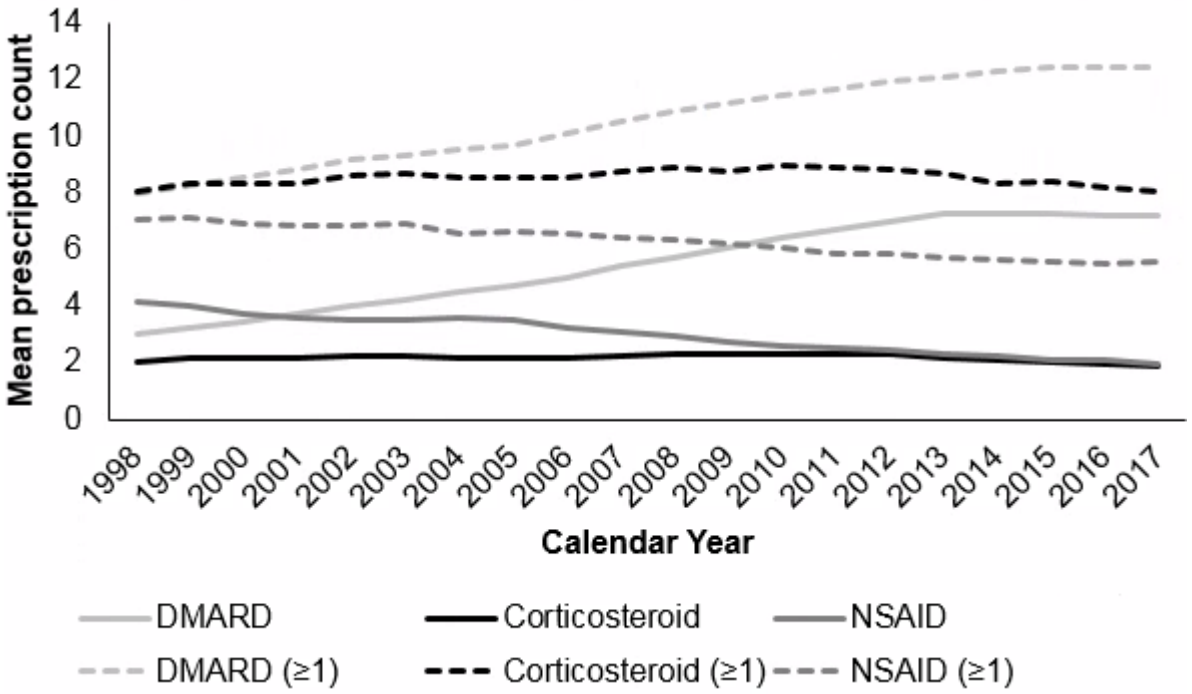
9.3.1.1 Prescription Counts

During follow-up, 59.6% (42,545) of RA patients and 67.0% of matched RA patients had prescribed DMARDs. In 43,870 patients with follow-up in sensitivity analysis RA1 during 1998-2016, 74.6% (32,724) had prescribed DMARDs. This was higher in sensitivity analysis RA2, 94.7% (41,536 of 43,870), as by definition a DMARD post-diagnosis was required.

In RA patients, the mean DMARD prescription count per person-year was 3.00 in 1998 and 7.22 in 2017, peaking at 7.25 in 2013 (Figure 67). In sensitivity analyses RA1 and RA2, this was 4.00 and 5.06 in 1998, rising continually to 7.78 and 9.52 in 2016 (Figure 68). In patients receiving ≥1 DMARD in a given year, the mean prescription count per person-year was 9.97: 7.96 in 1998, 12.12 in 2013 and 12.48 in 2017 (Table 37). In sensitivity analysis RA1 this was 10.5: 8.00 in 1998, 12.14 in 2013 and 12.53 in 2016. The most common number of DMARD prescriptions in a year in these patients was 6, 12 and 13 (Figure 69). In these patients (receiving ≥1 DMARD in a given year) the average annual median in five-year bands was 7.72 for 1998-2002, 9.13 for 2003-2007, 10.65 for 2008-2012 and 10.98 for 2013-2017.

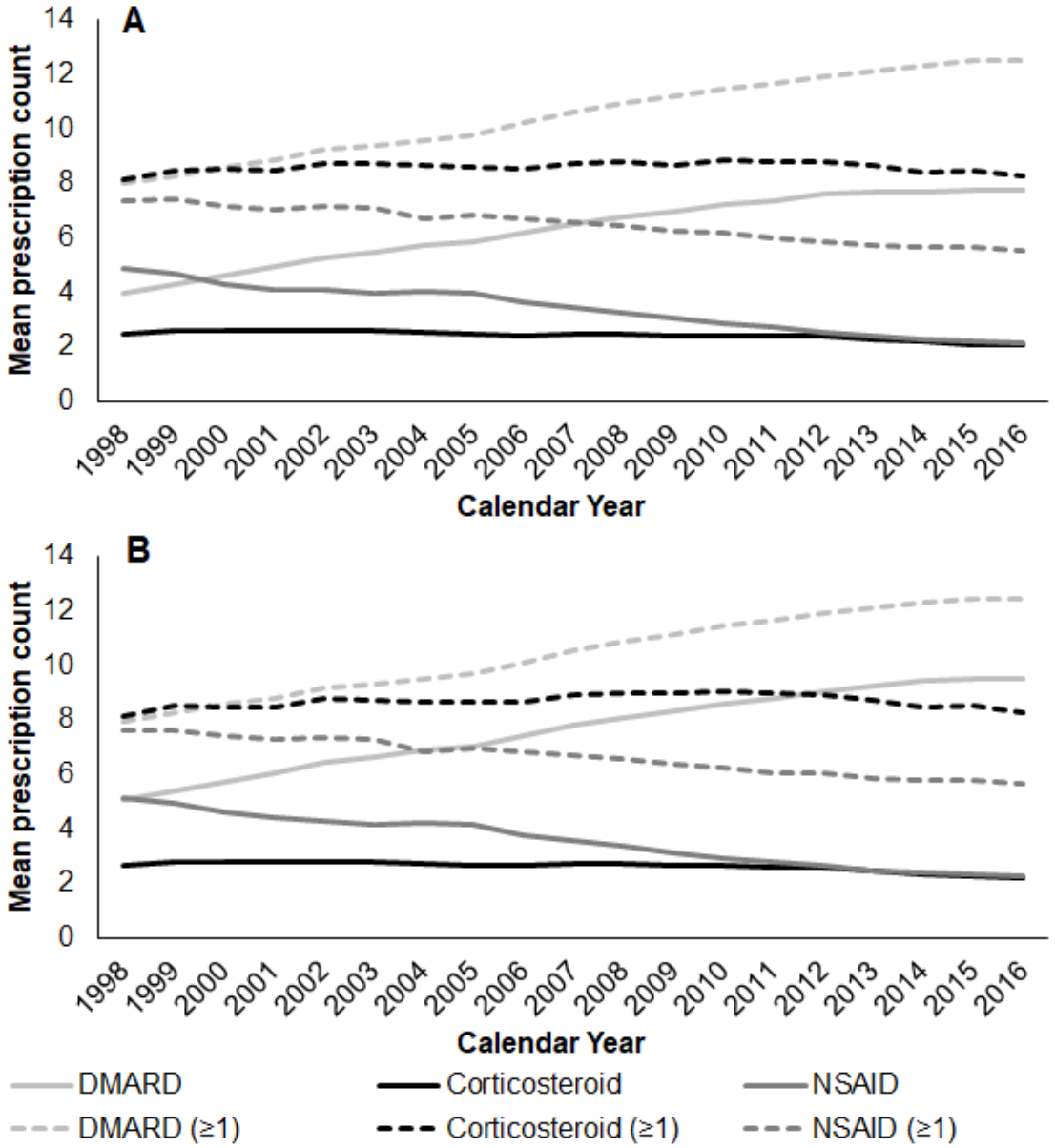
In patients with a full year of GP registration, the annual proportion of patients with ≥6 DMARD prescriptions rose from 24.5% in 1998 to 48.2% in 2013, reaching 48.7% in 2017 (Figure 70). In sensitivity analysis RA1, this rose from 32.4% in 1998 to 51.9% in 2014, reaching 51.9% in 2016 (Figure 69, Figure 71). In sensitivity analysis RA2 this rose from 41.1% in 1998 to 63.2% in 2016.

Figure 67. Annual mean prescription count per person-year, for all RA patients (N = 71,411) and those with ≥1 prescription in a given year (N = 62,306), 1998-2017



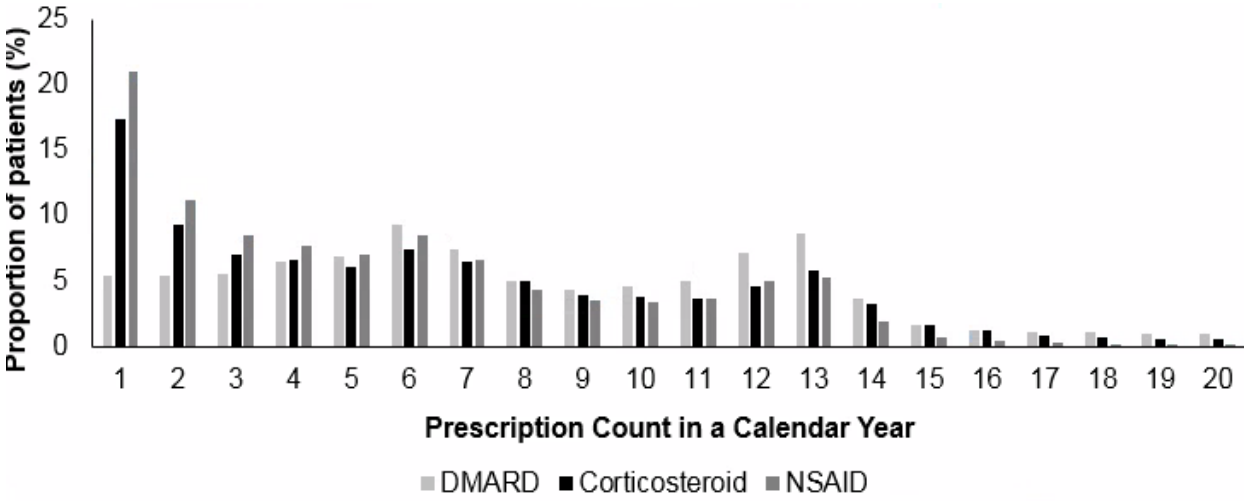
Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Figure 68. Annual mean prescription count per person-year in sensitivity analyses, 1998-2016: A) all RA patients in RA1 (N = 43,870), and those with ≥1 prescription in a given year (N = 41,307); B) all RA patients in RA2 (N = 44,523), and those with ≥1 prescription in a given year (N = 43,597)



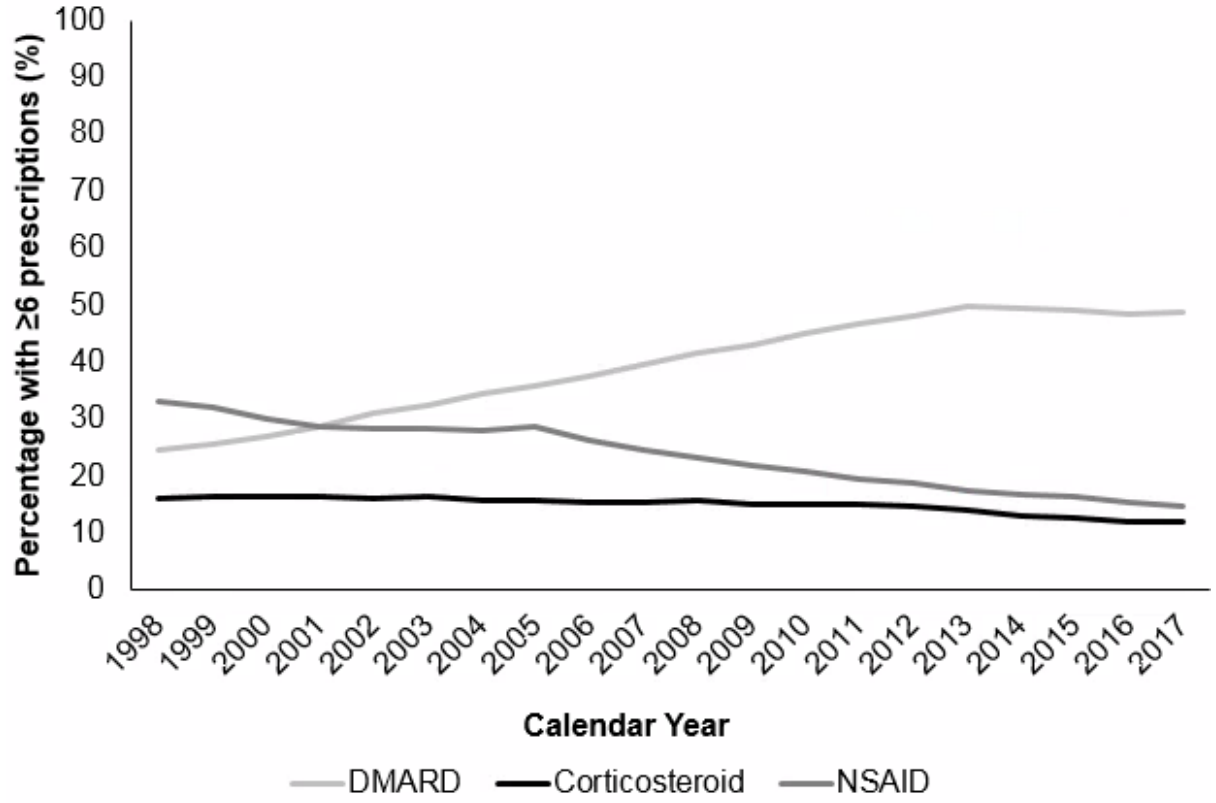
Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug; RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug

Figure 69. Percentage of patients with 1-20 prescriptions in a year across the period 1997-2017, for patients receiving ≥1 prescription in a given year (N = 62,306)



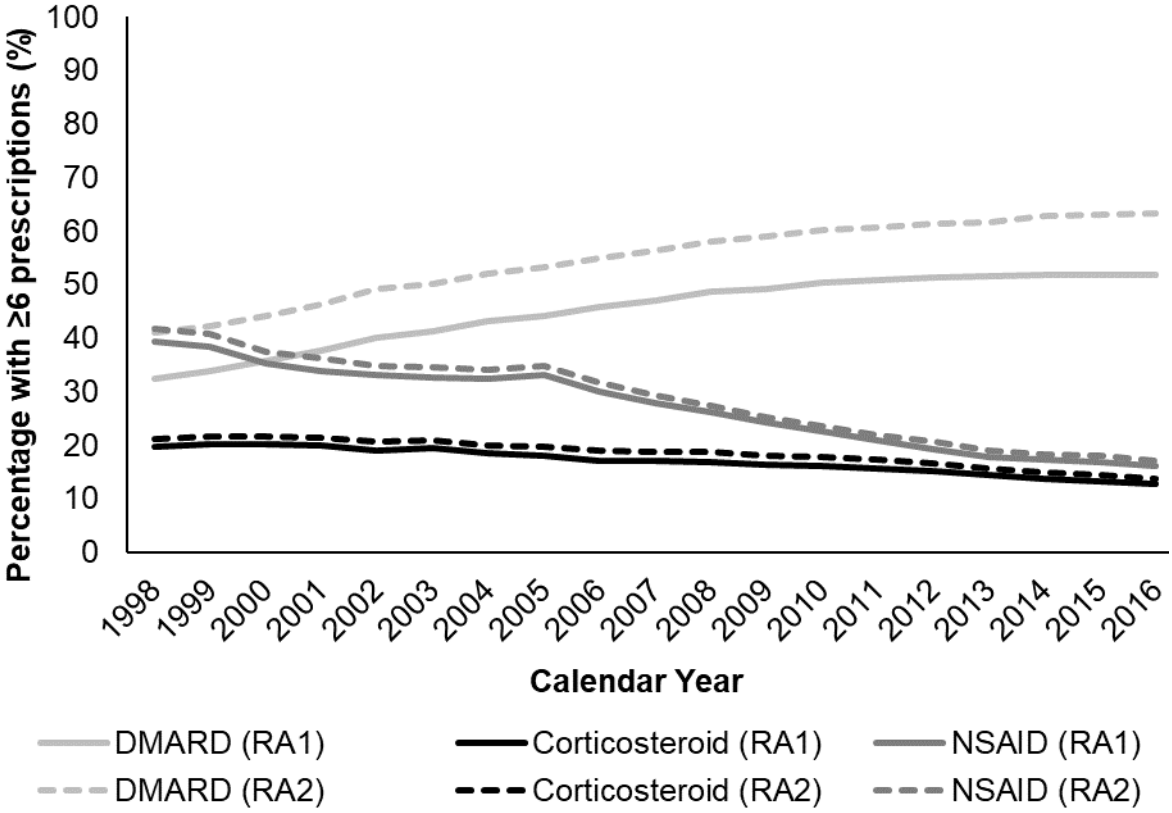
Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Figure 70. Annual percentage of patients with ≥6 prescriptions, 1998-2017 (N = 71,411)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Figure 71. Annual percentage of RA patients with ≥6 annual prescriptions in sensitivity analyses RA1 (N = 43,870) and RA2 (N = 44,523), 1998-2016



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug; RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug

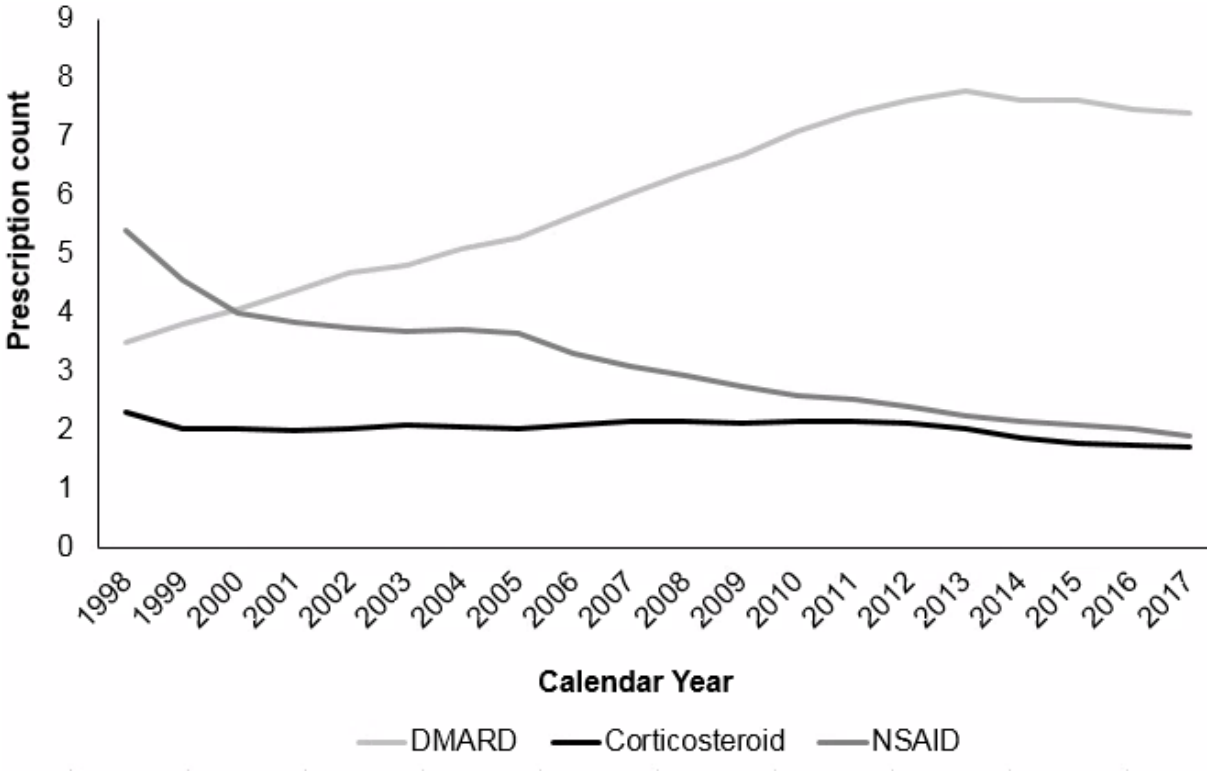
Table 37. Measures of annual prescription utilisation, for patients receiving ≥1 prescription in a given year (N = 62,306), 1997-2017

Prescription Measure	Corticosteroid	DMARD	NSAID
Percentage with 1-12 prescriptions in a year (%)	81.1	72.5	90.3
Greatest number prescribed per patient in a year	169	232	119
Mean number prescribed per patient in a year	7.79	9.97	5.83
Modal number prescribed per patient in a year	1	6	1

Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

During follow-up, 72.9% (30,024) of matched RA patients (having an RA code during the study period) had prescribed DMARDs. The mean count of prescriptions per person-year for matched RA patients rose from 3.51 in 1998 to 7.41 in 2017, having peaked at 7.78 in 2013 (Figure 72). The APC declined from +9.21 in 1998-1999 to -1.91 in 2013-2014, reaching -0.88 in 2016-2017 (Mean APC +4.07).

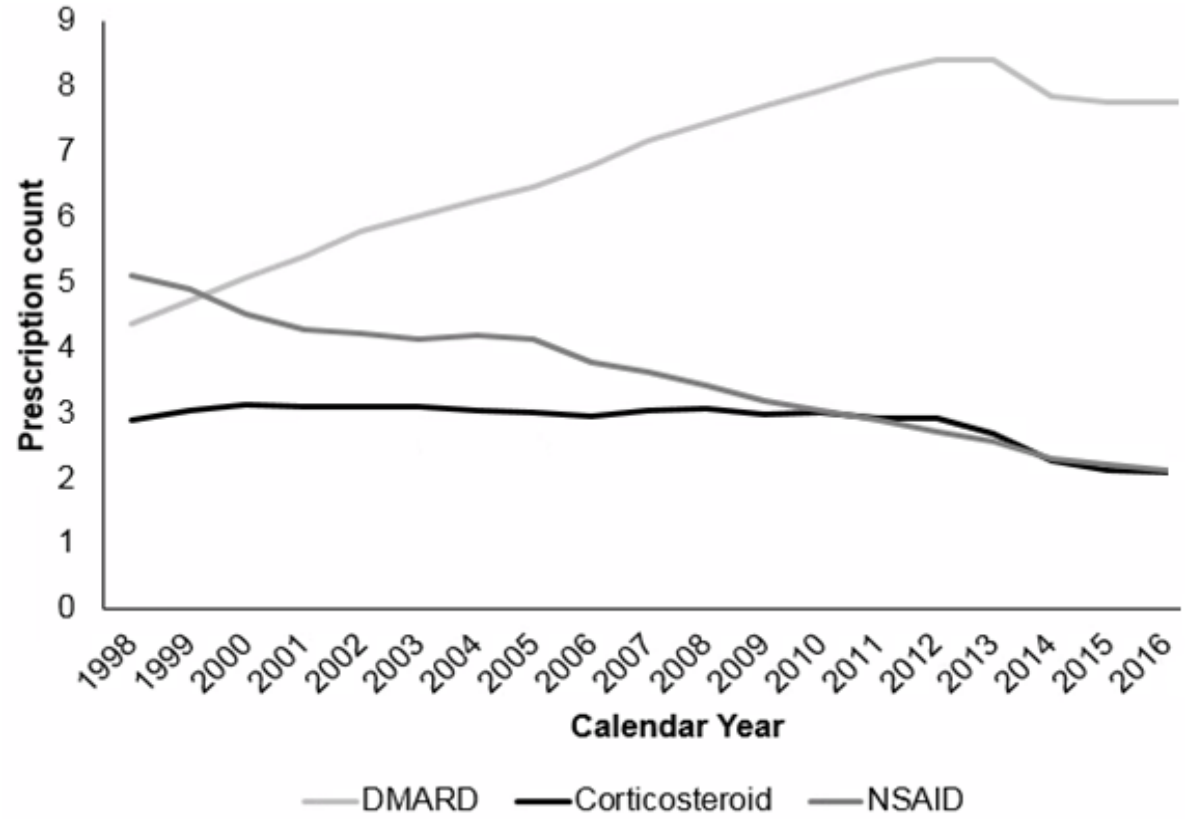
Figure 72. Annual mean prescription count per person-year for matched RA patients in 1998-2017 (N = 41,198)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

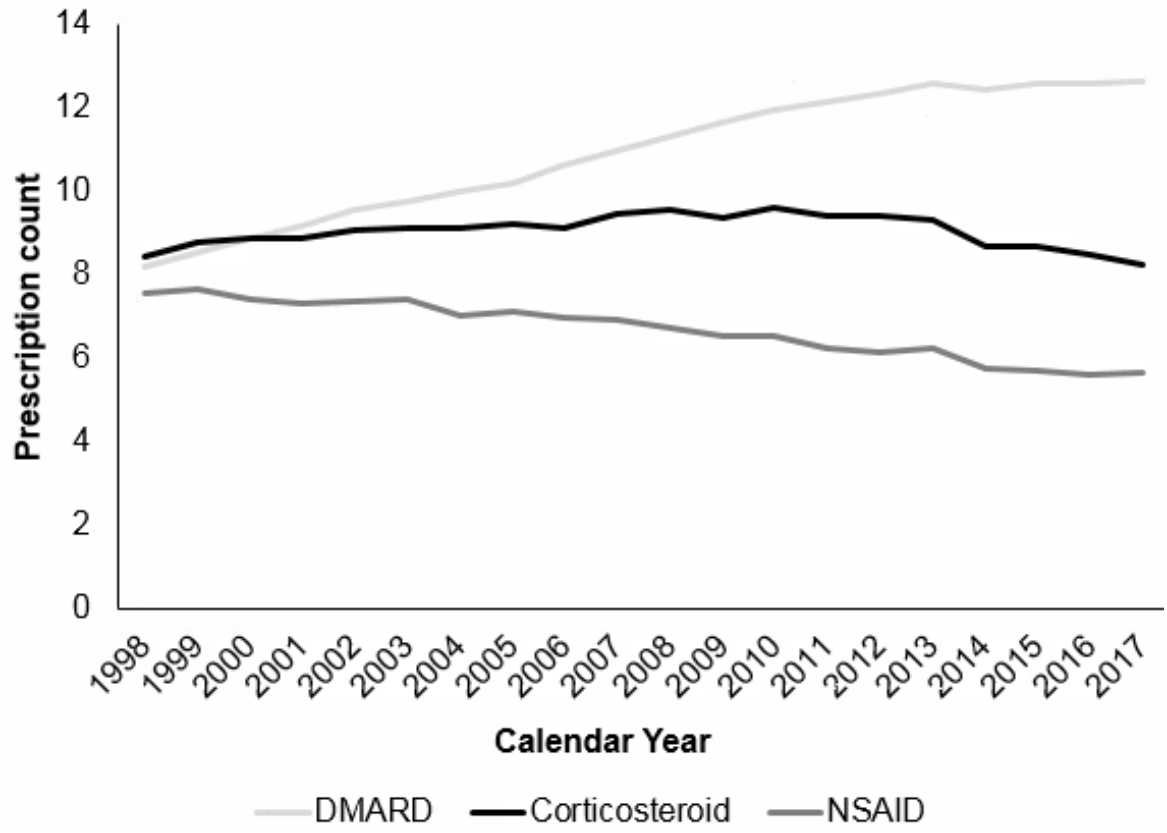
In sub-analysis A of sensitivity analysis RA1, where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis, the mean count of DMARD prescriptions per person-year was higher: 4.37 in 1998 and 7.75 in 2017, with a peak of 8.42 in 2012 (Figure 73). In patients with ≥1 prescription in a given year, in sub-analysis A this was 8.20 in 1998, rising to 12.55 in 2013 and then less steeply to 12.63 in 2017 (Figure 74).

Figure 73. Annual Mean prescription count per person-year in sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), 1998-2016 (N = 44,426)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

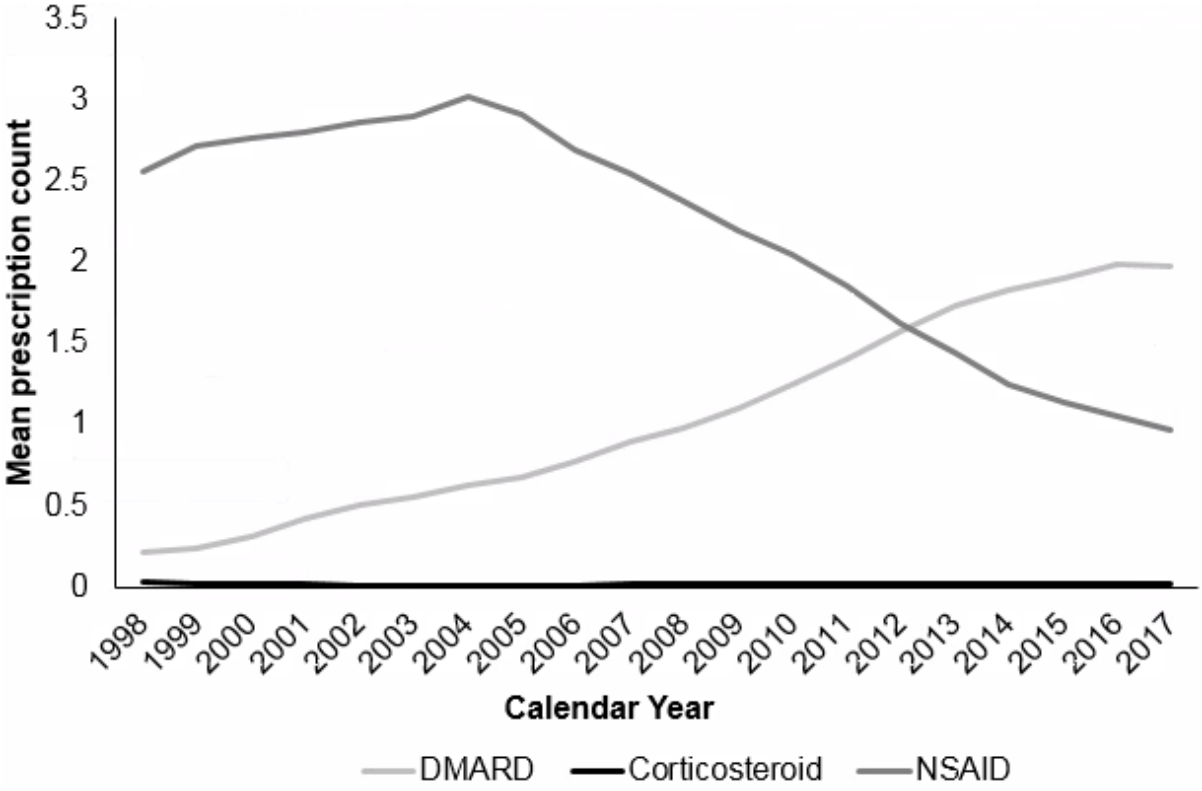
Figure 74. Annual mean prescription count per person-year in RA patients with ≥1 prescription in a given year, in sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), 1998-2017 (N = 39,581)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

In sub-analysis B, with medication selected from specific BNF chapters, the prescription count was lower among RA patients (Figure 75). The DMARD prescription count per person-year rose from 0.21 in 1998 to 1.97 in 2017. Among RA patients with ≥1 prescription in a given year, in sub-analysis B this was 5.63 in 1998 and 8.64 in 2017.

Figure 75. Annual mean prescription count in RA patients, in sub-analysis B with BNF chapter constraints applied, 1998-2017 (N = 71,411)



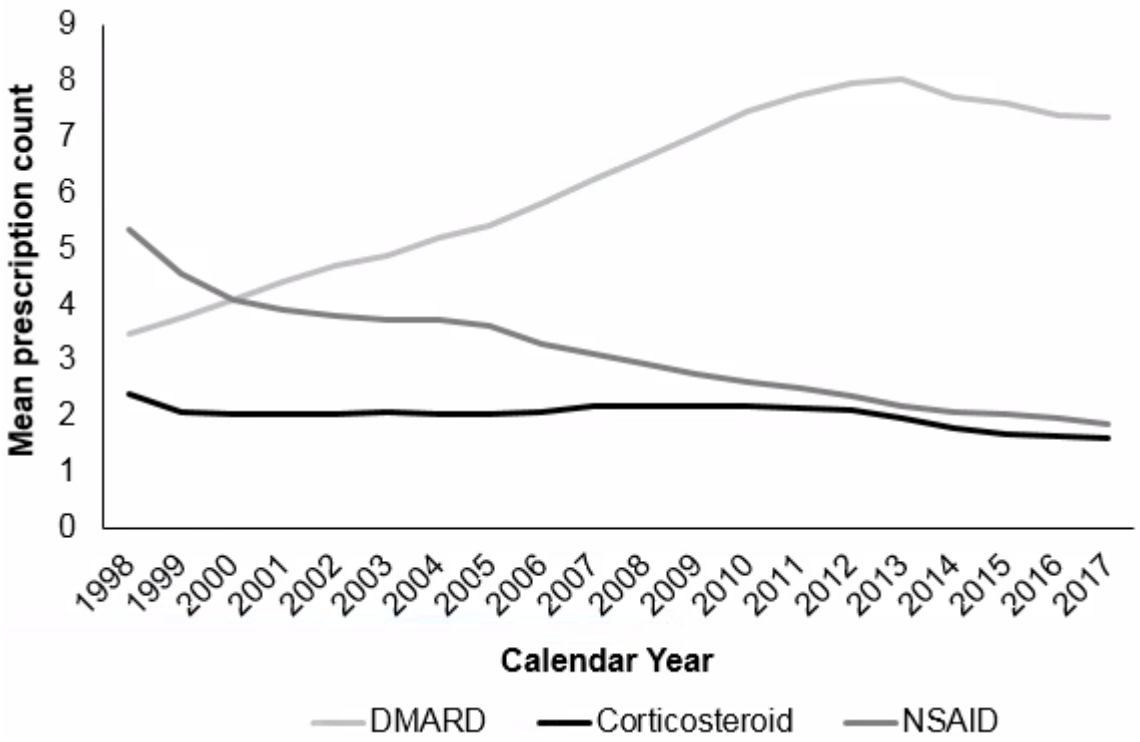
Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

9.3.1.1.1 Incident Cohort

In patients with incident RA during follow-up, 68.7% (21,826) received a DMARD prescription. This was 83.3% (15,541 of 18,657) in sensitivity analysis RA1 and 99.2% (21,131 of 21,295) in sensitivity analysis RA2, between 1998-2016. In this incident cohort, the mean DMARD prescription count per person-year rose from 3.47 in 1998 to 8.05 in 2013 before falling to 7.37 in 2017 (Figure 76). In sensitivity analysis RA1 this rose from 4.69 in 1998 to 8.54 in 2013 before falling to 8.15 in 2016 (Figure 77). In sensitivity analysis RA2 this rose from 5.77 in 1998 to 9.88 in 2014, before falling to 9.77 in 2016. The average annual median prescription count per year in five-year bands was 3.00 for 1998-2002, 4.82 for 2003-2007, 6.63 for 2008-2012 and 6.48 for 2013-2017. In the year post-diagnosis, the mean prescription count per person-year showed a similar trend; it rose from 3.66 in 1998 to 7.42 in 2016, having peaked at 8.50 in 2010 (Figure 78). In sensitivity analysis RA1 this was 5.08 in 1998 and 8.96 in 2016, having peaked at 9.27 in 2011. In sensitivity analysis RA2 this was 6.03 in 1998 and 11.22 in 2016. The proportion with ≥1 prescription in the year post-diagnosis was 48.5% in 1998 and 60.8% in 2016, having peaked at 77.0% in 2009 (Figure 79). In sensitivity analysis RA1 this was 66.1% in 1998 and 78.6% in 2016, having peaked at

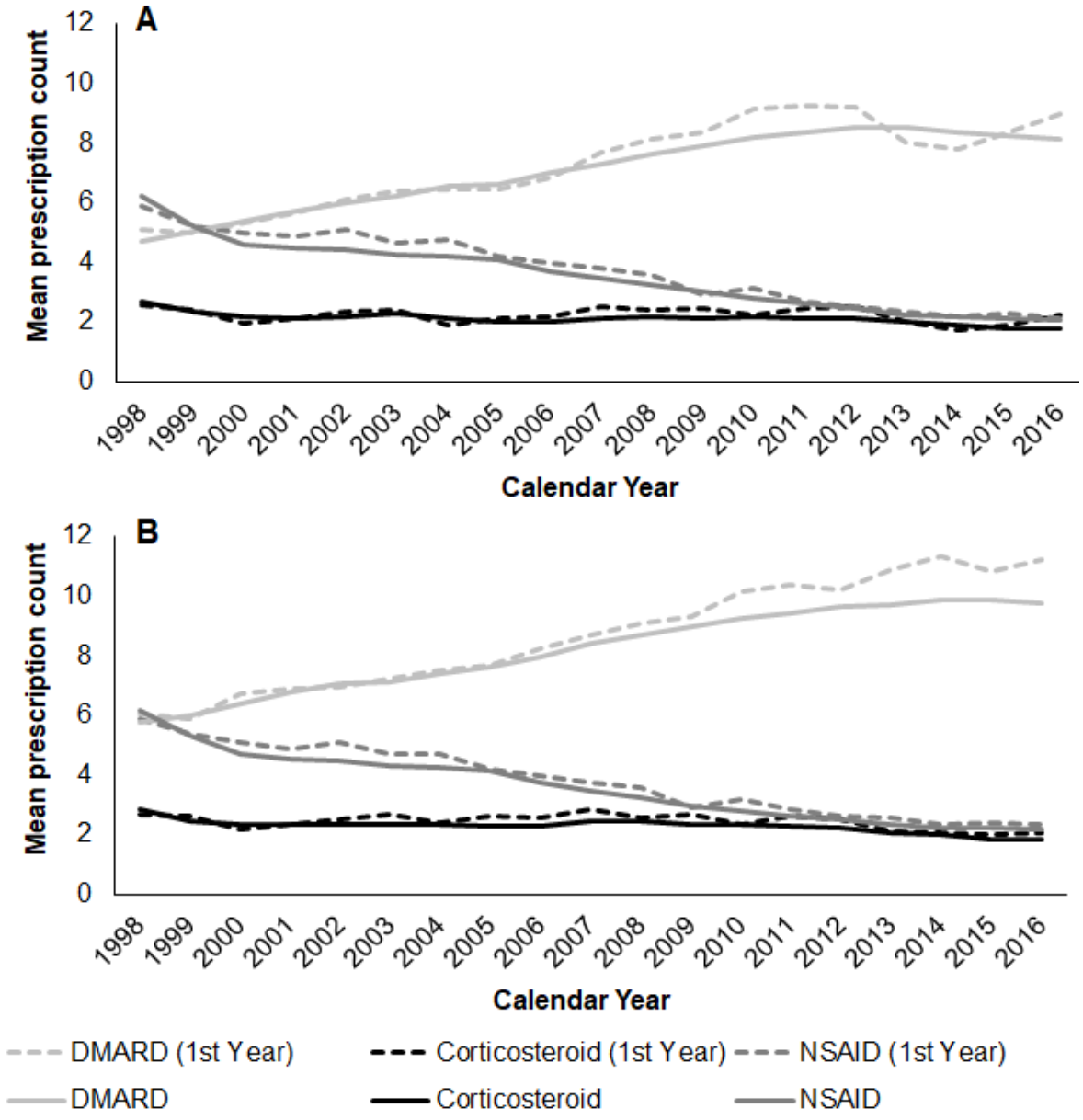
84.0% in 2011 (Figure 80). In sensitivity analysis RA2 this was 80.3% in 1998 and 98.6% in 2016.

Figure 76. Annual mean prescription count per person-year for patients with incident RA, 1998-2017 (N = 31,768)



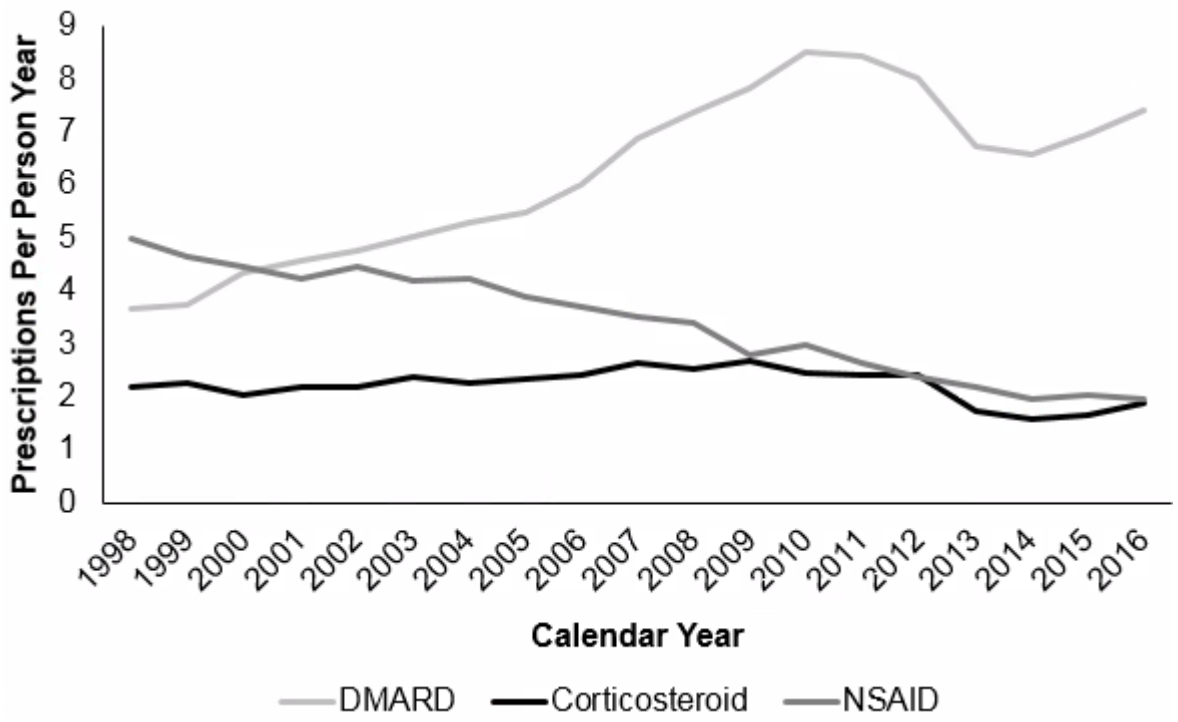
Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Figure 77. Annual mean prescription count per person-year for patients in sensitivity analyses with incident RA, and in the year post-diagnosis, 1998-2016: A) RA1 (N = 18,657); B: RA2 (N = 21,295)



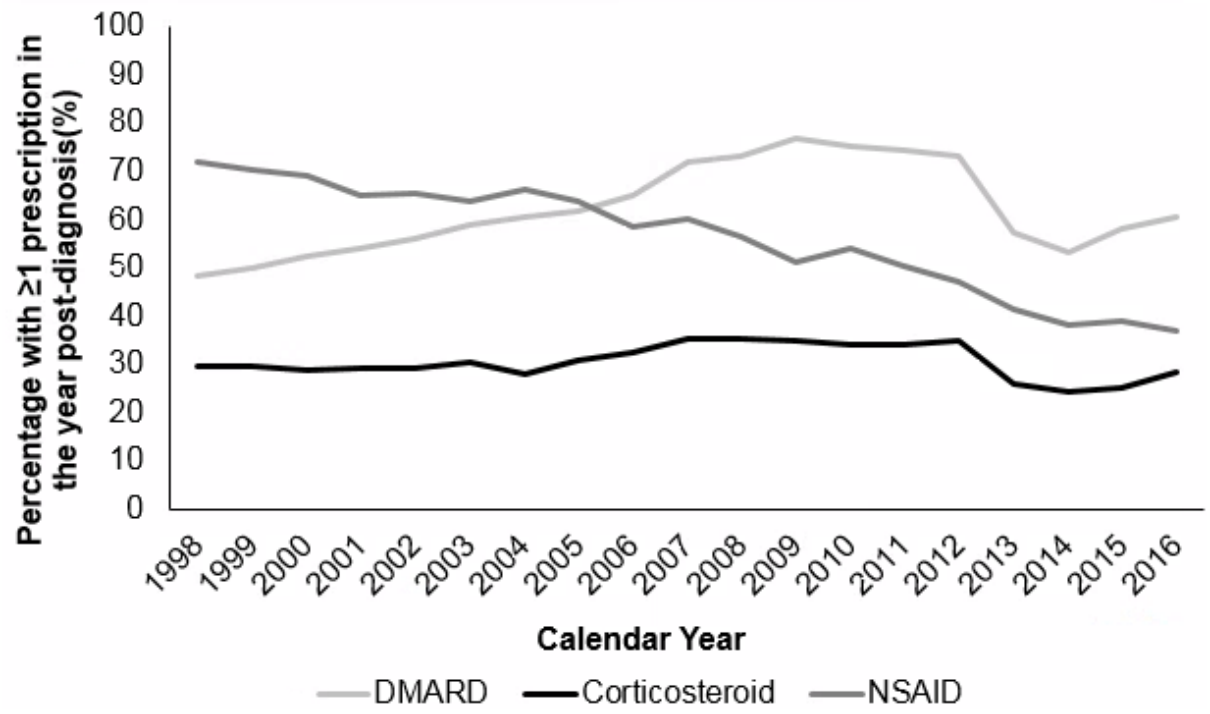
Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug; RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug

Figure 78. Annual mean prescription count per person-year for RA patients in the year post-diagnosis, 1998-2016 (N = 29,918)



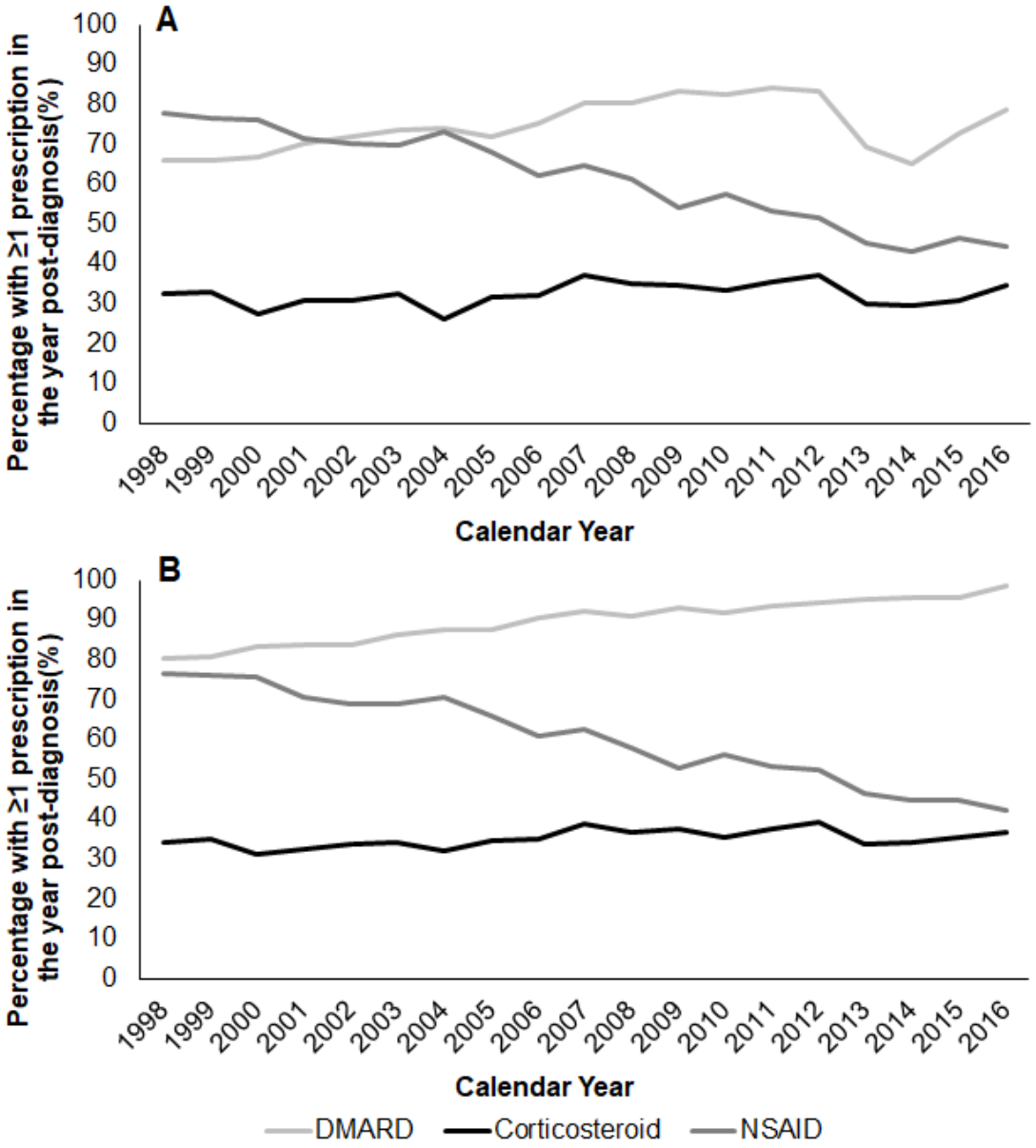
Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Figure 79. Annual percentage of RA patients with ≥1 prescription in the year post-diagnosis, 1998-2016 (N = 29,918)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Figure 80. Annual percentage of RA patients with ≥1 prescription in the year post-diagnosis in sensitivity analyses, 1998-2016: A) RA1 (N = 18,657); B) RA2 (N = 21,295)

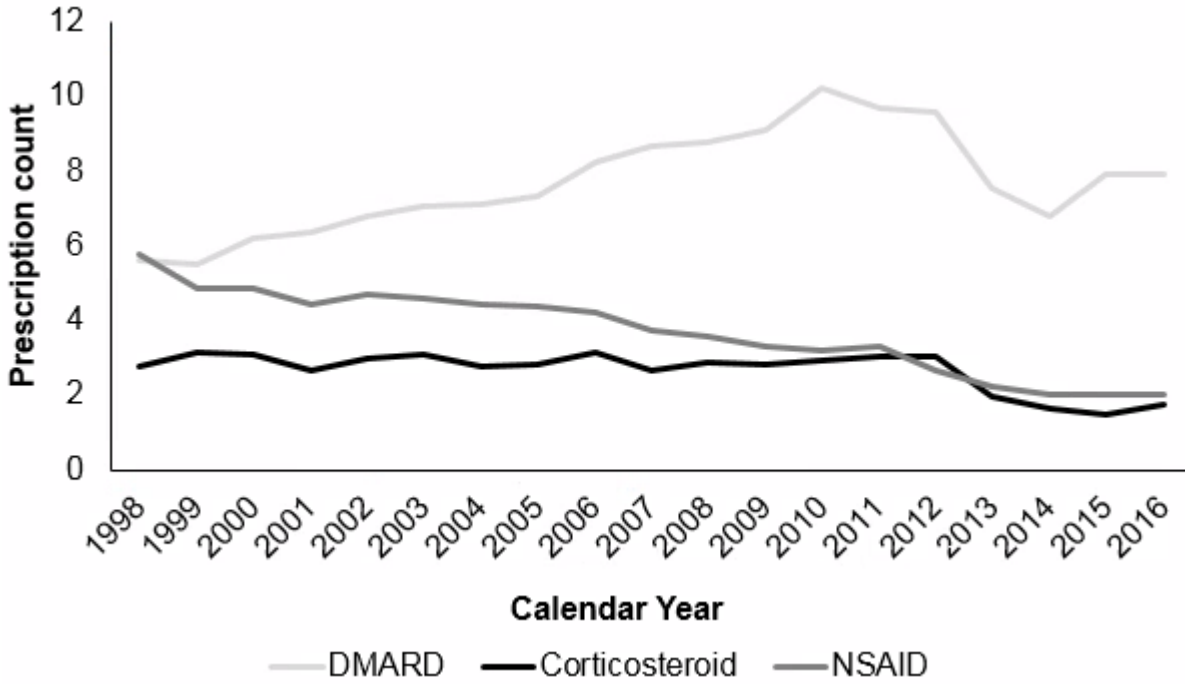


Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug; RA1 = additional code >180 days later; RA2 = subsequent disease modifying anti-rheumatic drug

In patients with incident RA during the study period, in sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), the mean DMARD prescription count rose from 5.80 in 1998 to 8.68 in 2013 before falling to 7.81 in 2014, with modest change thereafter. In the year post-diagnosis, the prescription count rose from 5.60 in 1998 to 9.58 in 2012 before falling to

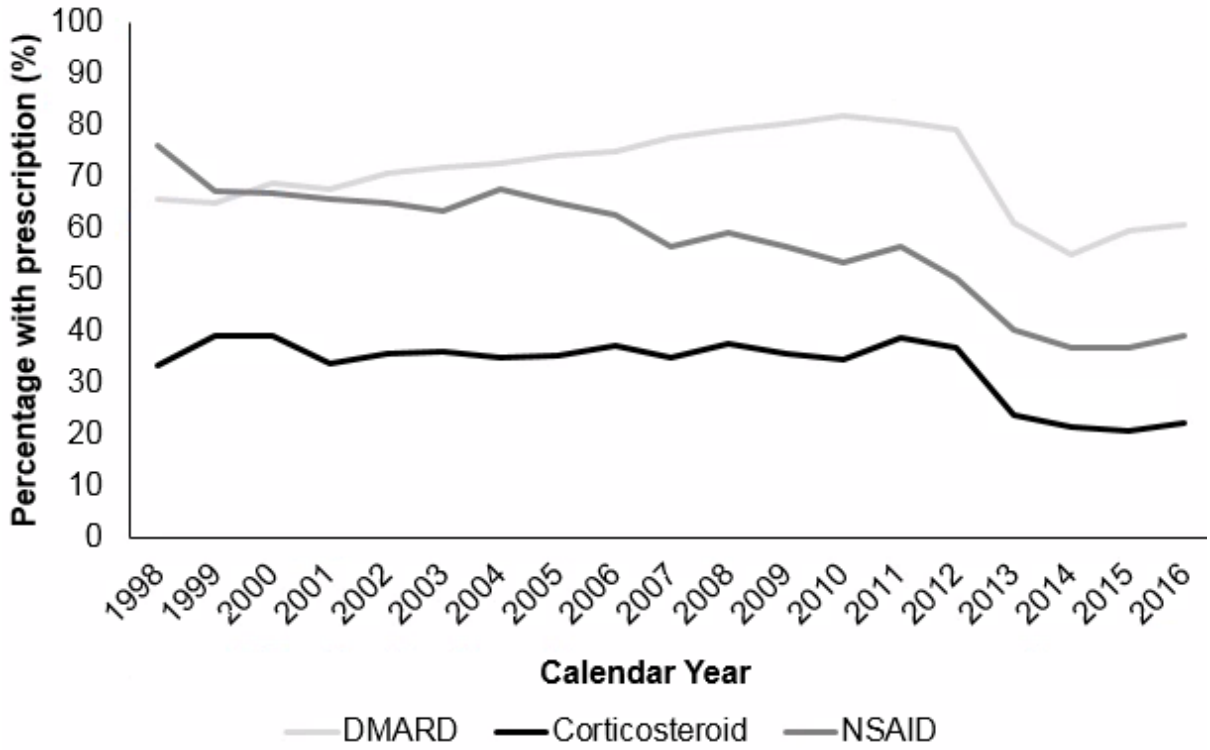
6.78 in 2014 and then rising to 7.94 in 2016 (Figure 81). The proportion of RA patients with a DMARD prescription in the year post-diagnosis was 65.6% in 1998, rising to 79.0% in 2012 before falling to 54.8% in 2014 and then rising to 60.7% in 2016 (Figure 82).

Figure 81. Annual mean prescription count per person-year in the year post-diagnosis, in sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), 1998-2016 (N = 29,403)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

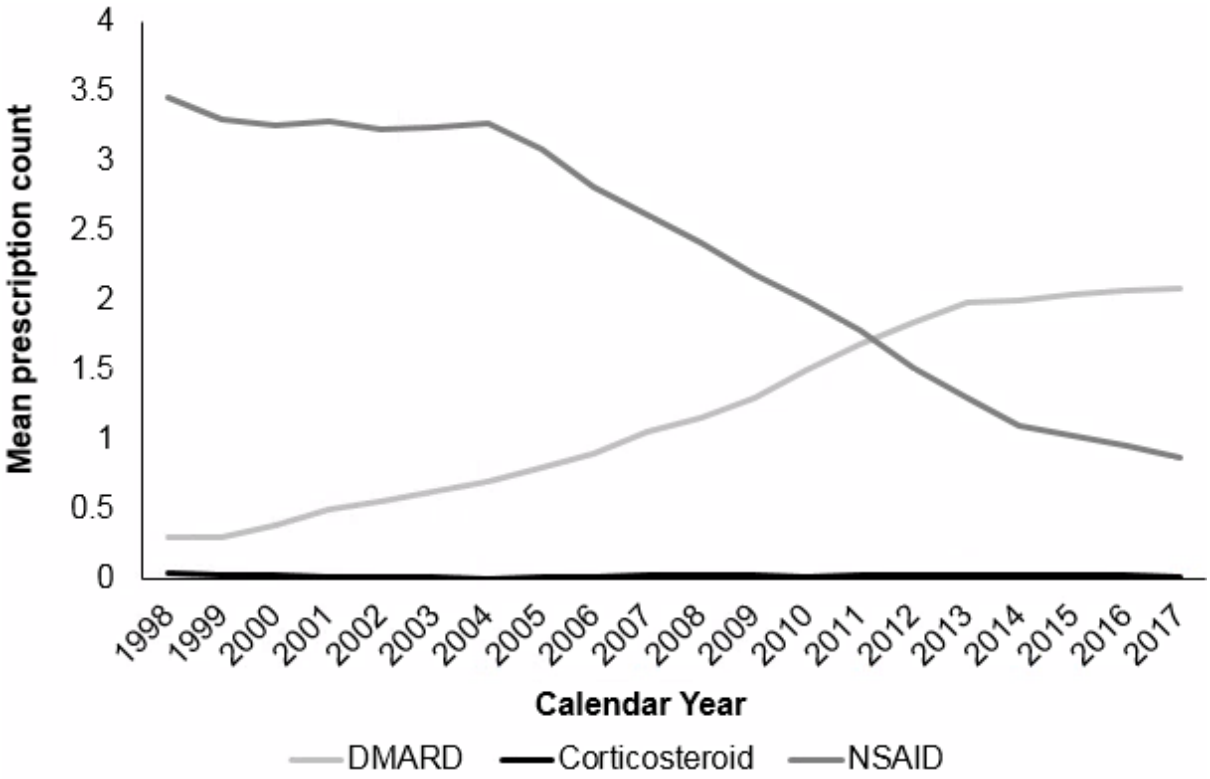
Figure 82. Annual percentage of RA patients with a prescription in the year post-diagnosis, in sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), 1998-2016 (N = 29,403)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

In the incident RA cohort, in sub-analysis B with medication selected from specific BNF chapters, the DMARD prescription count per person-year was only 0.30 in 1998 and 2.09 in 2017 (Figure 83). In the year post-diagnosis, in sub-analysis B this was 0.22 in 1998, rising to 1.36 in 2012 before falling to 1.11 in 2013, then returning to 1.40 in 2016.

Figure 83. Annual mean prescription count for incident RA patients in sub-analysis B with medication selected from specific BNF chapters, in 1998-2017 (N = 71,411)

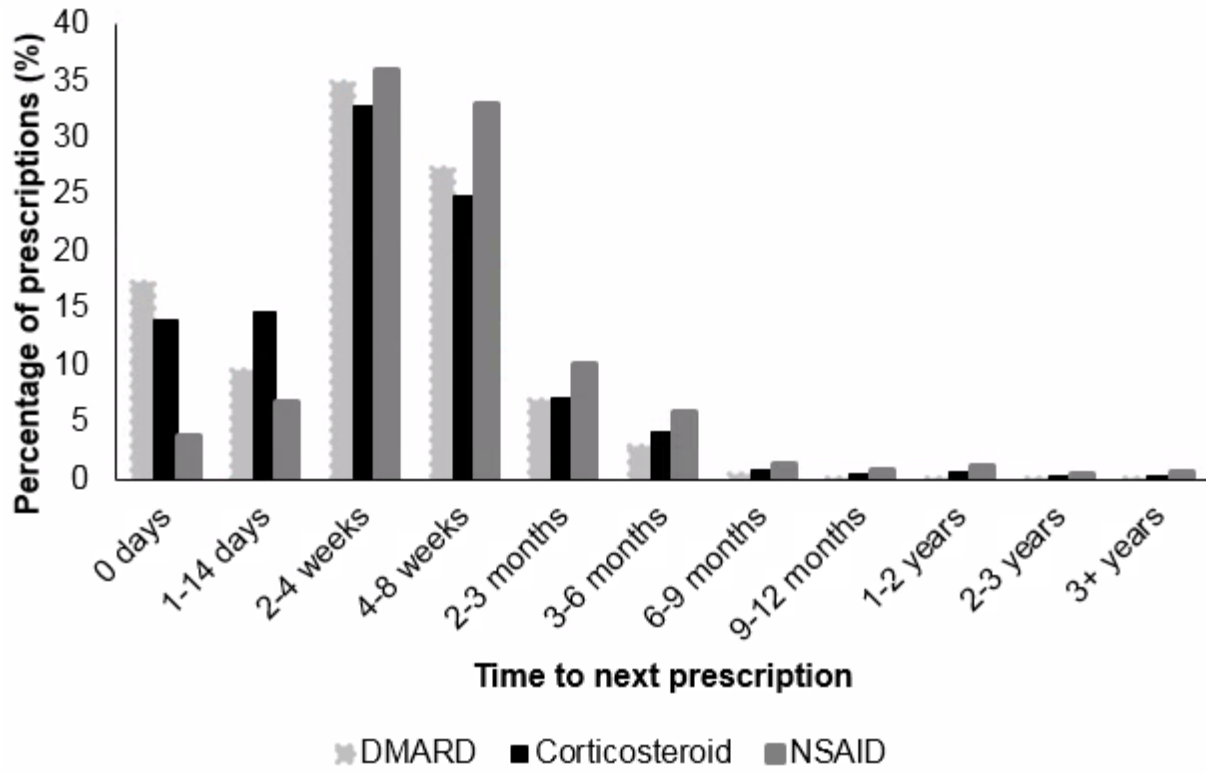


Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

9.3.1.2 Prescribing Duration

During follow-up, 59.6% of RA patients received DMARD prescribing; 55.6% received ≥90 days in a year (long-term DMARD prescribing) and 49.3% received ≥180 days in a year. Across 1998-2016, in sensitivity analysis RA1 this was 74.7%, 70.1% and 62.4% respectively; in sensitivity analysis RA2 this was 93.6%, 86.1% and 75.5%. Among 2,569,898 DMARD prescriptions during follow-up, the modal number of days between prescriptions was 28 days, with the gap between prescriptions being ≥2 months in only 10.7% of cases (Figure 84).

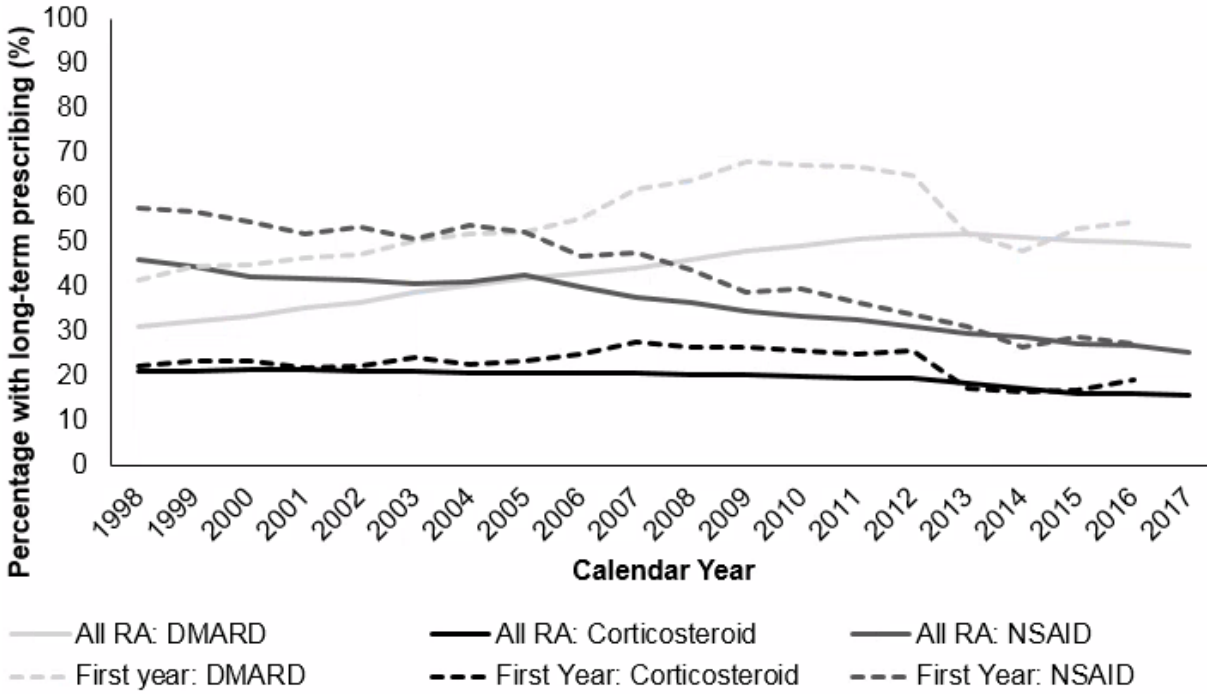
Figure 84. The proportion of prescriptions followed by a gap of 0 days to ≥3 years before the next prescription, in RA patients (N = 62,306)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

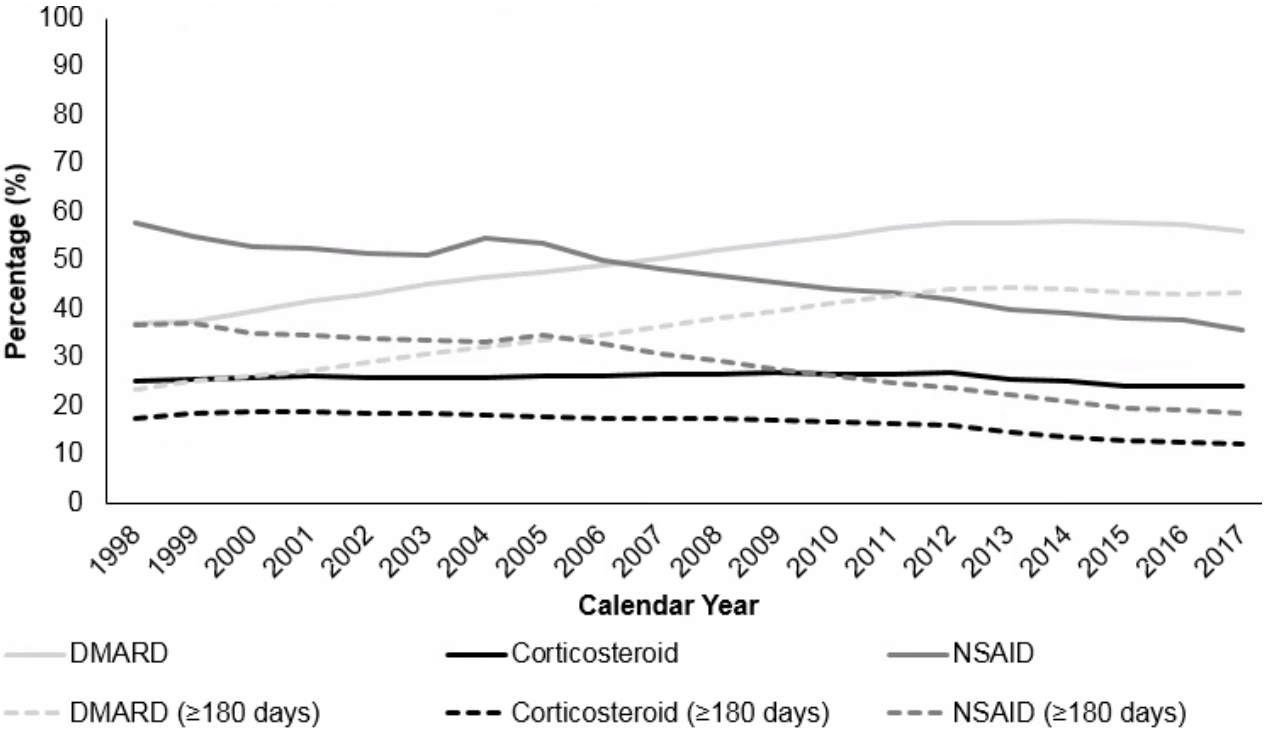
The proportion of RA patients with any, ≥90 days and ≥180 days of DMARD prescribing in a year rose from 37.0%, 31.0% and 23.5% in 1998 to 57.8%, 52.0% and 44.7% in 2013 before it fell to 56.3%, 49.3% and 43.6% in 2017 (Figure 85, Figure 86, Table 38). In sensitivity analysis RA1, this was 49.4%, 41.1% and 31.2% in 1998; 62.9%, 56.0% and 47.5% in 2013; and 61.4%, 53.6% and 46.0% in 2016. In sensitivity analysis RA2, this was 62.8%, 52.2% and 39.5% in 1998; 76.1%, 67.0% and 56.7% in 2013; and 76.4%, 66.3% and 56.7% in 2016. The mean APC in the proportion of patients with long-term prescribing for five-year bands was +4.19 in 1998-2002, +3.92 in 2003-2007, +3.15 in 2008-2012 and -0.92 in 2013-2017.

Figure 85. Annual percentage with ≥90 days prescribing: all RA patients, 1998-2017 (N = 68,939) and in the year post-diagnosis, 1998-2016 (N = 29,918)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Figure 86. Annual percentage of RA patients with ≥1 prescription (N = 71,411) and with ≥180 days of prescribing, 1998-2017 (N = 66,147)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Table 38. Annual proportion with ≥90 days prescribing: all RA patients (1998-2017) and in the year post-diagnosis (1998-2016)

Year	Proportion (%) with long-term medication prescribing (N = 68,939)			Proportion (%) with long-term medication prescribing in the first year post-diagnosis (N = 29,918)		
	DMARD	Corticosteroid	NSAID	DMARD	Corticosteroid	NSAID
1998	31	21	45.9	41.6	22.2	57.7
1999	32.2	20.9	44.5	44.4	23.2	56.7
2000	33.3	21.5	42.2	44.9	23.3	54.4
2001	35.3	21.6	42	46.4	21.7	51.7
2002	36.5	21.2	41.4	47.4	22.2	53.4
2003	38.6	21	40.7	50.3	24.1	50.6
2004	40.2	20.8	40.9	51.8	22.7	53.7
2005	41.7	20.8	42.7	52.3	23.3	52.1
2006	42.9	20.6	40	55.3	24.7	46.8
2007	44.2	20.5	37.8	62	27.5	47.6
2008	46.2	20.4	36.4	63.9	26.3	43.6
2009	47.8	20.2	34.7	67.9	26.3	38.6
2010	49.1	19.9	33.2	67.1	25.8	39.5
2011	50.9	19.6	32.6	67.1	24.8	36.6
2012	51.6	19.4	31.1	65.1	25.5	33.8
2013	52	18.2	29.6	51.9	17.3	31.2
2014	51.3	17.1	28.5	47.8	16.4	26.6
2015	50.3	16.2	27.3	53.2	16.7	28.9
2016	50.1	16.1	27	54.7	19.1	27.1
2017	49.3	15.5	25.1			

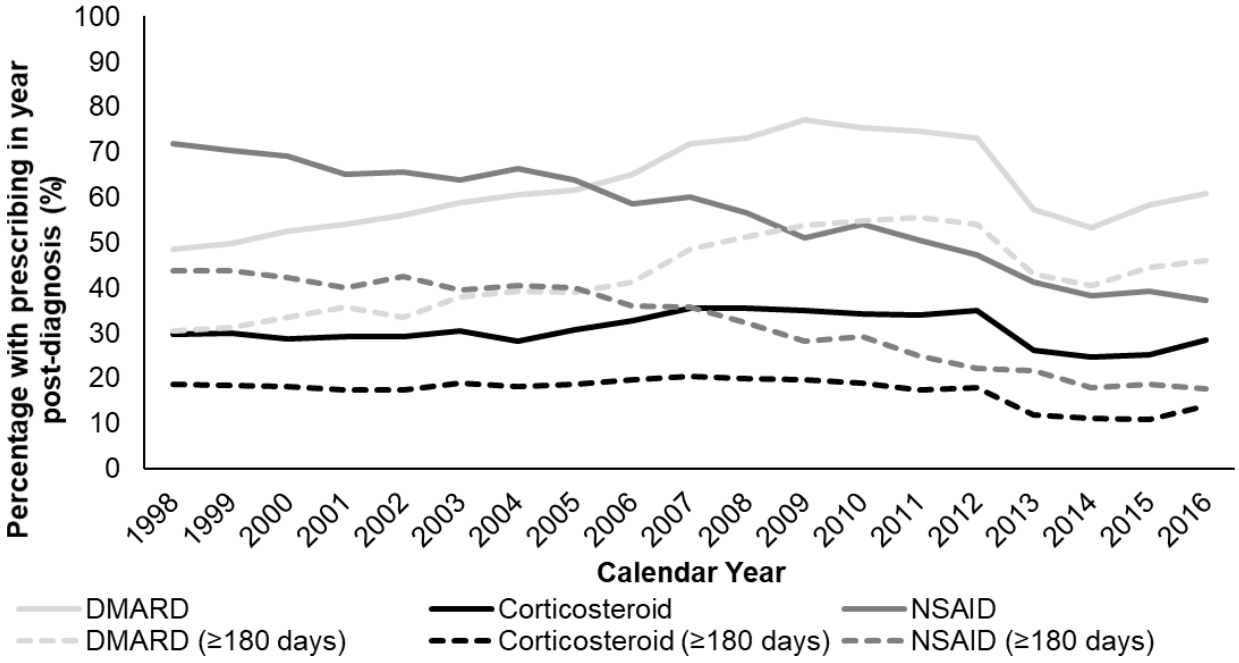
Note: NSAID = Non-steroidal anti-inflammatory; DMARD = Disease-modifying anti-rheumatic drug

9.3.1.2.1 Incident Cohort

For incident RA patients, the proportion with any, ≥90 days and ≥180 days of DMARD prescribing in the year post-diagnosis was 62.5%, 54.6% and 43.3% respectively. Between 1998 and 2016, this was 75.1%, 66.2% and 52.3% in sensitivity analysis RA1 and 90.7%, 78.2% and 61.5% in sensitivity analysis RA2.

The proportion with any, ≥90 days and ≥180 days of DMARD prescribing in the year post-diagnosis rose from 48.5%, 41.6% and 30.3% in 1998 to 77.0%, 67.9% and 53.9% in 2009 before plateauing, reaching 73.0%, 65.1% and 54.1% in 2012 (Figure 85, Figure 87). The proportions subsequently fell to 53.4%, 47.8% and 40.5% in 2014 before showing a return trend, reaching 60.8%, 54.7% and 45.9% by 2016. The pattern was similar in sensitivity analysis RA1: 66.1%, 56.4% and 42.1% in 1998; 83.1%, 73.7% and 58.9% in 2009; 69.4%, 62.6% and 51.3% in 2013; and 78.6%, 69.3% and 57.8% in 2016. In sensitivity analysis RA2, it rose steadily from 80.3%, 68.7% and 50.0% in 1998 to 98.6%, 84.7% and 70.1% in 2016.

Figure 87. Annual percentage of RA patients in the year post-diagnosis with ≥1 prescription (N = 30,742) or ≥180 days prescribing, 1998-2016 (N = 29,164)



Note: NSAID = Non-steroidal anti-inflammatory; DMARD = Disease-modifying anti-rheumatic drug

9.3.2 Trends in Oral Corticosteroid Prescribing

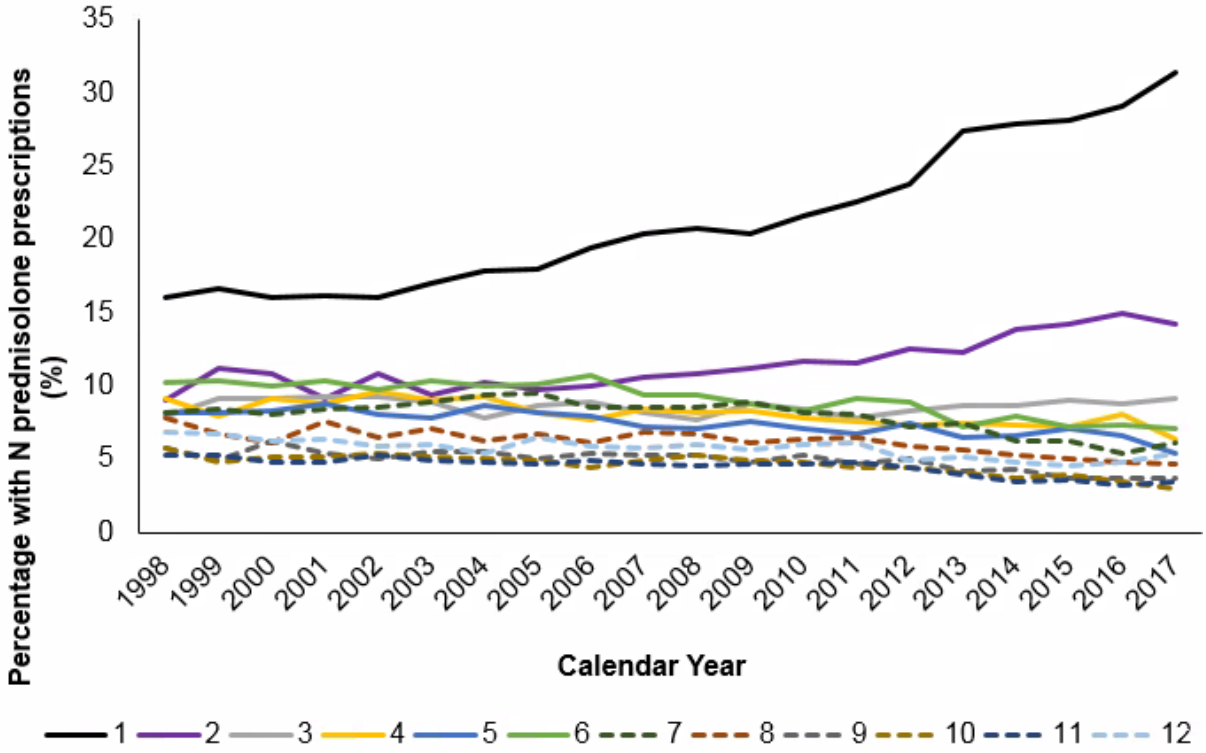
9.3.2.1 Prescription Counts

During follow-up, 32,220 (45.1%) RA patients had prescribed corticosteroids. This was 23,392 (53.3%) patients in sensitivity analysis RA1 and 24,416 (58.8%) patients in sensitivity analysis RA2. The mean corticosteroid prescription count per person-year was 2.04 in 1998 and 1.89 in 2017 (Figure 67). In sensitivity analyses RA1 and RA2

this was 2.50 and 2.84 in 1998 and 2.07 and 1.83 in 2016, with a mean APC of -1.01 and -1.05 (Figure 68). Excluding DMARD-naïve patients (i.e. excluding follow-up prior to a first DMARD prescription), the mean corticosteroid prescription count per person-year was higher: 3.08 in 1998 and 2.15 in 2017 (N = 42,545).

In patients receiving ≥1 corticosteroid in a given year, the mean prescription count per person-year was 8.54 and changed little: 8.03 in 1998, peaking at 8.89 in 2008, and 8.02 in 2017 (Table 37). In sensitivity analyses RA1 and RA2 this was 8.60 and 8.69: 8.17 and 8.12 in 1998 and 8.31 and 8.30 in 2016. Where RA patients received corticosteroids, the median number of prescriptions per year was 6 between 1998 and 2012 and then 5 from 2013 to 2017. In addition, the most common number of prescriptions in a year was 1, 2 and 6 (Figure 69). However, among patients receiving 1-12 prescriptions of prednisolone in a year, over time an increasing proportion received only 1-2 prednisolone prescriptions (16.00% having 1 in 1998 compared with 31.43% in 2017) (Figure 88).

Figure 88. Annual percentage of RA patients with 1-12 oral prednisolone prescriptions issued in a year, 1998-2017 (N = 30,948)



The annual proportion of RA patients with ≥6 corticosteroid prescriptions and a full year of GP registration in a given year was 16.1% in 1998 and initially stable, reaching 15.7% in 2008 (mean APC -0.29) before showing steady decline to 12.0% in 2017

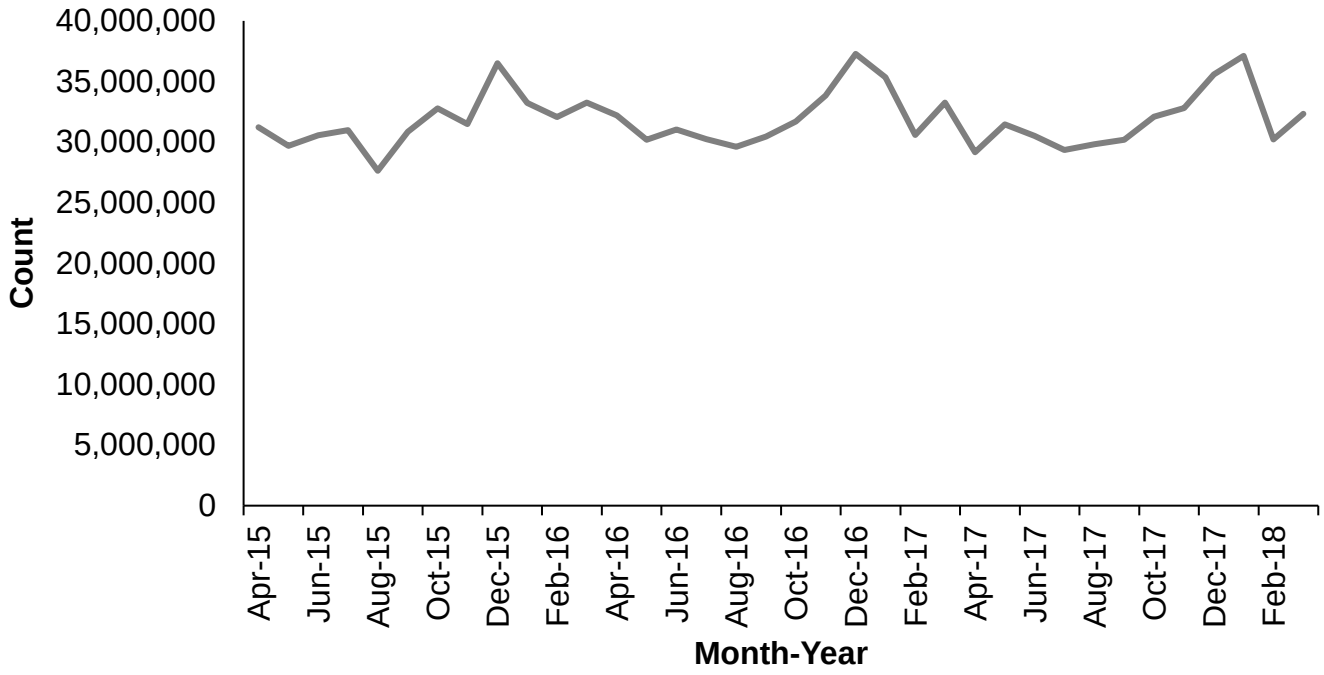
(overall mean APC -1.54) (Figure 70). In sensitivity analyses RA1 and RA2 this was 19.7% and 21.2% in 1998 and 12.8% and 13.8% in 2016 (Figure 71).

During follow-up, 19,880 (43.5%) matched RA patients (having RA coded in the study period) had prescribed corticosteroids. The mean count of prescriptions per person-year for matched RA patients fell from 2.31 in 1998 to 1.74 in 2017 (Figure 72). The APC was -11.33 in 1998-1999 before showing little change until 2013-2014 (APC -7.97); returning to -0.40 in 2016-2017. Excluding DMARD-naïve patients, the mean count of prescriptions per person-year was higher: 3.16 in 1998 and 1.94 in 2017 (N = 29,963).

In sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), the mean count of corticosteroid prescriptions per person-year was 2.90 in 1998 and stable until 2013 before declining to 2.05 in 2017 (Figure 73). In patients with ≥1 prescription in a given year, in sub-analysis A this was 8.41 in 1998 and 9.31 in 2013 (mean APC 1998-2013: +0.70) before declining to 8.68 in 2013 and then less steeply to 8.24 in 2017 (Figure 74). Excluding DMARD-naïve patients, the mean count of corticosteroid prescriptions per person-year was higher: 3.08 in 1998 and 2.20 in 2017 (N = 33,084).

Prednisolone prescribing across England was stable from 2015-2018; 31,226,876 prescriptions in April 2015 and 32,333,377 in March 2018 (Figure 89).

Figure 89. Monthly count of prednisolone prescriptions made across England, 2015-2018

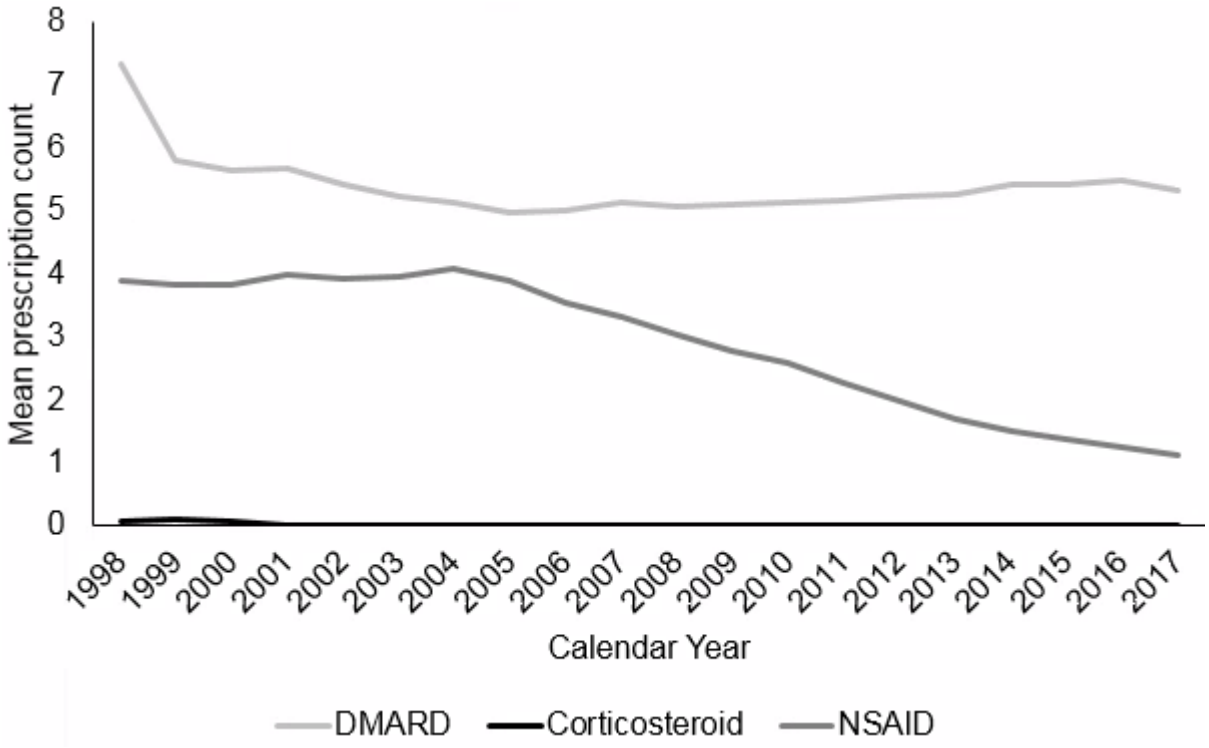


In patients with incident RA during follow-up, 45.6% (14,498) received a corticosteroid prescription during the study period. This was 52.8% (9,935 of 18,809) in sensitivity analysis RA1 and 52.1% (11,406 of 21,880) in sensitivity analysis RA2. In this incident cohort the mean corticosteroid prescription count per person-year was 2.39 in 1998, falling to 2.07 in 1999 before slowly rising to 2.12 in 2012 before falling to 1.62 in 2017 (Figure 76). The mean prescription count per person-year across five-year bands was 2.11 in 1998-2002, 2.08 in 2003-2007, 2.17 in 2008-2012 and 1.74 in 2012-2017. In sensitivity analyses RA1 and RA2 this was 2.30 and 2.47 in 1998-2002, 2.12 and 2.34 in 2003-2007, 2.15 and 2.34 in 2008-2012 and 1.88 and 1.94 in 2013-2016 (four-year band) (Figure 77). Excluding DMARD-naïve patients, the mean prescription count per person-year was 2.59 in 1998 and 2.54 in 2012 before declining to 2.00 in 2017.

In the year post-diagnosis, the mean prescription count per person-year was 2.18 in 1998, rising to 2.40 in 2012 with a peak at 2.66 in 2009. This fell to 1.58 in 2014 before rising to 1.89 in 2016. In sensitivity analyses RA1 and RA2 this was 2.55 and 2.68 in 1998 and 2.25 and 2.09 in 2016, having fallen to 1.75 and 2.06 in 2014. The proportion with ≥1 prescription in the year post-diagnosis was stable; 29.6% in 1998 and 28.4% in 2016 (Figure 79). In sensitivity analyses RA1 and RA2 this was 32.4% and 34.1% in 1998 and 34.5% and 36.7% in 2016 (Figure 80).

Prescription counts were lower in sub-analysis B, with medication selected from specific BNF chapters (Figure 75). The corticosteroid prescription count per person-year fell from 0.05 in 1998 to 0.03 in 2017. Among RA patients with ≥1 prescription in a given year, this was 3.34 in 1998 and 1.50 in 2017. Excluding DMARD-naïve patients, in sub-analysis B, the prescription count per person-year in the RA cohort was 0.08 in 1998 and fell more steeply to 0.01 in 2017 (Figure 90). With the corticosteroid terms listed by Black et al. (486), 41.7% of RA patients received a corticosteroid between 1992 and 2009; this was 3.8% when the BNF chapter constraint was additionally applied (n = 50,155).

Figure 90. Annual mean prescription count for RA patients following their first DMARD prescription, in sub-analysis B with BNF chapter constraints applied, 1998-2017 (N = 71,411)



Note: NSAID = Non-steroidal anti-inflammatory; DMARD = Disease-modifying anti-rheumatic drug

9.3.2.1.1 Incident Cohort

In the incident RA cohort, in sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), the mean corticosteroid prescription count was stable from 1998 (2.68) to 2012 (2.84) before declining to 1.87 in 2016. There was a similar trend in the year post-diagnosis; 2.78 in 1998, 3.03 in 2012 and 1.76 in 2016 (Figure 81). The proportion of RA patients with a corticosteroid

prescription in the year post-diagnosis was 33.4% in 1998 and 36.9% in 2012 and 22.4% in 2016 (Figure 82).

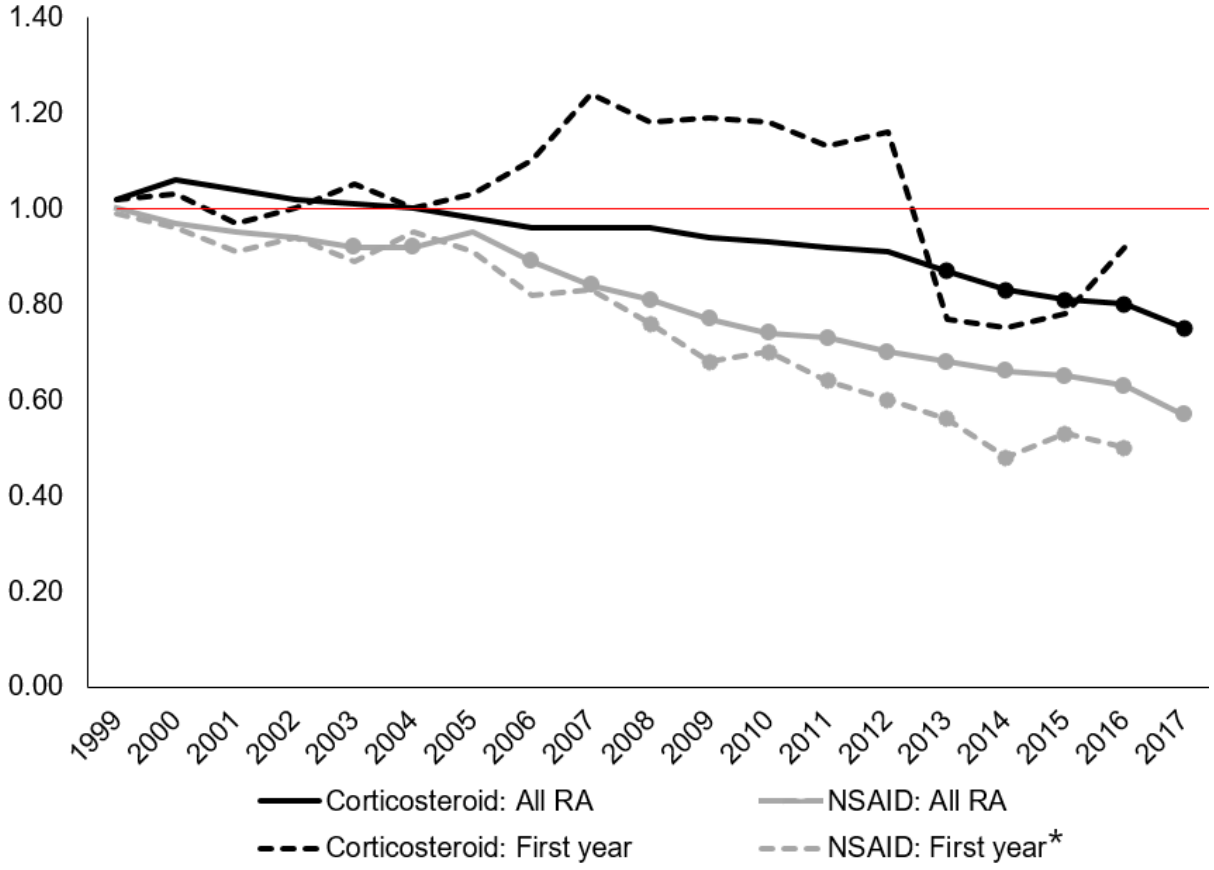
In the incident RA cohort in sub-analysis B, with medication selected from specific BNF chapters, the corticosteroid prescription count per person-year was lower: 0.04 in 1998 and 0.02 in 2017 (Figure 83).

9.3.2.2 Prescribing Duration

During follow-up, 45.1% of RA patients received corticosteroid prescribing; 32.2% received ≥90 days in a year and 25.7% received ≥180 days in a year. In sensitivity analysis RA1 this was 52.5%, 37.9% and 30.1% respectively; in sensitivity analysis RA2 this was 53.5%, 39.5% and 31.8%. Among 1,008,623 corticosteroid prescriptions during the study period, the modal number of days between prescriptions was 28 days, with the gap between prescriptions being ≥2 months in 13.7% of cases (Figure 84).

The proportion with any, ≥90 days and ≥180 days of corticosteroid prescribing in a year was 25.1%, 21.0% and 17.6% in 1998 and 26.9%, 19.4% and 16.0% in 2012 and 24.0%, 15.5% and 12.3% in 2017 (Figure 85, Table 38, Figure 85). The decline in the proportion receiving long-term corticosteroid prescribing was significant between 2013 (IRR 0.87, 95% CI 0.81-0.94) and 2017 (IRR 0.75, 95% CI 0.70-0.80) (mean APC -1.54) (Figure 91). In sensitivity analysis RA1, this was 30.1%, 25.7% and 21.7% in 1998; 27.9%, 20.4% and 17.3% in 2012; and 25.4%, 17.2% and 13.4% in 2016. In sensitivity analysis RA2, this was 32.2%, 27.4% and 23.2% in 1998; 30.0%, 22.2% and 18.3% in 2012; and 26.8%, 18.5% and 14.5% in 2016. The mean APC in the proportion of patients with long-term prescribing for five-year bands was +0.23 in 1998-2002, -0.63 in 2003-2007, -1.06 in 2008-2012 and -4.36 in 2013-2017.

Figure 91. Annual adjusted IRRs for having ≥90 days medication prescribing in 1999-2017 compared with 1998: all RA patients and in the year post-diagnosis



Note: Adjusted for sex and age-group. NSAID = Non-steroidal anti-inflammatory drug: IRR = incidence risk ratio

*GP practice included as a random intercept

Dot indicates P value from Wald test <0.001 in that year

Red line indicates IRR = 1

Women were slightly less likely to receive long-term corticosteroids than men (IRR 0.96, 95% CI 0.94-0.97), with this difference predominantly being in the year post-diagnosis (IRR 0.88, 95% CI 0.85-0.92) (Table 39). Compared with age 18-29, prescribing significantly increased with age from age 50 to 89 (age 50-59: IRR 1.27, 95% CI 1.16-1.39; age 80-89: IRR 2.18, 95% CI 1.99-2.40), then less so for patients aged 90-99 (IRR 1.60, 95% CI 1.44-1.78) (Figure 92). The predicted annual count of patients with long-term corticosteroid prescribing, per sex and per age group, differed little from the observed data (Figure 93). Little variation was explained by adding GP practice as a random intercept, (variance 0.15, standard deviation 0.39) and the AIC was higher (222,143 compared with212,113.8) and so the quasi-Poisson fixed effects model was chosen as the dispersion parameter was 1.62 in the fixed effects model (Appendix D: Table D 1). Given the study focus on temporal change, the model was compared with an equivalent model excluding calendar year via a Chi^2 test, which

revealed a significant increase in residual deviance (1150.28 compared with 481.18; P<0.001). In the subset with IMD recorded, there was little variation in prescribing across IMD in comparison to the first quintile (quintile 2: 0.96, 95% CI 0.94-0.98; quintile 5: IRR 0.97, 95% CI 0.94-0.99) and the model AIC was lower when IMD was excluded (8631.9 compared with 197,479).

Figure 92. Annual percentage of RA patients with ≥90 days corticosteroid prescribing by age-group and sex, 1998-2017 (N = 21,726)

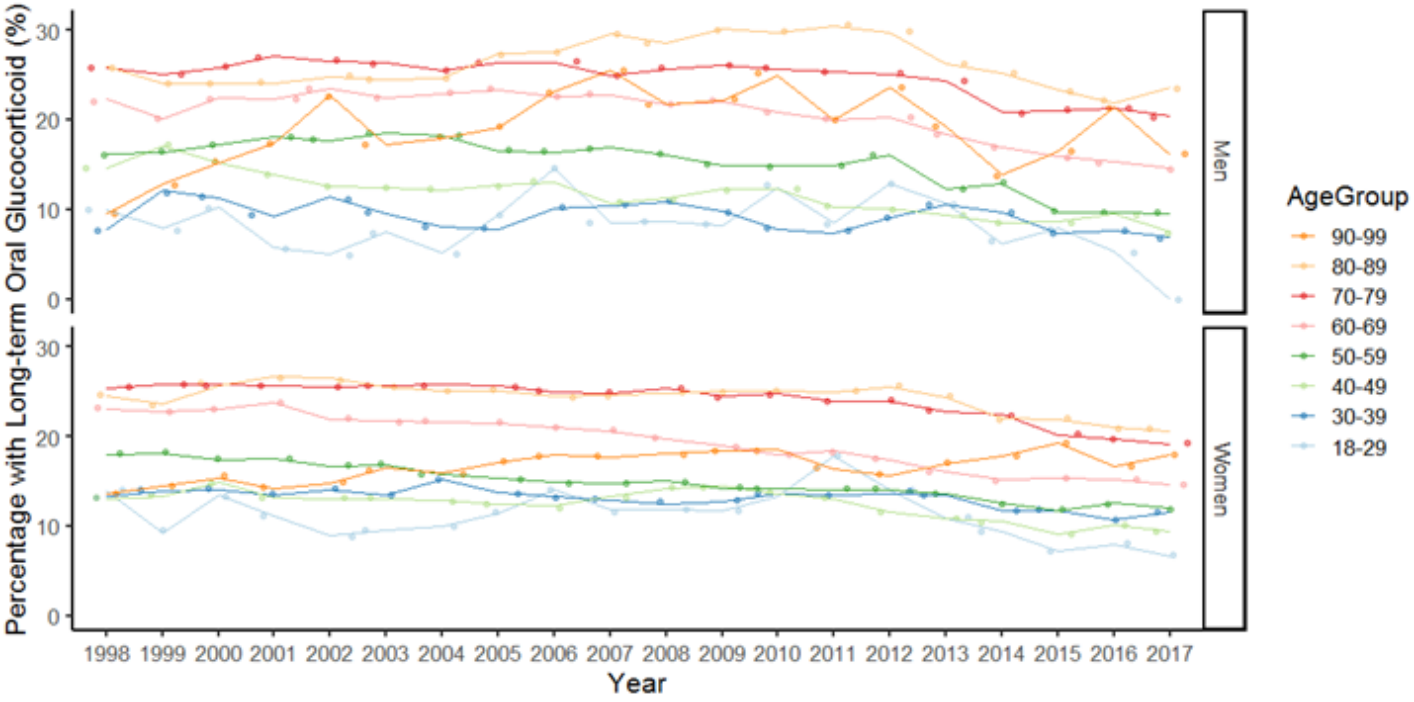
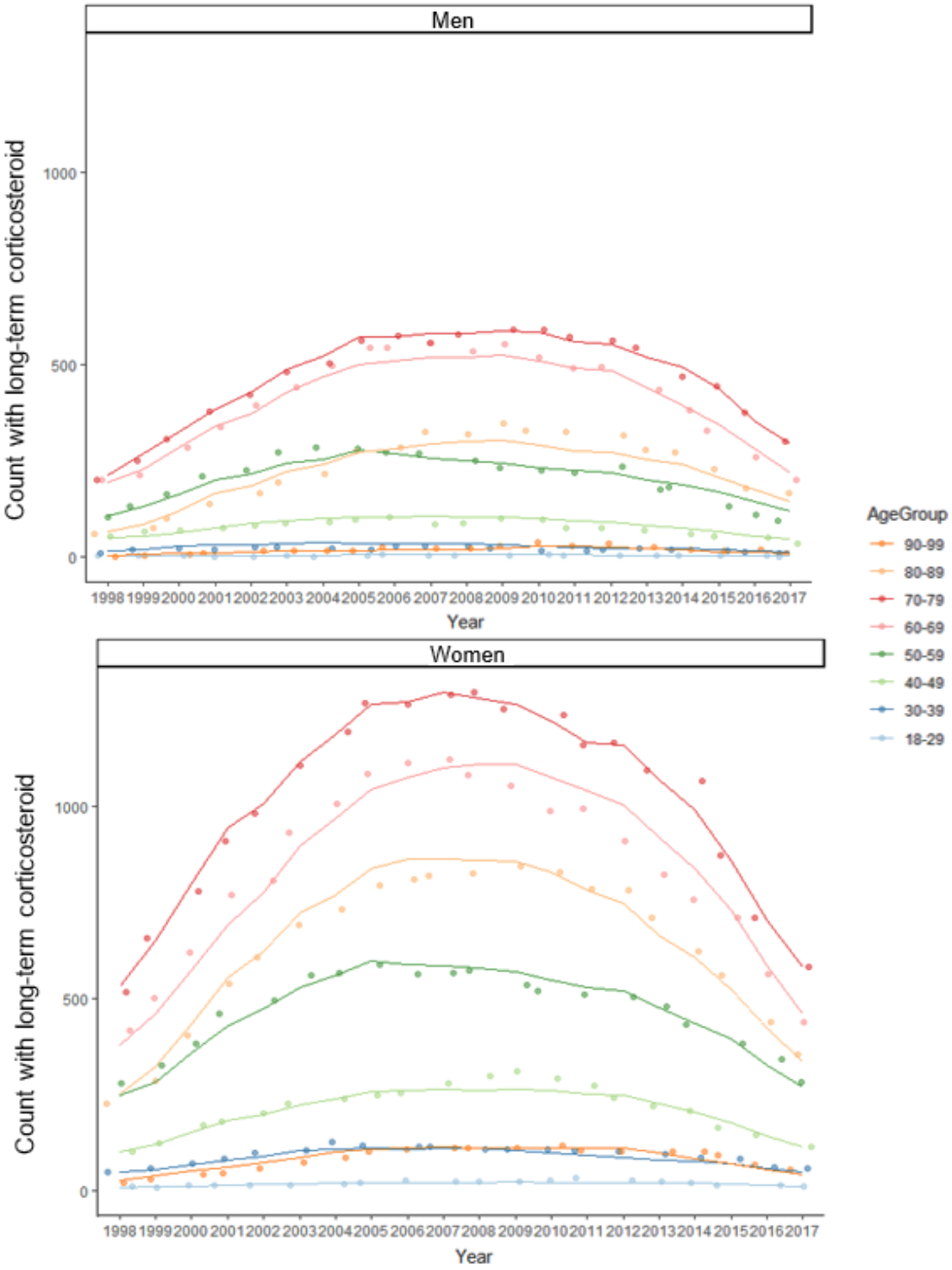


Figure 93. Annual predicted (line) and observed (dot) count of RA patients with ≥90 days corticosteroid prescribing by age group and sex, 1998-2017 (N = 71,411)



9.3.2.2.1 Incident Cohort

In incident RA patients, the proportion with any, ≥90 days and ≥180 days of corticosteroid prescribing in the year post-diagnosis was 30.4%, 22.5% and 16.9% respectively. This was 32.4%, 23.9% and 17.5% in sensitivity analysis RA1 and 34.3%,

25.7% and 19.0% in sensitivity analysis RA2. The proportion with any, ≥90 days and ≥180 days of corticosteroid prescribing in the year post-diagnosis was 29.6%, 22.2% and 18.6% in 1998 and 34.8%, 25.5% and 17.9% in 2012, then fell in 2013-2014 before reaching 28.5%, 19.1% and 13.8% in 2016 (Figure 85, Figure 87). The pattern was similar in sensitivity analysis RA1: 32.4%, 26.3% and 21.9% in 1998; 37.1%, 27.5% and 19.5% in 2012; and 34.7%, 24.8% and 17.6% in 2016. In sensitivity analysis RA2, it was 33.9%, 26.1% and 22.0% in 1998; 38.2, 28.3 and 19.5% in 2012; and 34.0%, 23.5% and 16.9% in 2016.

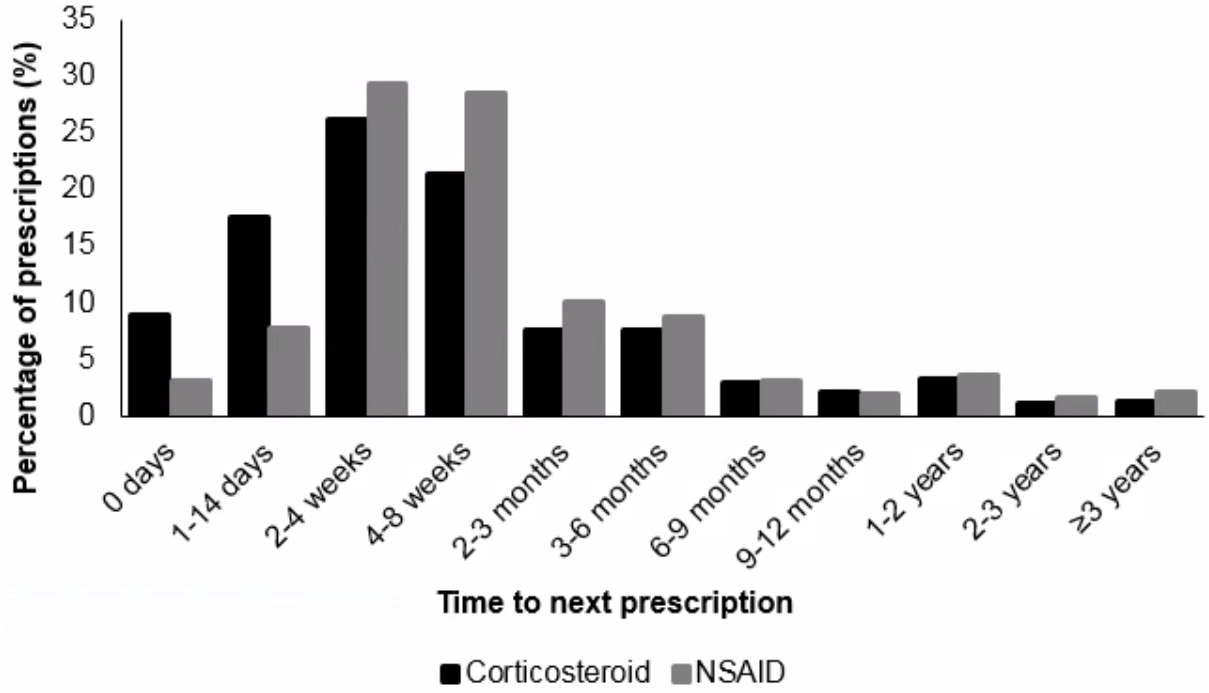
Long-term corticosteroid prescribing in the year post-diagnosis remained stable except for being significantly lower in 2013-14 (in 2013: IRR 0.77, 95% CI 0.69-0.86; in 2014: IRR 0.75, 95% CI 0.68-0.83). Women were less likely to receive long-term corticosteroids in the year post-diagnosis (IRR=0.88, 95% CI: 0.85-0.92). Prescribing was low until age 49, then increased with age from 50-59 (IRR 1.33, 95% CI 1.09-1.62) to 80-89 (IRR 2.60, 95% CI 2.12-3.19) before starting to decline. In the subset with IMD, long-term corticosteroid prescribing in the year post-diagnosis did not change over time. It was lower among women (0.88, 95% CI 0.84-0.92) and increased with age (50-59: 1.70, 95% CI 1.28-2.27; 80-89: 3.45, 95% CI 2.57-4.62) before starting to fall among patients aged 90-99 (2.66, 95% CI 1.85-3.82). There was no significant difference between IMD quintiles.

In Poisson regression modelling, the coefficients were similar when GP practice was included (Appendix D: Table D 2) and the effect of the random intercept was not very dominant (variance: 0.09, SD: 0.30). The AIC was slightly lower in the random intercept model (29,772 compared with 29,975) but the Hausman test indicated that the fixed effects model was suitable (P<0.001). The data was not over-dispersed (dispersion parameter: 0.70); a generalised Poisson regression model was attempted in case of under-dispersion but this produced multiple warnings. There were only 21 (6.62%) data groups with 0 patients having corticosteroids, and the zeroes were not caused by a different process and so a zero-limited model was not performed. In the subset with IMD, the Hausman test supported the alternative hypothesis that the fixed effect model sufficed (P<0.001) and the data was not over-dispersed (dispersion parameter: 0.80); a fixed effect Poisson regression model was chosen. The IMD quintile coefficients had no significance and the AIC was higher with their inclusion in the model (19,121 compared with 19,116); therefore IMD was excluded from the model.

9.3.2.2.2 Non-Rheumatoid Arthritis Cohort

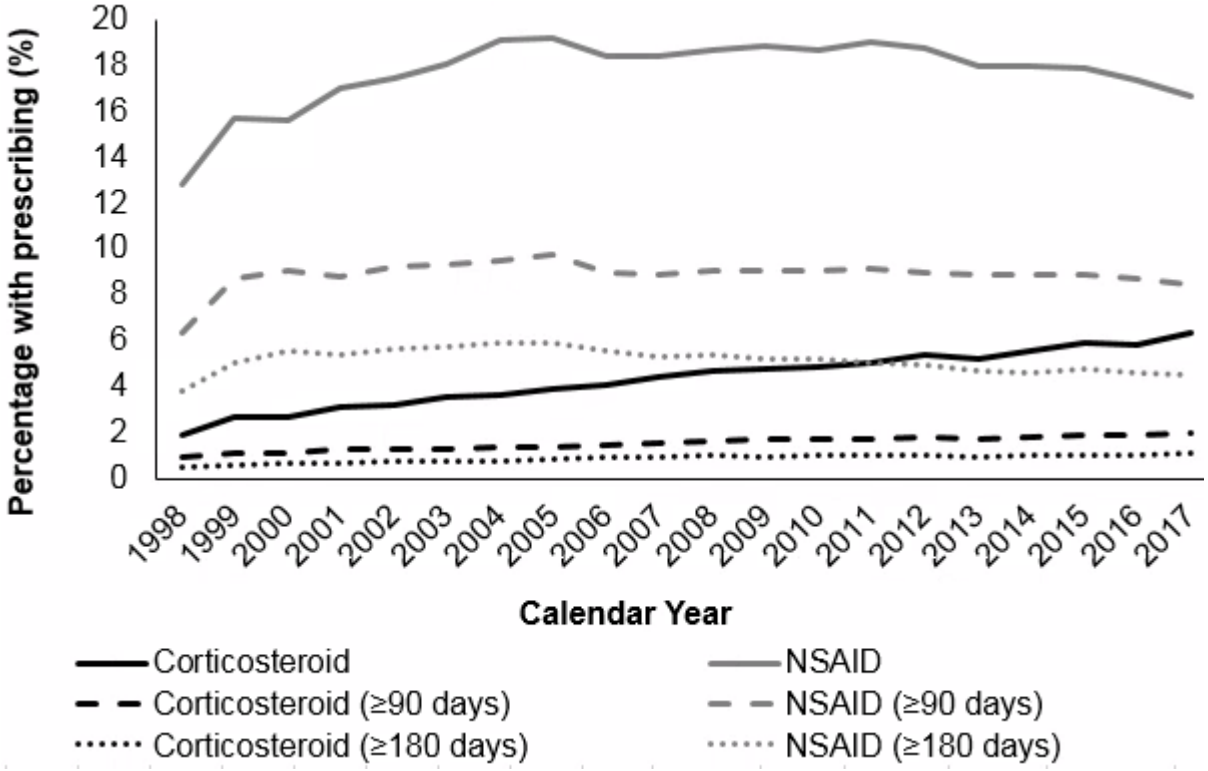
In the non-RA cohort, 11.8% had a corticosteroid during follow-up; 3.8% and 1.9% had ≥90 and ≥180 days of corticosteroid prescribing in a year. The most common duration between prescriptions was 28 days (5.3% of cases); however, in 8.9% of instances ≥2 prescriptions were recorded on the same day (Figure 94). The proportion of non-RA patients with any, ≥90 and ≥180 days of corticosteroid prescribing per person-year was 1.9%, 0.9% and 0.5% in 1998 and 6.3%, 2.0% and 1.1% in 2017, with no change in 2013-14 in the upward trend (Figure 95).

Figure 94. The proportion of prescriptions followed by a gap of 0 days to ≥3 years before the next prescription, in the non-RA cohort (N = 87,611)



Note: NSAID = Non-steroidal anti-inflammatory drug

Figure 95. Annual percentage of non-RA patients with any (N = 205,188), ≥90 (N = 195,636) and ≥180 days (N = 183,803) medication prescribing, 1998-2017



Note: NSAID = Non-steroidal anti-inflammatory drug

9.3.3 Trends in NSAID Prescribing

9.3.3.1 Prescription Counts

During follow-up, 49,431 (69.2%) RA patients had prescribed NSAIDs; 34,201 (77.0%) and 34,232 (75.3%) in sensitivity analyses RA1 and RA2.

In RA patients, the NSAID prescription count fell from 1998 to 1999, before stabilising across 2000-2005 and then declining at a steadily reducing APC (i.e. at a flattening rate). The mean NSAID prescription count per person-year was 4.17 in 1998 and 1.96 in 2017 with a mean APC of -3.80 (Figure 67). In sensitivity analyses RA1 and RA2, this was 4.88 and 5.15 in 1998 and 2.14 and 2.27 in 2016, with a mean APC of -4.44 and -4.43 (Figure 68). Excluding DMARD-naïve patients, the mean NSAID prescription count per person-year was higher: 6.11 in 1998 and 2.16 in 2017 (N = 42,545).

In RA patients receiving ≥1 NSAID in a given year, the mean prescription count per person-year was 7.08 in 1998 and 5.58 in 2017 (mean APC: -1.41) (Table 37). In

sensitivity analyses RA1 and RA2 this was 7.39 and 7.59 in 1998 and 5.69 and 8.30 in 2016. Where RA patients received NSAIDs, the most common number of prescriptions in a year was 1, 2 and 6 (Figure 69). In patients prescribed NSAIDs, the median number of prescriptions per year was 6 in 1998-1999, 5 in 2000-2010, 4 in 2011-2014 and 3 in 2015-2017.

The annual proportion of RA patients with ≥6 NSAID prescriptions and a full year of GP registration in a given year showed a similar trend. This was 33.0% in 1998, 28.7% in 2005 and 14.8% in 2017 (mean APC: -4.12) (Figure 70). In sensitivity analyses RA1 and RA2, this was 39.4% and 41.8% in 1998 and 16.1% and 17.2% in 2016 (Figure 71).

During follow-up, 28,134 (68.3%) matched RA patients had prescribed NSAIDs. The mean count of prescriptions per person-year for matched RA patients fell from 5.40 in 1998 to 1.92 in 2017 (Figure 72). The decline was greatest in 1998-2000 (mean APC -13.91) and then limited until 2005 (mean APC -1.81) before declining with reducing APC between 2006 and 2017 (mean APC -5.18). Excluding DMARD-naïve patients, the mean count of prescriptions per person-year was higher: 6.67 in 1998 and 2.11 in 2017 (N = 29,963).

In sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), the mean count of NSAID prescriptions per person-year was 3.07 in 1998 and rose to 3.47 in 2005 before declining to and stable until 2013 before declining to 1.65 in 2013 and then more slowly to 1.04 in 2017 (Figure 73). Excluding DMARD-naïve patients, the mean count of NSAID prescriptions per person-year was higher: 4.11 in 1998 and 1.19 in 2017 (N = 33,084). In RA patients with ≥1 prescription in a given year, in sub-analysis A this was 6.59 in 1998 and 6.89 in 2005 (mean APC 1998-2005: +0.66) before declining to 5.92 in 2014 (mean APC 1999-2014: -1.66) and then reaching 6.24 by 2017 (Figure 74).

In sub-analysis B (where medication was selected from specific BNF chapters), prescription counts were lower (Figure 75). The NSAID prescription count per person-year fell from 2.56 in 1998 to 0.97 in 2017. In RA patients with ≥1 prescription in a given year, there was little variation in prescription count: 6.24 in 1998 and 6.08 in 2017. Excluding DMARD-naïve patients, the NSAID prescription count per person-year in the RA cohort was higher: 3.90 in 1998 and 1.13 in 2017 (Figure 90).

9.3.3.1.1 Incident Cohort

In patients with incident RA, 68.8% (21,848) received an NSAID prescription during the study period. This was 77.3% (14,544) in sensitivity analysis RA1 and 73.7% (16,135) in sensitivity analysis RA2. In this incident cohort, the mean NSAID prescription count per person-year was 5.36 in 1998, falling to 1.86 in 2017(Figure 76). Excluding DMARD-naïve patients, the mean prescription count per person-year was lower: 4.57 in 1998 and 0.99 in 2017. In the incident RA cohort, the mean prescription count per person-year across five-year bands was 4.34 in 1998-2002, 3.50 in 2003-2007, 2.64 in 2008-2012 and 2.03 in 2012-2017. In sensitivity analyses RA1 and RA2, this was 4.98 and 5.04 in 1998-2002, 3.94 and 3.98 in 2003-2007, 2.83 and 2.84 in 2008-2012 and 1.16 and 2.25 in 2013-2016 (four-year band) ().

In the year post-diagnosis, the mean prescription count per person-year was 5.00 in 1998, falling to 1.90 in 2016. In sensitivity analyses RA1 and RA2, this was 5.89 and 5.87 in 1998 and 2.09 and 2.35 in 2016, having fallen to 1.75 and 2.06 in 2014. The proportion with ≥1 prescription in the year post-diagnosis was 71.8% in 1998 and 32.1% in 2016 (Figure 79). In sensitivity analyses RA1 and RA2, this was 77.8% and 76.5% in 1998 and 44.1% and 42.3% in 2016 (Figure 80).

In sub-analysis A (where the subsequent RA code ≥6 months after the first was used to assign the date of RA diagnosis), in the incident RA cohort the mean NSAID prescription count declined from 4.31 in 1998 to 0.98 in 2016. There was a similar trend in the year post-diagnosis; 3.91 in 1998 and 0.75 in 2016 (Figure 81). The proportion of RA patients with an NSAID prescription in the year post-diagnosis was 59.2% in 1998 and 12.3% in 2016 (Figure 82).

In the incident RA cohort in sub-analysis B (where medication was selected from specific BNF chapters), the NSAID prescription count per person-year was 3.47 in 1998 and 0.87 in 2017 (Figure 83). In the year post-diagnosis, this was 3.43 in 1998 and 0.94 in 2016.

9.3.3.2 Prescribing Duration

During follow-up, 69.0% of RA patients received NSAID prescribing; 55.8% received ≥90 days in a year and 43.2% received ≥180 days in a year. In sensitivity analysis RA1, this was 76.8%, 63.4% and 49.7% respectively; in sensitivity analysis RA2, this

was 75.1%, 62.8% and 50.4%. Among 1,342,661 NSAID prescriptions during the study period, the modal number of days between prescriptions was 28 days, with the gap between prescriptions being ≥2 months in 20.9% of cases (Figure 84).

The proportion of RA patients with any, ≥90 days and ≥180 days of NSAID prescribing in a year was 57.9%, 45.9% and 36.6% in 1998 and 35.6%, 25.1% and 18.4% in 2017 (Figure 85, Figure 86, Table 38). The proportion with long-term NSAID prescribing fell by mean APC -0.99 between 1998 and 2005 (45.9% to 42.7%) and then by mean APC -4.33 between 2006 and 2017 (40.0% to 25.1%). The decline in the proportion receiving long-term NSAID prescribing was significant from 2003 (IRR 0.92, 95% CI 0.88-0.95) to 2017 (IRR 0.57, 95% CI 0.54-0.60) (mean APC -3.10), excepting in 2005 (Table 39). In sensitivity analysis RA1, this was 65.3%, 53.4% and 43.3% in 1998 and 39.0%, 27.9% and 20.0% in 2016. In sensitivity analysis RA2, this was 66.9%, 55.8% and 46.0% in 1998 and 40.2%, 18.5% and 21.4% in 2016. The mean APC in the proportion of patients with long-term prescribing for five-year bands was -2.53 in 1998-2002, -1.76 in 2003-2007, -3.79 in 2008-2012 and -4.21 in 2013-2017.

Table 39. Adjusted IRRs for having ≥90 days medication prescribing in a year: all RA patients and in the year post-diagnosis

	Adjusted IRR (95% CI)			
	Corticosteroid		**NSAID**	
	All RA patients (N = 68,939)	**First year from diagnosis (N = 30,799)**	**All RA patients (N = 68,939)**	**First year from diagnosis (N = 30,799)†**
Calendar year				
1998	1	1	1	1
1999	1.02 (0.95 to 1.10)	1.02 (0.89 to 1.17)	1.00 (0.96 to 1.05)	0.99 (0.91 to 1.08)
2000	1.06 (0.99 to 1.13)	1.03 (0.90 to 1.17)	0.97 (0.92 to 1.01)	0.96 (0.89 to 1.05)
2001	1.04 (0.97 to 1.12)	0.97 (0.85 to 1.10)	0.95 (0.91 to 0.99)	0.91 (0.84 to 0.99)
2002	1.02 (0.96 to 1.08)	1.00 (0.89 to 1.12)	0.94 (0.90 to 0.97)	0.94 (0.86 to 1.01)
2003	1.01 (0.95 to 1.07)	1.05 (0.94 to 1.18)	0.92 (0.88 to 0.95)*	0.89 (0.82 to 0.97)
2004	1.00 (0.94 to 1.06)	1.00 (0.89 to 1.12)	0.92 (0.89 to 0.95)*	0.95 (0.88 to 1.03)
2005	0.98 (0.92 to 1.04)	1.03 (0.92 to 1.15)	0.95 (0.92 to 0.98)	0.91 (0.85 to 0.99)
2006	0.96 (0.91 to 1.02)	1.10 (1.00 to 1.22)	0.89 (0.86 to 0.92)*	0.82 (0.76 to 0.89)
2007	0.96 (0.91 to 1.02)	1.24 (1.09 to 1.41)	0.84 (0.81 to 0.88)*	0.83 (0.76 to 0.90)
2008	0.96 (0.90 to 1.01)	1.18 (1.07 to 1.31)	0.81 (0.78 to 0.84)*	0.76 (0.70 to 0.82)*
2009	0.94 (0.89 to 1.00)	1.19 (1.06 to 1.33)	0.77 (0.74 to 0.80)*	0.68 (0.62 to 0.74)*

2010	0.93 (0.87 to 0.99)	1.18 (1.08 to 1.30)	0.74 (0.71 to 0.77)*	0.70 (0.64 to 0.76)*
2011	0.92 (0.87 to 0.98)	1.13 (1.01 to 1.26)	0.73 (0.70 to 0.75)*	0.64 (0.58 to 0.70)*
2012	0.91 (0.85 to 0.98)	1.16 (1.04 to 1.31)	0.70 (0.67 to 0.72)*	0.60 (0.54 to 0.66)*
2013	0.87 (0.81 to 0.94)*	0.77 (0.69 to 0.86)	0.68 (0.66 to 0.70)*	0.56 (0.51 to 0.61)*
2014	0.83 (0.77 to 0.89)*	0.75 (0.68 to 0.83)	0.66 (0.64 to 0.68)*	0.48 (0.44 to 0.53)*
2015	0.81 (0.76 to 0.86)*	0.78 (0.68 to 0.90)	0.65 (0.62 to 0.67)*	0.53 (0.47 to 0.59)*
2016	0.80 (0.75 to 0.85)*	0.92 (0.82 to 1.03)	0.63 (0.61 to 0.65)*	0.50 (0.44 to 0.56)*
2017	0.75 (0.70 to 0.80)*		0.57 (0.54 to 0.60)*	
Sex				
Male	1	1	1	1
Female	0.96 (0.94 to 0.97)*	0.88 (0.85 to 0.92)*	1 (0.99 to 1.01)	1.02 (0.99 to 1.05)
Age-group				
18-29	1	1	1	1
30-39	1.07 (0.97 to 1.19)	1.05 (0.84 to 1.30)	1.15 (1.10 to 1.20)*	1.13 (1.03 to 1.25)
40-49	1.05 (0.95 to 1.15)	1.03 (0.83 to 1.27)	1.32 (1.26 to 1.37)*	1.17 (1.07 to 1.29)
50-59	1.27 (1.16 to 1.39)*	1.33 (1.09 to 1.62)	1.34 (1.28 to 1.39)*	1.16 (1.06 to 1.27)
60-69	1.67 (1.52 to 1.84)*	1.73 (1.42 to 2.11)*	1.24 (1.20 to 1.30)*	1.09 (1.00 to 1.20)
70-79	2.08 (1.90 to 2.27)*	2.35 (1.93 to 2.85)*	0.99 (0.95 to 1.04)	0.91 (0.83 to 1.00)
80-89	2.18 (1.99 to 2.40)*	2.6 (2.12 to 3.19)*	0.74 (0.70 to 0.77)*	0.72 (0.65 to 0.81)*
90-99	1.60 (1.44 to 1.78)*	2.13 (1.65 to 2.75)*	0.56 (0.52 to 0.61)*	0.75 (0.61 to 0.91)

Note: adjusted for calendar year, sex and age-group as appropriate. IRR = incidence risk ratio

†GP practice included as a random intercept

*P value from Wald test <0.001

There was no sex difference in long-term NSAID prescribing. It rose to peak with age among patients aged 50-59 (IRR 1.34, 95% CI: 1.28-1.39) before declining to age 90-99 (IRR 0.56, 95% CI: 0.52-0.61). Little variation was explained by adding GP practice as a random intercept (variance 0.04, standard deviation 0.21) and a Hausman test showed that the fixed effect model was sufficient (P<0.001), although the AIC was lower (271,101.4 compared with 275,668). A fixed effects quasi-Poisson model was selected due to over-dispersion (dispersion parameter=1.39) (Appendix D: Table D 3).

In the subset of RA patients with IMD recorded, long-term NSAID prescribing was greater among the most deprived (quintile 5: IRR 1.04, 95% CI 1.02-1.06), with no significant variation among the first four IMD quintiles. Long-term NSAID prescribing fell from 2003 (IRR: 0.92, 95% CI 0.89-0.95) to 2017 (IRR: 0.49, 95% CI 0.46-0.52). There

was no sex difference. Prescribing increased with age until 50-59 (IRR: 1.34, 95% CI 1.25-1.43) then declined until age 90-99 (IRR: 0.58, 95% CI 0.54-0.63), as with the main cohort. In the subset, the AIC was lower when IMD was included in the model (9,335.3 compared with 9,346) although had little effect on coefficients or their significance; the Hausman test supported the alternative hypothesis that the fixed effect model sufficed (P<0.001). The data in the fixed effect model was not over-dispersed (dispersion parameter: 0.69).

9.3.3.2.1 Incident Cohort

In incident RA patients, the proportion with any, ≥90 days and ≥180 days of NSAID prescribing in the year post-diagnosis was 54.3%, 42.1% and 31.2% respectively. This was 58.9%, 46.0% and 34.3% in sensitivity analysis RA1 and 58.0%, 45.8% and 34.7% in sensitivity analysis RA2. The proportion with any, ≥90 days and ≥180 days of NSAID prescribing in the year post-diagnosis was 71.8%, 57.7% and 43.7% in 1998 and declined to 37.2%, 27.1% and 17.6% in 2016 (Figure 85, Figure 87). The mean AIC was -1.88 across 1998-2002, -2.11 across 2003-2007, -6.48 across 2008-2012, and -5.01 across 2013-2016. The pattern was similar in sensitivity analysis RA1: 77.8%, 64.8% and 51.5% in 1998 and 44.1%, 31.4% and 19.8% in 2016. In sensitivity analysis RA2, it was 76.5%, 64.9% and 53.8% in 1998 and 42.3%, 31.1% and 21.3% in 2016.

Long-term NSAID prescribing in the year post-diagnosis declined significantly from 2008 (IRR 0.76, 95% CI 0.70-0.82) to 2016 (IRR 0.50, 95% CI 0.44-0.56). There was no difference between the sexes. There was a trend toward lower likelihood of prescribing with age above 50-59, with this being significant for age-group 80-89 (IRR 0.72, 95% CI 0.64-0.82). GP practice accounted for slight but significant variability (variance: 0.01, standard deviation: 0.11, Hausman P=0.12). Therefore, GP practice was included as a random intercept in Poisson regression analysis (Appendix D: Table D 4), giving a marginally lower AIC (40,959.7 compared with 40,987).

In the subset with IMD, long-term NSAID prescribing in the year post-diagnosis reduced over time and significantly from 2006 (IRR 0.78, 95% CI 0.72-0.86), to 2017 (IRR 0.38, 95% CI 0.29-0.50). There was a non-significant trend toward higher NSAID prescribing among less deprived patients, and no association with sex. Prescribing was lower in the older age groups, especially at 80-89 (0.76, 95% CI 0.64-0.90). The addition of IMD brought minimal improvement (AIC 5102.8 vs 5104.5) and minimal reduction in residual deviance (1020 vs 1029.6) at a cost of degrees of freedom (1411 vs 1415) and risk of over-fitting. A Chi^2 analysis of deviance showed that the model

was not significantly different with IMD (P = 0.05). GP practice was not included as a random intercept, despite marginal improvement in AIC (25,415.5 compared with 25,428), because the Hausman test supported the alternative hypothesis that the fixed effect model sufficed (P<0.001). The data was not over-dispersed (dispersion parameter = 0.79).

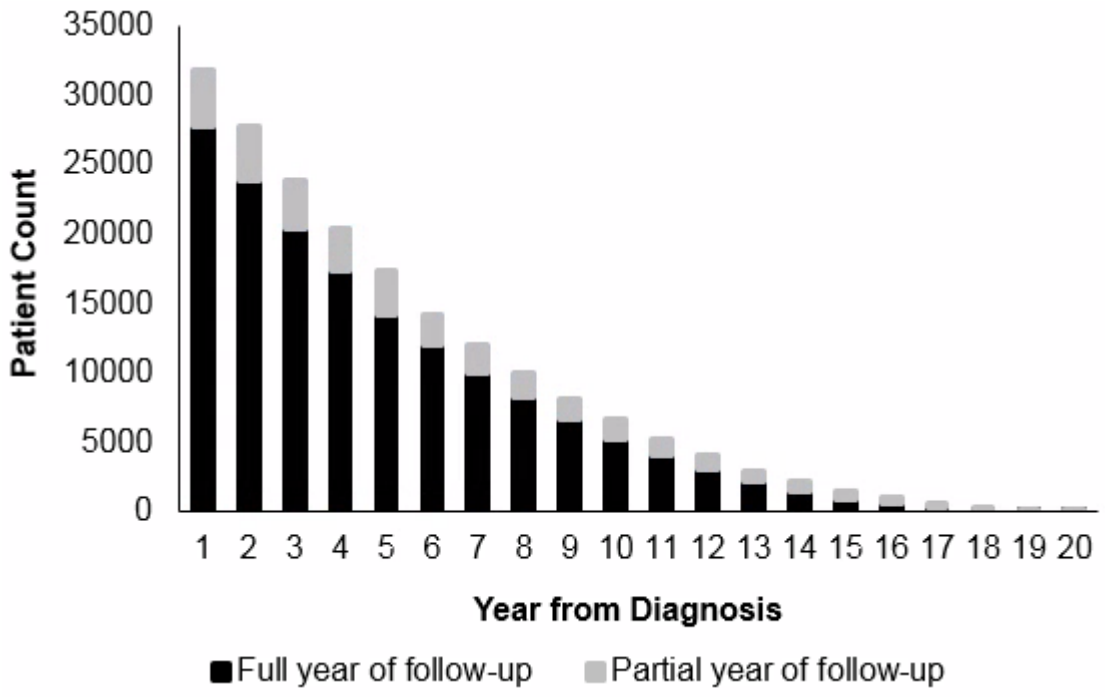
9.3.3.2.2 Non-Rheumatoid Arthritis Cohort

Among the non-RA cohort, 37.7% had an NSAID during follow-up; 19.7% and 9.7% had ≥90 and ≥180 days of corticosteroid prescribing in a year. The most common duration between prescriptions was 28 days (6.9% of cases); ≥2 prescriptions were recorded on the same day in only 3.1% of instances (Figure 94). The proportion of non-RA patients with any, ≥90 and ≥180 days of NSAID prescribing per person-year was 12.8%, 6.4% and 3.8% in 1998, rising to 19.1%, 9.5% and 6.0% in 2004 before declining to 16.7%, 8.4% and 4.5% in 2017 (Figure 95).

9.3.4 Prescribing over the life-course

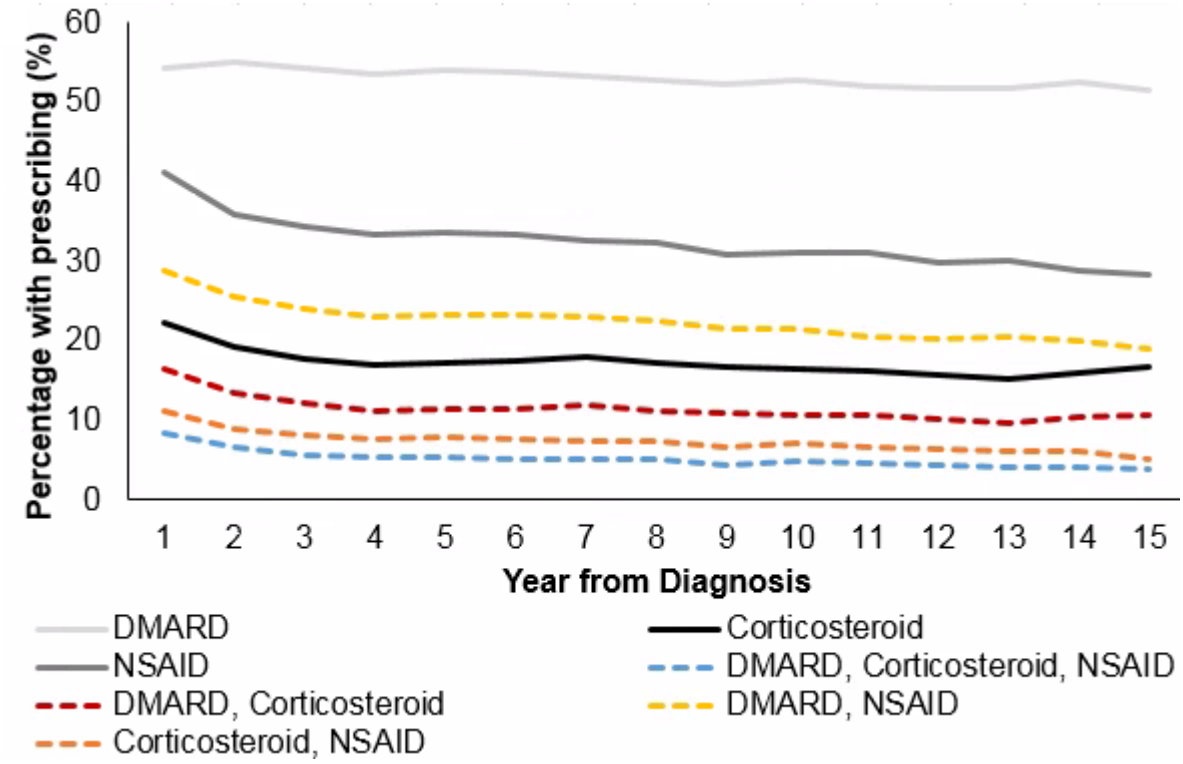
In patients with an incident diagnosis of RA during follow-up, 16.5% (6,604) had 10 years follow-up and 3.0% (1,460) had 15 years (Figure 96). Prescribing over the life-course was assessed across 15 years, after which the cohort follow-up size became <1000.

Figure 96. Count of incident RA patients with 1-20 years (full or partial year) of follow-up post-diagnosis (N = 31,678)



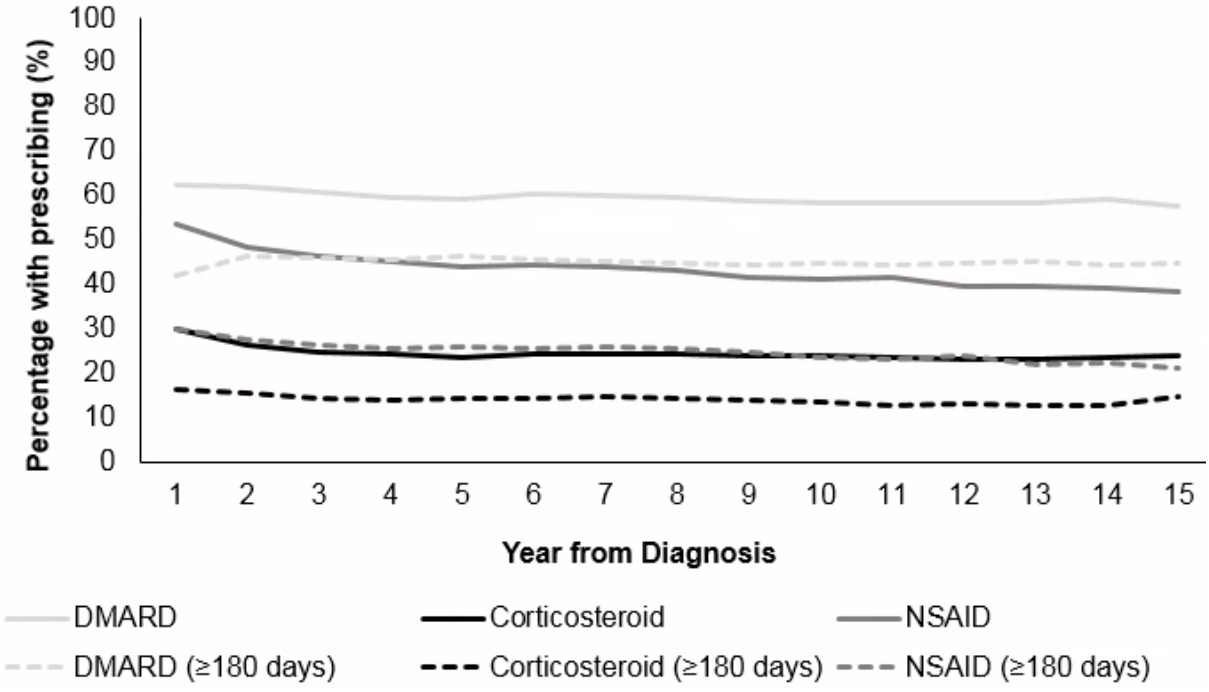
The proportion of RA patients with long-term DMARD prescribing declined non-significantly over the life-course, from 54.4% (53.9-55.0%) in the first year to 51.6% (48.9-54.3%) in the fifteenth year (mean APC: -0.37) post-diagnosis (Figure 97). The trend was similar among patients with any and ≥180 days of prescribing, excepting a delay in many patients receiving ≥180 days of DMARD prescribing until at least the second year post-diagnosis: 62.6% and 42.2% in the first year and 57.8% and 44.9% in the fifteenth year (mean APC: -0.58 and +0.48) (Figure 98).

Figure 97. Percentage of RA patients with ≥90 days prescribing in the 1-15 years post-diagnosis (N = 30,807)



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

Figure 98. Percentage of RA patients prescribed with medication (N = 31,768) or ≥180 days prescribing (N = 29,790) in the 1-15 years post-diagnosis



Note: DMARD = Disease-modifying anti-rheumatic drug; NSAID = Non-steroidal anti-inflammatory drug

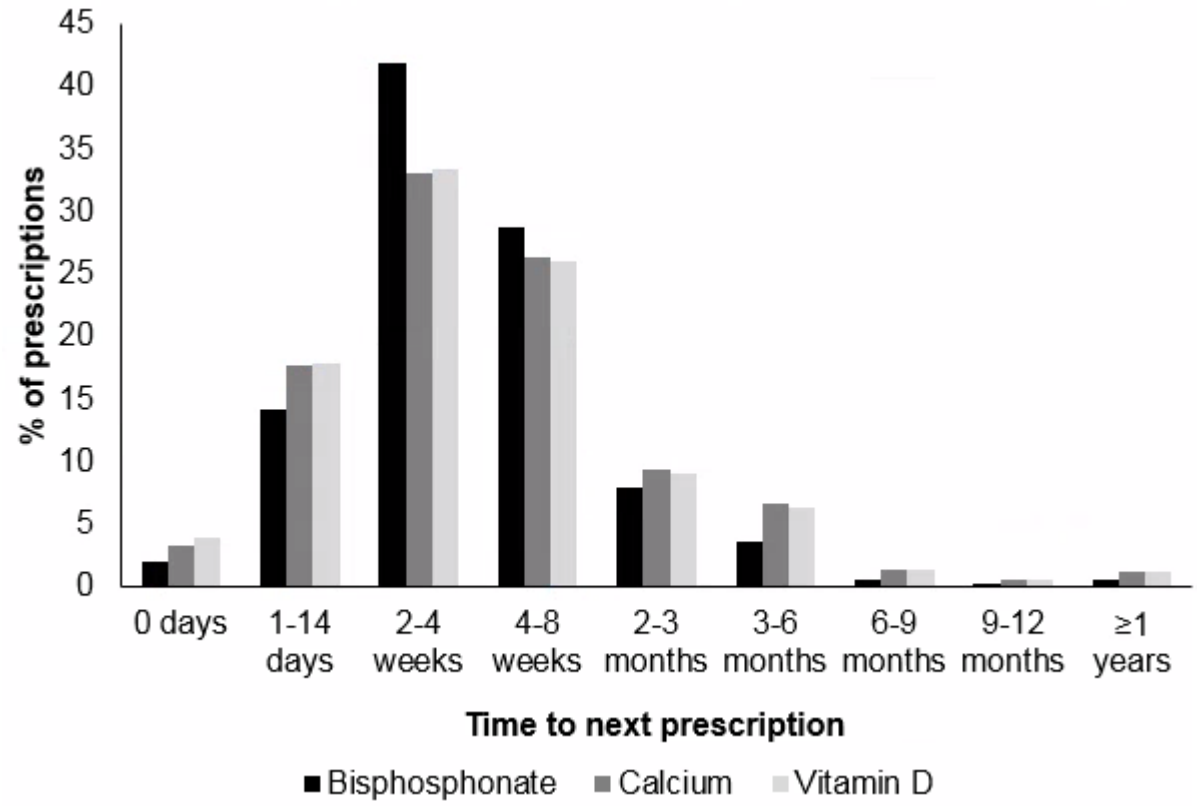
In RA patients, corticosteroid and NSAID prescribing declined over the first three years post-diagnosis (mean APC -8.25 (corticosteroids) and -6.67 (NSAIDs)). The proportion with any, ≥90 and ≥180 days of corticosteroid prescribing was 30.1%, 22.2% and 16.3% in the first year; 24.7%, 17.9% and 14.4% in the third year, with little change thereon to 23.8%, 16.8% and 14.5% in the fifteenth year. The proportion with any, ≥90 and ≥180 days of NSAID prescribing was 53.6%, 41.2% and 30.0% in the first year; 60.8%, 34.3% and 26.2% in the third year, with little change thereon to 57.8%, 28.4% and 21.2% in the fifteenth year. Assessments of combination prescribing (the proportion with ≥90 days of, for example, DMARDs and NSAIDs) showed consistent patterns.

9.3.5 Prophylaxis co-prescribing

9.3.5.1 Prednisolone

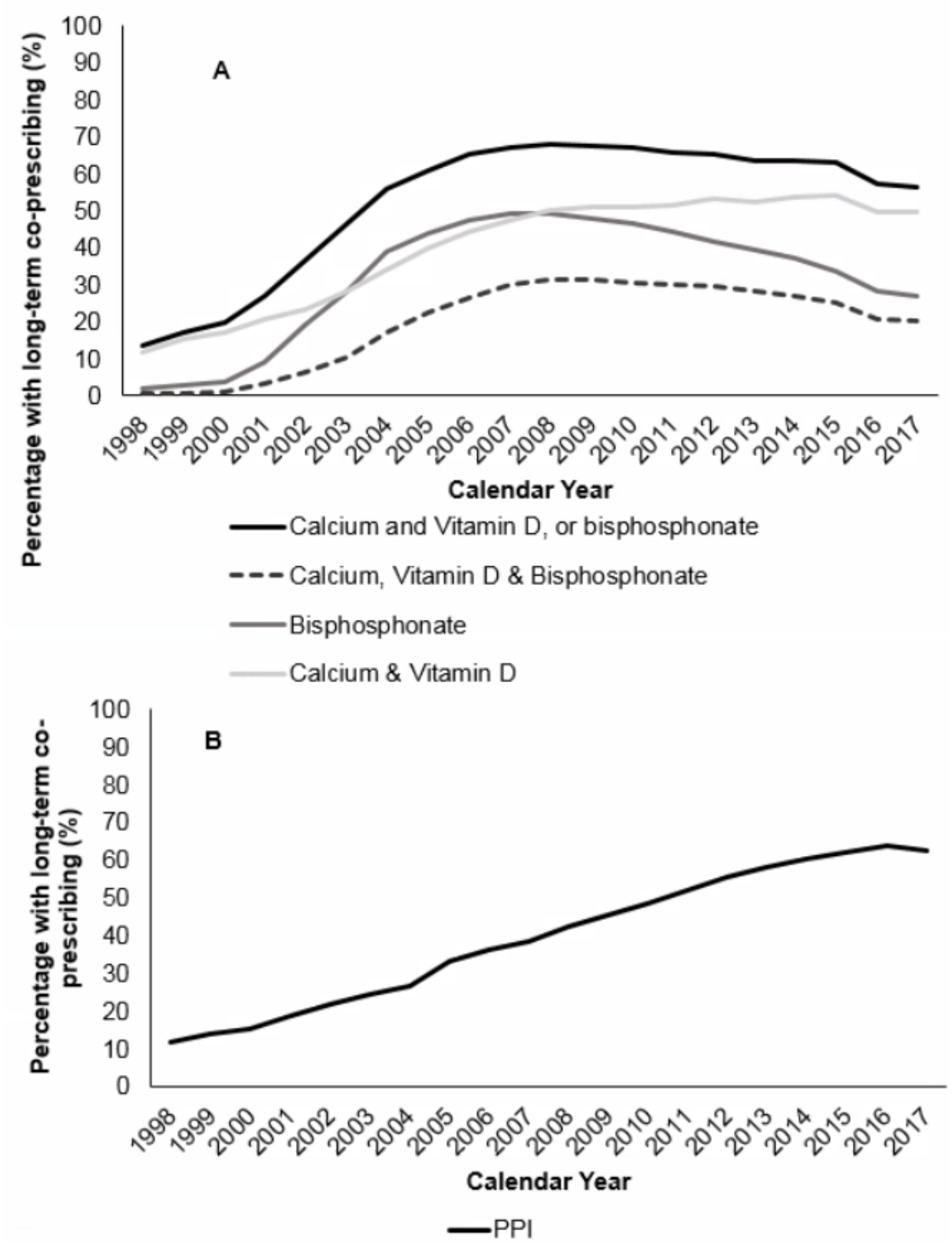
During follow-up, 15,069 women with RA had long-term prednisolone in a year; 775 of these had osteoporosis diagnosed prior to RA, leaving 14,314 patients in which bone protectant co-prescribing was assessed. The most common prescription durations for bone protectants were 7 days (10.9% [32,428] calcium, 10.9% [33,522] vitamin D) and 28 days (11.7% [31,911] bisphosphonate) (Figure 99). The mean and maximum annual prednisolone dose was calculated per patient; the median of these were 5.73 mg (IQR = 5.09) and 10.37 mg (IQR = 17.50), respectively.

Figure 99. The proportion of bone protectant prescriptions followed by a gap of 0 days to ≥1 years before the next prescription, in women with RA and a prednisolone prescription (N = 15,069)



Of 1,561 patients with long-term prednisolone in 1998, 2.1% (95% CI 1.35-2.75) were prescribed long-term bisphosphonate; 11.8% (95% CI 10.19-13.39) calcium and vitamin D and 0.4% (95% CI 0.33-0.80) calcium, vitamin D and bisphosphonate (Figure 100). This rose to 26.8% (95% CI 24.66-28.86), 49.8% (95% CI 47-45-52.20) and 20.2% (95% CI 18.30-22.10) in 2017. The proportions with calcium and with vitamin D were also plotted separately, each showing similar trends (Figure 101). Long-term bisphosphonate prescribing rose steeply to 49.4% (95% CI 47.82-50.90) in 2007 before slowly declining. The proportion with bisphosphonate co-prescribing increased with age from 9.4% among patients aged 20-39 to 50.6% among patients aged 80-99 (Figure 102). Among patients prescribed bisphosphonate in a given year, the number of prescriptions per year changed little over time, especially from 2002 (Figure 103).

Figure 100. Annual percentage of RA patients with ≥90 days of RA medication and protectant/s, 1998-2017: A) corticosteroid and bone protectant (bisphosphonate, calcium and vitamin D) (N = 14,314); B) NSAID and PPI (N = 38,480)



PPI = proton-pump inhibitors

Figure 101. Annual percentage of women with RA prescribed for ≥90 days with oral prednisolone and calcium or vitamin D, 1998-2017 (N = 14,314)

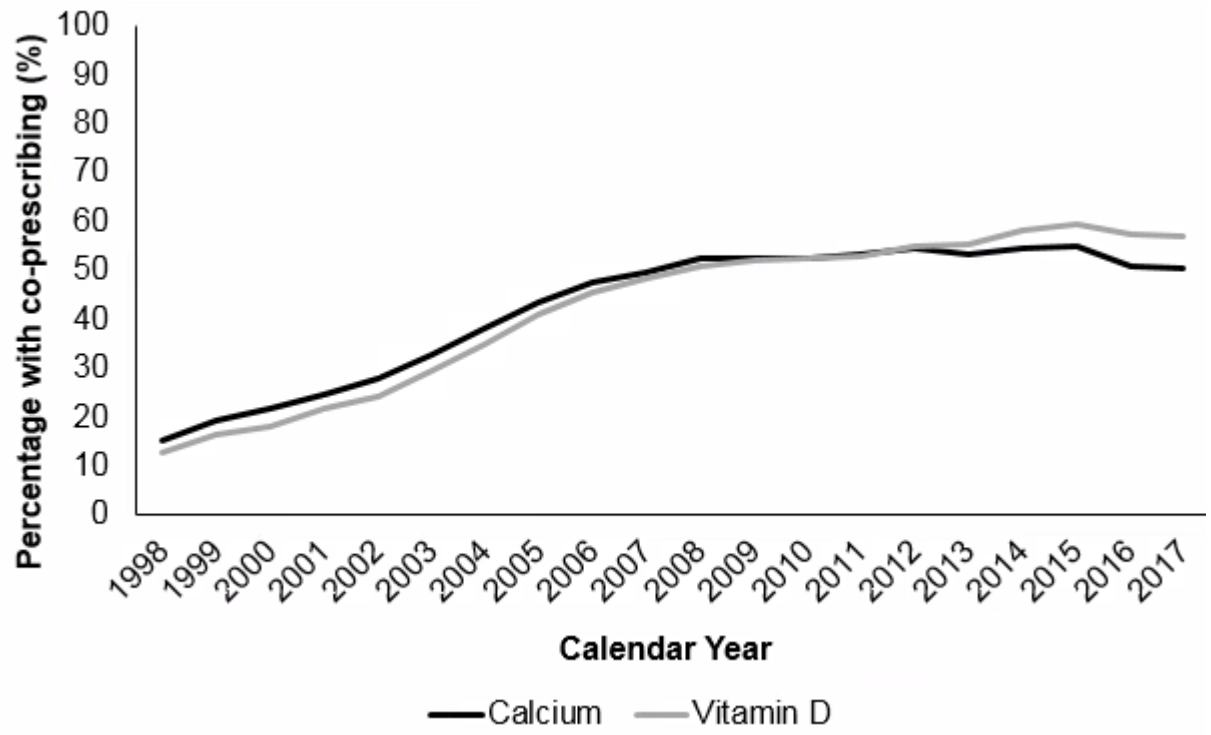


Figure 102. Annual percentage of women with RA prescribed for ≥90 days with oral prednisolone and bisphosphonate, by age-group, 1998-2017 (N = 14,314)

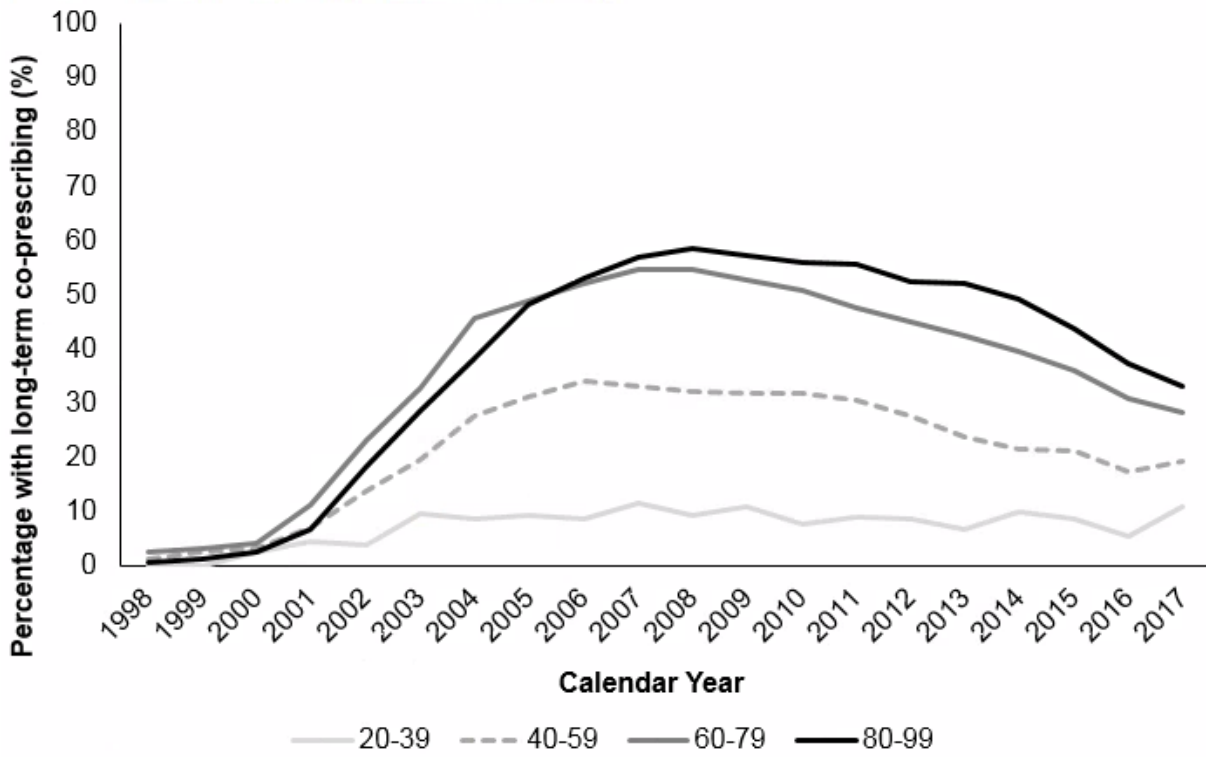
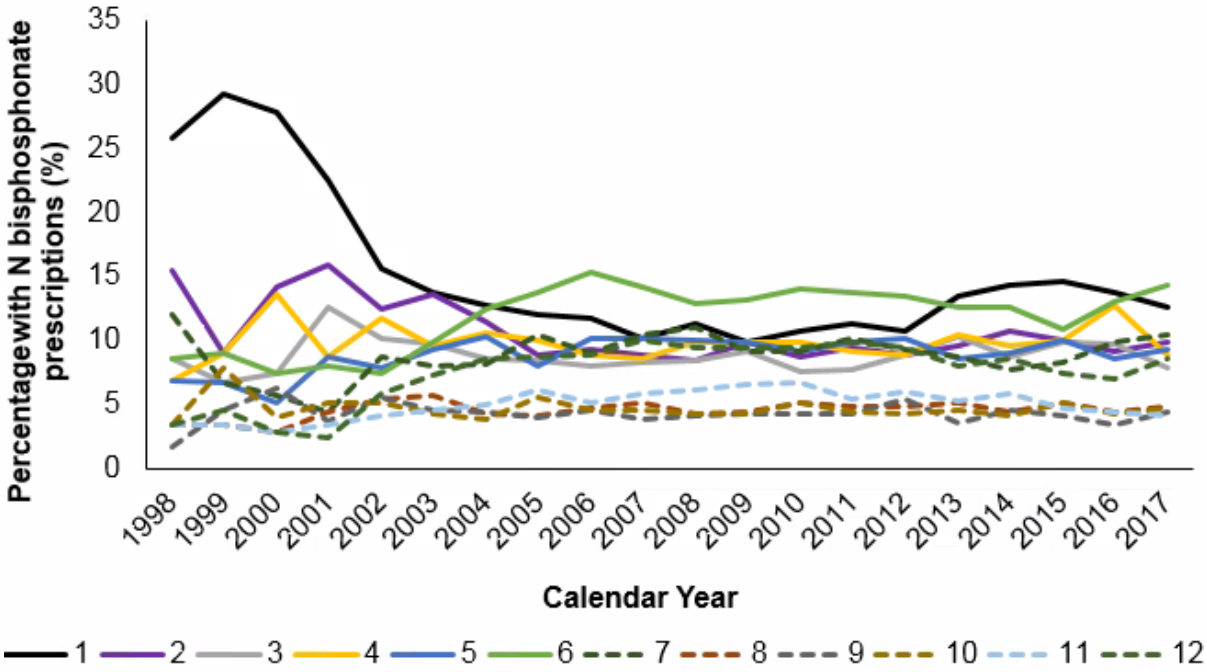
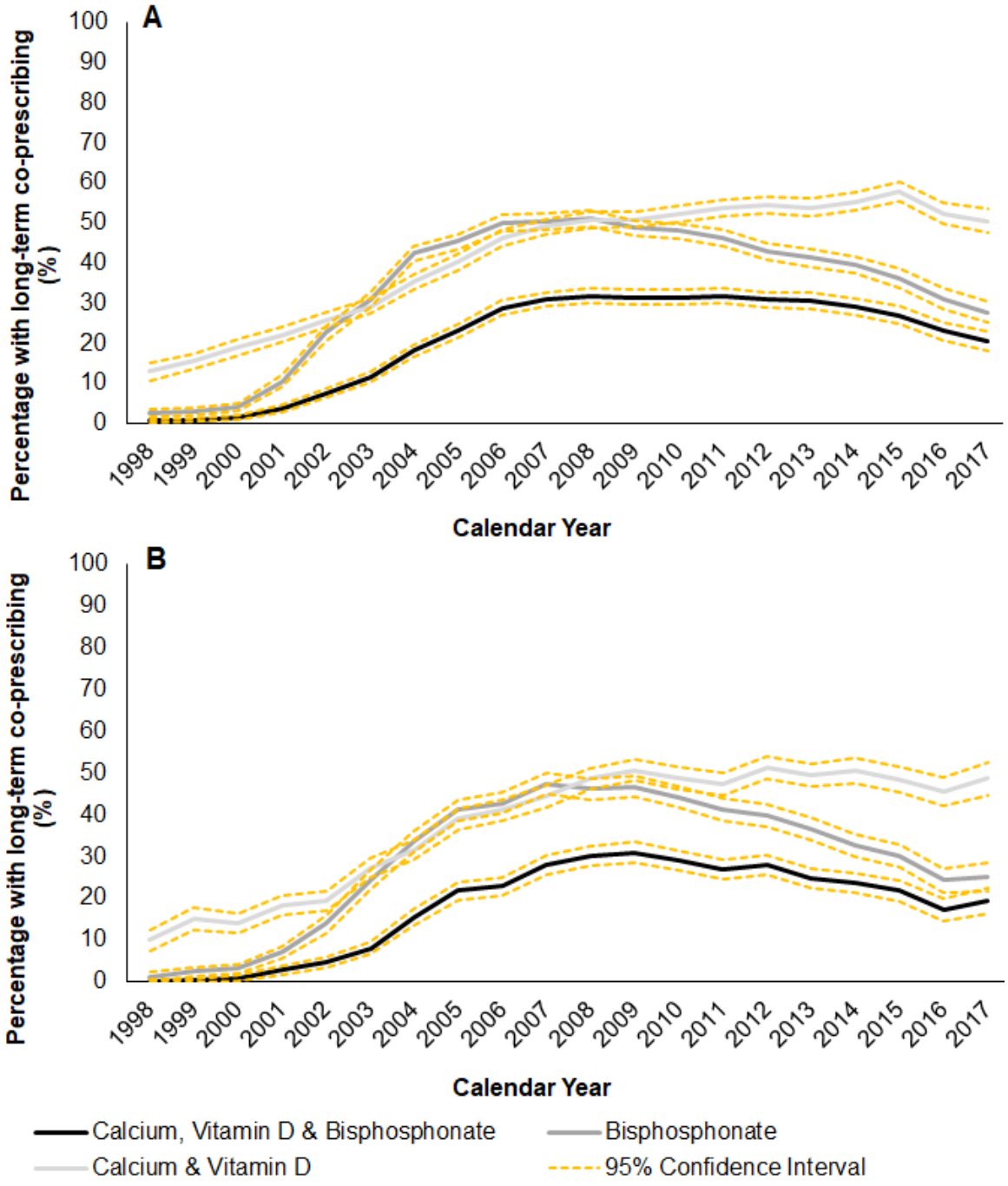


Figure 103. Annual percentage of women with RA prescribed prednisolone, with 1-12 bisphosphonate prescriptions, 1998-2017 (N =14,314)



Based on the maximum prescribed dose per year, in RA patients, the annual prescribing pattern was similar in high (n = 8,986) and low (n = 11,832) prednisolone dose cohorts (Figure 104). However, the latter cohort received fewer long-term co-prescriptions (44.5% and 36.8% bisphosphonate; 56.0% and 48.1% calcium, 55.6% and 47.7% vitamin D, 53.6% and 45.7% calcium and vitamin D, 31.8% and 25.8% calcium, vitamin D and bisphosphonate).

Figure 104. Annual percentage of women with RA prescribed for ≥90 days with high/low dose oral prednisolone and bone protectant medication, 1998-2017: A) high dose (≥7.5 mg) prednisolone (N = 8,986); B) low dose (<7.5 mg) prednisolone (N = 11,832)

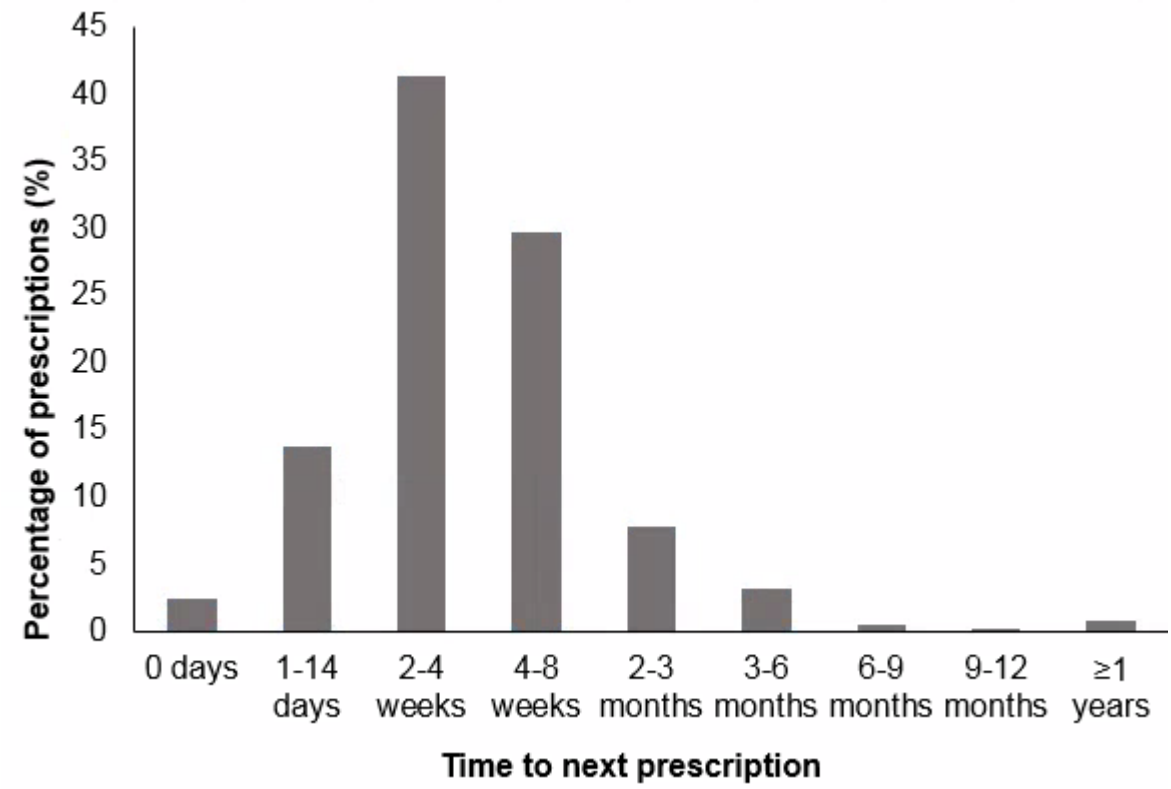


The prescribing patterns were similar in patients with ≥90 days of high-dose prednisolone prescribing (n = 5,952), compared with patients with ≥90 days of prednisolone prescribing not at high dose (n = 13,061) (e.g. 59.9% (95% CI 54.9-65.0) and 55.4% (95% CI 52.8-58.1) with calcium and vitamin D or bisphosphonate in 2017 respectively).

9.3.5.2 NSAID

In 38,480 RA patients with long-term NSAID prescribing during follow-up, there were 1,532,420 PPI prescriptions and PPI was prescribed long-term in the same year in 50.5% (19,444) of cases. This rose from 11.5% (95% CI 10.62-12.41) in 1998 to 62.64% (61.21-64.08) in 2017. The rate of change slowed across the period from APC +19.82 in 1998-1999 to -1.75 in 2016-2017 (mean APC +9.52). The duration of most PPI prescriptions was 28 days (11.1%) or 7 days (7.5%) (Figure 105).

Figure 105. The proportion of PPI prescriptions followed by a gap of 0 days to ≥1 years before the next prescription, in RA patients with an NSAID prescription (N = 38,480)



9.4 Discussion

The study addressed the objective of describing trends in real-world patient management using EHR data, specifically evaluating the pharmaceutical management of RA over two decades. The expected increase in GP DMARD prescribing in the year post-diagnosis was found to have stalled from 2009 and prescribing did not increase across the life-course. It seems that modern use of DMARDs has not facilitated reduced long-term prescribing of corticosteroids although NSAID prescribing has declined. Although corticosteroid and NSAID prescribing declined across the first three

years post-diagnosis, prescribing remained substantial three years post-diagnosis (17.9% and 34.3% respectively) and persisted over the next 12 years of follow-up. Initial increases in bone prophylaxis reversed from 2008, with declining bisphosphonate co-prescribing while vitamin D-calcium co-prescribing plateaued around 50%. GI prophylaxis co-prescribing increased more than five-fold but with a slowing pace.

Initial improvements in annual GP DMARD prescribing have stalled and there was no change across the life-course. The previously reported increase in prescribing in the year post-diagnosis between 1995 and 2010 and especially following the publication of BSR guidelines in 2006 was not apparent in this study in extant RA cases and was followed by a decline, despite national guidelines being published in 2009 (280, 477-479). However, the decline in prescribing in the year post-diagnosis from 2013 may be an artefact of QOF changes (highlighted in Chapter 2), suggesting that prescribing changes were limited to incident diagnoses and not extant cases that first received an RA code in 2013. The difference between incidence and prevalence was greatest in sensitivity analysis RA2 (Chapter 7), also indicating that improvements in DMARD prescribing targeted incident cases. While rheumatologists may be responsible for additional DMARD prescribing in the year following diagnosis, it is concerning that prescribing by GPs does not increase in the following years when a transfer to GP-led prescribing is expected. With 48.4% not receiving long-term DMARD prescribing 15 years after diagnosis, there may be scope for improved prescribing or increased GP-led management, excepting cases involving biologic DMARD prescribing in rheumatology. Policy, payment and guideline changes should support opportunities for long-term multi-disciplinary review, including of extant RA cases.

Long-term corticosteroid prescribing rates changed little over time and were 14.5 times higher among RA than non-RA patients. Corticosteroid prescribing was higher in sensitivity analyses, consistent with the higher prescribing reported by Black et al. using a more specific RA case definition (485, 524). Once initiated, corticosteroid prescribing was persistent, with little change across the patient life-course after three years and the annual mean prescription count among patients with a corticosteroid prescription remaining high (8.03 in 1998 and 8.02 in 2017). Following a non-significant trend toward increased prescribing in the year post-diagnosis pre-2013, the decline in prescribing post-2013 may be an artefact of incident coding of extant RA cases. Continued prescribing is problematic as it may mask the symptoms of poor disease control - preventing this from being addressed through a treat to target approach - and associates with diabetes, hypertension and cardiovascular disease risk in RA patients (501, 535, 536). It is concerning that older patients, at higher risk of toxicity, had higher

rates of long-term prescribing, especially in the year post-diagnosis. However, there are difficulties in translating guidelines and trial-based evidence into practice regarding corticosteroid reduction, such as adrenal suppression, tendonitis and comorbidities. This is suggested by the rising prescribing rate among matched non-RA patients and stable prescribing in the English population from 2015-2018.

Further effort to taper corticosteroids is required, as a recent study reported favourable outcomes from discontinuing corticosteroids after 34 weeks in early RA (537). An alternative to the guideline recommendation for short-term corticosteroid prescribing, when initiating DMARD therapy (280, 479), may be required to avoid initiation of persistent prescribing. Intramuscular corticosteroids provide a fixed tapered dose and may be useful for short-term use. An efficient treat-to-target approach with DMARDs, and liaison between care providers to address cases of persistent prescribing, should reduce GP prescribing of corticosteroids in new and existing RA patients. This is especially important given the higher prescribing rate, especially in the model including socio-economic deprivation, among older patients who are more susceptible to bone fractures (538) and should aim to address the sex difference in prescribing. Corticosteroid prescribing persisted across GP practices and socio-economic deprivation levels and so wide-scale rather than localised change is indicated. Provision is required for referral and multi-disciplinary review to invoke this. Clinicians may also need to be aware of the unmet need for RA pain control, with pain being an important component in RA disease activity scores (539-541).

Annual long-term NSAID prescribing halved between 1998 and 2017 among RA patients, in contrast to an upward trend among non-RA patients. This followed rising awareness of NSAID toxicity through the early-2000s (488, 542). It was predominantly driven by changing practice for newly diagnosed patients, suggesting the involvement of rheumatology reviews in the improvement of practice. Patients aged ≥60, who are most vulnerable to comorbidity and adverse events, received less NSAID prescribing. However, prescribing varied between GP practices and levels of patient socio-economic deprivation. Further, once initiated, NSAID prescribing seemed to persist as the annual mean prescription count among patients with an NSAID prescription only fell by mean APC: -1.41 and there was little change in the proportion with long-term prescribing over the life-course. With the rate of decline in long-term NSAID prescribing slowing in recent years, further effort is required to promote the review of persistent NSAID prescribing, especially among extant RA cases.

DMARD, corticosteroid and NSAID prescribing trends were similar in the main and sensitivity analyses although the prescribing counts were higher in sensitivity analyses. The sensitivity analyses were designed to exclude suspected but dismissed RA diagnoses and so the higher rates may be more accurate and should be inferred to apply to the prescribing duration calculations also. Prior to 2013, having multiple RA codes may have indicated disease severity and visit frequency as DMARD, corticosteroid and NSAID prescription counts were higher among RA patients after receiving an RA code ≥6 months following a previous RA code compared with all RA patients, with a larger drop in 2013. It also suggests that frequency of RA-related consultations, in which RA codes may be coded, may relate to likelihood of receiving medication and that patients receiving DMARDs are more commonly prescribed corticosteroids and NSAIDs. The difference in prescribing counts is consistent with the finding in Chapter 4 that medication prescribing estimates are higher in EHR studies using more specific disease definitions (485).

Prophylaxis uptake in primary care seems to require reinforcement. Initial increases in bone prophylaxis among women prescribed long-term prednisolone reversed from 2008, with declining bisphosphonate prescribing and vitamin D-calcium co-prescribing plateauing around 50%. Among RA patients in Germany a similar rate of bone protectant prophylaxis prescribing was reported in 2016; 47% (263). The decline in bisphosphonate prescribing follows drug safety announcements and was also observed Canada and in a CPRD study of bone-protectant prescribing after hip fracture (543, 544). While slightly higher co-prescribing was seen alongside high prednisolone doses, even daily doses <5 mg are harmful (545). GI prophylaxis increased more than five-fold and especially in 2005 following the withdrawal of rofecoxib in October 2004, reflecting rising awareness of NSAID toxicity (546). Suh et al. had reported a lower prescribing rate of 10% between 2002 and 2003, suggesting that prophylaxis may be more common among RA patients (547). However, the rate of improvement has slowed and even reversed in 2017. With growing RA prevalence among the elderly that are most susceptible to multi-morbidity and adverse drug reactions (285, 548), renewed efforts to increase bone and GI prophylaxis are crucial, and must target extant as well as incident RA cases.

9.4.1 Strengths and limitations

The methods appraisal is presented in Chapter 10, with chapter-specific strengths and limitations reported here.

The study strengths included the large cohort size, even 15 years post-diagnosis. Sensitivity analyses RA1 and RA2 (requiring additional RA codes or DMARDs) improved the specificity of the definition of RA and helped to confirm study findings. Most of the RA cohort (76.6%, n = 54,685) were eligible for ≥1 sensitivity analysis and 64.3% (n = 35,179) were eligible for both, suggesting reasonable specificity of the RA definition used in assessing the secondary outcomes. Matching to non-disease patients by age and sex is a common approach in studies and was important given the consistent variation in disease incidence shown in this study by these variables (549-551). Matching by GP practice and following up from a matched index date addressed practice-level and temporal variation in diagnostic coding and prescribing practices. It was particularly relevant given that practice-level variation in NSAID prescribing was identified. The GP prescriptions used were recorded automatically in the EHRs of patients contributing data to the CPRD, providing complete information on GP prescribing. All UK GP EHRs record prescriptions automatically, so prescribing trends in other research databases should reflect those reported here.

The assessments made of RA management over the disease course were designed to enable comparison with guideline recommendations and other studies (238, 255, 280, 302). Defining long-term prescribing as ≥90 days in 1 year enabled comparison with prior studies (527) and allowed for cases of shared prescribing with rheumatologists, while defining long-term as ≥180 days enabled conservative estimates that confirmed general trends. The sub-analysis of 'any' prescribing highlighted patterns among any patients receiving medication therapy. Poisson regression was a suitable model selection given the predominantly small counts of annual prescriptions (commonly 0 to 6). The large sample size enabled calendar year to be modelled as a factor to facilitate interpretation. The offset was able to handle the arc in the population size in the CPRD adequately without the need to model calendar year as an integer and squared for the quadratic effect to preserve degrees of freedom. Modelling of IMD and GP practice indicators highlighted where practice- and patient- level factors contribute to prescribing patterns and require consideration in interventions or policy-making. The matched non-RA cohort facilitated in discerning RA-specific prescribing patterns. Excluding non-RA patients with RA diagnosis in the six months following the index date facilitated the comparison of long-term management.

Study limitations include the change in coding practice in 2013 that affected the RA cohort definitions. Prescribing rates pre- and post- 2013 should be compared with caution. Assessment of baseline levels of COPD may also have been affected by coding practices as the coding accuracy improved in 2004 (525) when a COPD-related

QOF payment indicator was introduced. This would not affect comparisons between the RA and Non-RA cohort as each match was based on an index date. Temporal variability metrics can help to quantify temporal variation in coding practice, though sensitivity analyses and understanding of QOF account for the key factors here (552). Although changes in external factors over time could affect the estimates, comparison to non-RA individuals that were matched on GP practice registration minimised any bias (473). Interpretation of the analyses should consider such external factors. Propensity matching on asthma and COPD prevalence would facilitate comparisons of corticosteroid prescribing between RA and non-RA patients, though the rates were low (asthma: 14.1% RA, 11.5% Non-RA; COPD: 6% RA, 3.6% Non-RA) and should not affect comparison of rate change over time. We did not examine change in corticosteroid dosages, which would inform understanding of exposure and medication tapering, however toxicity is increased for all doses (501, 536, 545, 553) and we showed prescribing for 15 years post-diagnosis, beyond the recommended duration for tapering.

GP EHRs do not fully capture patient prescribing as multi-disciplinary prescribing is expected, nor actual medication usage. The relevance of this in an EHR-based study will be influenced by the medication of interest and its likelihood of being prescribed in different settings or being ‘stored’ or ‘conserved’ by patients. Some prescribed medications, including NSAIDs and corticosteroids, are more likely than others such as DMARDs to be ‘saved’ for use during flares or episodes of increased pain. In studies estimating medication exposure, the likelihood of patient adherence to a full prescription course should also be considered based on patient factors as well as the likelihood of adverse events and the condition being studied, as course completion may particularly be an issue in acute conditions. Further, prescribing may differ between settings and countries. Similarly, there may be a temporal lag in RA diagnosis as siloed EHRs introduce a delay between any rheumatology-led diagnoses being coded in GP EHRs. DMARD initiation in secondary care, or secondary care prescribing of intravenous bisphosphonates and denusomab could not be assessed. DMARD prescribing may have continued to rise through biologic availability in secondary care, however these are typically second-line therapeutics and GP DMARD prescribing did not change across the life-course, suggesting that the apparent plateau from 2009 requires investigation. It could not be distinguished where DMARDs were unsuitable or ineffective and long-term corticosteroids or NSAIDs form part of an informed therapeutic approach, however such cases are uncommon (554).

Sub-optimal DMARD prescribing, as indicated by the annual proportion with long-term DMARD prescribing not reaching above 55.1% in the 15 years following diagnosis, will have reduced the sensitivity of sensitivity analysis RA2. DMARD prescribing patterns in sensitivity analysis RA2 cannot be readily compared to those of all RA patients given the involvement of DMARDs in the case definition, though it indicates trends among RA patients prescribed DMARDs.

Other analgesics besides NSAIDs, such as paracetamol, also offer pain relief and may be used in RA management, however these were not studied. There is complexity in using EHR data to assess analgesics in relation to disease management as these are not disease-specific medications. To investigate this more comprehensively would require information on comorbid pain conditions such as back pain, osteoarthritis, shoulder pain and chronic headache. Such conditions may be prescribed analgesics during flares and episodes of pain. The timing of pain flares, and the condition to which a medication prescription should be attributed, are difficult to determine using EHR-based data. The use of analgesics is also under-estimated in EHR data, given their over-the-counter availability, especially among patients who pay for prescriptions (for example, in the UK, adults may pay for prescriptions until aged 60). In this study, the use of all analgesics in RA management was therefore not evaluated. However, in comparison to other analgesics, NSAIDs also reduce joint inflammation and are recommended in UK NICE guidelines for the control of pain or stiffness in RA (280). This may mean that, in patients with RA, NSAIDs (compared with other analgesics) are more specifically prescribed for treating RA than for other comorbid conditions. NSAID prescribing was therefore considered to be more pertinent to investigation in this study of RA management, than other analgesics. The comparison of NSAID prescribing in RA and matched non-RA patients indicated a difference in prescribing which may indeed relate to RA management. However, patients were not matched on other pain conditions such as back pain, which would be necessary to investigate the role of NSAIDs or other analgesics in RA management further.

Intramuscular, intravenous and intra-articular corticosteroids (which may also be given in secondary care) were not assessed as their use is intermittent, meaning the corticosteroid burden may be higher and their use may have changed over time. In addition, the corticosteroids and NSAIDs prescribed may not have been related to RA, although the non-RA cohort assists with understanding RA-related trends.

Using medication terms rather than BNF chapters in defining medications has the downside that the medication terms list would require updating for future studies if new

drugs are developed and assigned to existing chapters. This was deemed necessary given that some chapters incorporated diverse types of medication. Sub-analysis B confirmed the necessity as using only medication terms when linked to specific, RA-relevant chapters yielded very low prescription counts, although the prescribing trends were similar to in the main analysis (Figure 75, Figure 83, Figure 90). Sub-analysis B using medication terms published by a previous study, both with and without BNF constraints, revealed that other studies do not seem to use these BNF constraints as with them the corticosteroid prescription counts were lower than previously reported (524).

Where multiple prescriptions were prescribed on one day, one contributed to the prescription duration calculation per medication. This avoided over-counting from misplaced prescriptions yet is conservative where patients retain prescriptions to issue as required. Unascertainable prescription durations were set at 90 days, which may over-estimate use given the mode prescription duration was 28 days. However this affected <3% prescriptions, findings were similar in secondary analyses of ≥180 days prescribing, and this should not affect interpretation of change over time. The assumption is used in other studies also (555). These long-term prescribing definitions should also allow for unused prescriptions, given the mode prescription duration. Where the prescription duration was >90 days this was replaced as 90 days (504); only 6.2% of the prescriptions studied had a longer duration. The mode DMARD, corticosteroid and NSAID prescription duration remained 28 days after each step of duration calculation was applied. Only prescriptions made during follow-up were assessed; prescribing may seem lower where patients received prescriptions just prior to follow-up although the effect should be minimal given the mode prescription duration.

The annual prescription counts described prescribing based on the date of prescription, which does not describe prescription durations that span concurrent years. However, the prescription duration calculations that were also performed encapsulated the duration in a given year regardless of the initial year of prescription. Prescription count measurements did not consider variations in prescription length although prescription duration calculations did and so these were the focus of the analysis. The importance of this was highlighted when different answers (one month and two months) were given when two clinicians were asked, in relation to this study, about the mean prescription duration for DMARDs, corticosteroids and NSAIDs. Accordingly, in patients receiving >1 prescription, most receive 6 or 12 in a year so these could both be markers of 1 full year of medication use. Assessment of the number of prescriptions per year would be

more meaningful to an investigation of clinician prescribing burden. The duration-based assessment that was performed in this study (proportion of patients with at least 1, 60, 180 days prescribing) is more meaningful for this study of trends in prescribing (any duration) and long-term prescribing, with comparison to guideline recommendations . An alternative duration-based assessment of the length of prescribing per patient, and investigation of gaps in prescribing, would inform understanding of trends in medication adherence. All such measures of prescribing face the same issues discussed above in relation to estimating medication exposure using GP EHR data.

The trends highlighted in prophylaxis co-prescribing merit further investigation. While there is increasing evidence of glucocorticoid toxicity, the national guideline recommendation for bone-protectant co-prescribing is based on the assessed level of patient risk of osteoporotic fracture (302, 535). It is recommended that osteoporotic fracture risk is assessed in women over the age of 64, women aged 50-64 with current or frequent use of corticosteroids, patients aged younger than 50 with current or frequent use of corticosteroids, and patients aged younger than 40 prescribed high-dose oral corticosteroids for ≥3 months. Bisphosphonate co-prescribing alongside corticosteroids is recommended in patients who are assessed to have high risk of osteoporotic fracture. The level of risk was not examined in this study. Therefore, the inclusion of patients at low osteoporotic fracture risk, in whom co-prescribing may not be appropriate, may have affected the interpretation of apparent sub-optimal co-prescribing. Fracture risk could have been predicted based on factors including female sex, age, menopause, hormone replacement therapy, smoking status, body mass index and bone mineral density. However, co-prescribing was assessed only in women, given that in this study the median age at RA diagnosis was 57 (61 in the incident cohort) and the risk of osteoporotic fracture is more likely to be high in older, post-menopausal women. Also, it is not uncommon to have a ‘drug holiday’ in bisphosphonate prescribing, though this might affect a small proportion of patients at any one time (e.g. a one year break after ten years of treatment is recommended for high fracture risk patients) (556). A sensitivity analysis that assessed fracture risk and treatment duration could have been appropriate.

In analyses of co-prescribing and prescription combinations e.g. vitamin D and calcium, prescriptions were not required to be concomitant, simply prescribed in the same year. Vitamin D and calcium prescription numbers and the proportion with long-term prescribing of each were comparable so they were likely co-prescribed. Further, the assessment of 180 days co-prescribing gives a fair indicator of long-term concurrent use.

9.4.2 Conclusion

Despite modern RA treatment paradigms and the expected observed increase in DMARD prescribing, long-term corticosteroid prescribing in the year post-diagnosis has not appreciably reduced. While NSAID prescribing has decreased appropriately with modern understanding of risks and guidance (especially from mid-2000 and among newly diagnosed patients), it remains substantial. Further, the prescribing of both seems to persist once initiated, with implications for toxicity and masking poor RA disease control. Changes in NSAID and DMARD prescribing among RA patients reflect safety concerns and management guidelines but have predominantly been limited to new cases and the momentum has slowed. Review of prescribing among extant cases and across the life-course is required to optimise pharmaceutical management and to transfer to GP-led management once optimal DMARD therapy is reached. Bone and gastrointestinal prophylactic therapy has improved but remains sub-optimal.

With rising prevalence of RA among older patients (Chapter 7), it is increasingly important to address the persistent suboptimal corticosteroid, NSAID and prophylaxis prescribing practices. Persistent corticosteroid exposure, through its anti-inflammatory and immunosuppressive effects, is particularly pertinent for RA patients in the context of infectious respiratory diseases including 2019 novel coronavirus disease (Covid-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and tuberculosis (557-559). This is because respiratory mortality risk is elevated in RA patients and the immunosuppressive effect of the medication may increase the risk of serious respiratory infection (one recent study showed increased odds of hospitalisation with Covid-19 in patients prescribed corticosteroid for rheumatic diseases) (471, 560). Rheumatologists need to understand the causes of persistent prescribing and develop alternative strategies of pain management. Improved primary-secondary care co-management of RA and wider comorbidities, with an informed discussion of co-protectants, could facilitate such alternative strategies. Multi-disciplinary review should particularly be encouraged where corticosteroid prescribing is initiated upon commencement of DMARD therapy, to prevent persistent prescribing.

Chapter 10 Discussion and Future Directions

10.1 Introduction

Musculoskeletal disease (including inflammatory arthritis) is a major cause of disability worldwide and affects 14% of adults in the UK (24, 25). Despite improvements in therapy and the principles of disease management, a large proportion of patients with inflammatory arthritis receive suboptimal care. This thesis evaluated aspects of the epidemiology and / or management of three common inflammatory arthritides (gout, AS, RA), in which timely therapy is critical in maintaining quality of life, through literature reviews and using routine UK EHR data.

The studies in this thesis contribute to knowledge by providing the first systematic review of EHR-based studies in gout pharmacologic management, thematic scoping reviews of EHR-based studies in AS and RA (the latter being UK-specific), and evaluating trends in the epidemiology, diagnosis and pharmacologic management of AS and RA in the UK. The findings and discussion relating to the four thesis objectives were presented in earlier chapters, along with chapter-specific strengths and limitations. Chapters 4-6 reported findings for the first objective, whereby existing EHR-based studies of gout, AS and RA were described. Chapter 7 reported findings for objective two regarding epidemiologic trends, whereby spatio-temporal and demographic patterns in the incidence and prevalence of AS and RA were determined. Chapter 8 reported findings for objective three, regarding trends in the timeliness of diagnosis, by quantifying the time from first back pain symptom and rheumatology referral to diagnosis of AS over two decades. Chapter 9 reported findings for objective four, regarding trends in real-world management, by describing changes in RA medication and prophylaxis prescribing over two decades.

This chapter provides a summary of the key findings for the thesis objectives and a synthesis of these in relation to the underlying hypothesis. The strengths and limitations that were common across the studies using EHR data (Chapters 7-9) are then presented, followed by a discussion of the implications of the work in this thesis and recommendations for the future direction of research. Finally, an overall conclusion to the thesis is given.

10.2 Key Findings

In the context of inflammatory arthritis, this thesis described the existing EHR-based literature and investigated the incidence, prevalence, timeliness of diagnosis and management of disease using EHR data. This thesis provides evidence that EHR data can contribute novel understanding of disease epidemiology and management. The key findings from analyses conducted for the thesis objectives are presented below.

10.2.1 Objective 1: Existing EHR-based Studies

Chapters 4-6 presented literature reviews to describe the existing literature on EHR-based studies in inflammatory arthritis. Chapter 2 had identified a number of studies of the pharmacologic management of gout. The aim of Chapter 4 was to systematically review the methods, results, reporting and risk of bias in literature on EHR-based studies of the pharmacologic management of an inflammatory arthritis, and gout was selected as the exemplar. The diagnostic definitions of gout used, and the investigation of the real-world management of gout, were specific focuses given their relevance to this thesis on using EHR data to understand the epidemiology and management of disease. Following on from this comprehensive assessment of the methodology, reporting and risk of bias in EHR-based studies of pharmacologic management in an inflammatory arthritis, thematic scoping reviews were performed to understand the themes of research conducted in the two other inflammatory arthritides studied in this thesis: AS and RA (Chapters 5, 6). Together, Chapters 4-6 met the objective of describing existing EHR-based studies and showed the contribution made to understanding in a range of themes. Recommendations were made for improved reporting and quality assessment in EHR-based studies.

In Chapter 4, the systematic literature review quantified the rising rate in publication of EHR-based gout medication studies. It was determined that varied case-definitions and medication-related methodology were used in the literature, which affected the ability to evaluate and compare treatment outcomes. Studies that employed more specific diagnostic definitions reported higher rates of medication prescribing and titration. Nevertheless, the existing literature consistently reported that ULT is effective, safe and sub-optimally prescribed in the management of gout. It was concluded that researchers should improve the reporting of methods of EHR data use for the purposes of aiding understanding and reproducibility of research. Areas that were particularly important to improve were the reporting of diagnostic definitions through provision of code-lists, and reporting of data preparation steps and coding validation work, particularly in studies

using less specific diagnostic definitions. There was some evidence of temporal improvement in comprehensiveness of reporting. However, some aspects of CoR remained difficult to assess because researchers rarely had full access to the EHR database source or full understanding of any data processing applied prior to research access. Risk of Bias was generally low; however, the Cochrane items used to assess RoB were not applicable for EHR-based studies in a third of instances. Recommendations were made for adapted CoR and RoB tools for the evaluation of relevant factors that affect EHR-based research and that were commonly missed in the reviewed literature.

In Chapter 5, the study themes examined in AS using EHR data were identified, along with the healthcare settings in which the UK studies were based. The EHR-based research contributed to understanding in five themes, from determining risk factors for AS to reporting on the individual, societal and health economic cost of AS. The reported diagnostic delay and suboptimal prescribing practices require addressing, given the importance of biologics in early RA. Data from GP EHRs was under-used in investigating the diagnosis of AS and screening and management of EAMs and comorbidities.

In Chapter 6, the data sources, methodology, study themes and key findings in UK EHR-based studies of RA were reviewed. Most studies used GP data and high validity was reported for definitions of RA based on this data. Risk factors for RA were reported using IR and IRR along with the adjusted incidence rate and percentage prevalence of RA, although these measures were extended to more recent calendar years in Chapter 7. While in Chapter 4 the impact of different diagnostic definitions on measures of prescribing was reported, in this chapter the impact on incidence rate was described. The increased risk of comorbidity and higher mortality rate in RA patients is of known concern. The urgency of addressing the increased risk of CVD is compounded by the reported finding that CVD risk tools were less accurate in RA than non-RA patients (496). Addressing the suboptimal influenza and pneumococcal vaccine uptake could also help in reducing the rates of respiratory-related deaths (471). Rising DMARD prescribing was found between 1987 and 2010, while there was limited investigation of corticosteroid and NSAID prescribing, which was addressed in Chapter 9.

10.2.2 Objective 2: Epidemiology

For the objective of describing trends in disease epidemiology, the incidence and prevalence in the UK across two decades in AS and RA were investigated in Chapter

7, using GP EHR data. In AS, the incidence and prevalence in the UK were uncertain and in RA these had not been examined in recent calendar years. The incidence of AS was found to be stable in the last decade although a rising number of patients, especially those aged 30-39 and women, received only one code, suggesting improvement in screening. Sensitivity analyses, with differing time windows for confirmation of AS diagnosis, yielded comparable results, suggesting that a >7 day time window was sufficient for diagnosis confirmation in AS GP EHR studies. The pattern of RA incidence was less clear and measurements were affected by an apparent change in coding practice following the introduction of payments in 2013 related to maintaining a registry of RA patients (20). Sensitivity analyses, which required additional diagnostic coding or DMARD prescribing, helped to identify the apparent change in RA coding practice. The prevalence of AS and RA rose in women and men, particularly among older patients. The apparent change in RA coding practice may have improved the sensitivity of RA prevalence estimates over time.

The study finding of rising AS and RA prevalence, presented in Chapter 7, highlighted the growing importance of appropriate disease management in an ageing population. There are important implications in the context of infectious respiratory disease outbreaks, in which the growing population of AS and RA patients may be particularly susceptible, given the use of immunosuppressive agents in disease management and the higher respiratory mortality rate in RA patients (471). This study met the objective of describing epidemiologic trends and highlighted how measures to improve coding quality can affect secondary use of data in research.

10.2.3 Objective 3: Timeliness of Diagnosis

The objective of describing trends in the timeliness of diagnosis, given its importance for the success of early pharmacologic therapy, was investigated in Chapter 8 using population-based UK GP EHR data. The study investigated the time from first recorded back pain symptoms to rheumatology referral and diagnosis of AS, in which significant diagnostic delay had been previously reported from hospital and survey-based studies. Symptoms are often first presented in primary care and this study reported a trend to worsening in delay to rheumatology referral and AS diagnosis over two decades. The longer delay in women, even after rheumatology referral, with no apparent reduction in the sex difference over time, is of particular concern. Sensitivity analyses, requiring ≥2 and ≥3 years of quality (UTS) GP registration prior to AS diagnosis, suggested even greater diagnostic delay. This study met the objective of describing trends in the time to diagnosis and highlighted how EHR data quality affects estimates.

The study presented in Chapter 8 suggested that one approach to reduce diagnostic delay would be to aid the recognition and referral of IBP among non-rheumatologists, given the apparent contribution of delayed rheumatology referral to the time to diagnosis. A UK study written in 2007 had reported on the existence among GPs of a range of approaches to the diagnosis of AS (517), and as the time to rheumatology referral and diagnosis in this study did not improve in subsequent years, that finding potentially remained hugely relevant across the study period. That study had reported that 17% of GPs identified <1/2 of the features of IBP (517), which would contribute to diagnostic delay. The need for improved education and prompts for early referral are unlikely to be UK-specific. A study of healthcare professionals in the USA reported that less than one-third recognised that pain getting better with activity was suggestive of IBP, and once suspected, 39.6% opted to treat the patient themselves (through physical therapy in 81.4% of cases), with only 13.0% immediately referring patients to another specialist (520). In the UK, revised national guidelines published subsequently in 2019 have highlighted the range of symptoms and test results that may raise suspicion of AS, and prompt early rheumatology referral (223). The impact of such revised guidance on diagnostic delay in the forthcoming years could be investigated in a future study using the methodology described in Chapter 8.

10.2.4 Objective 4: Real-world Management

For the objective of describing trends in management, the patterns of RA therapy and prophylaxis prescribing were assessed using UK GP data in Chapter 9. Despite the shift toward DMARD prescribing, the long-term prescribing of corticosteroids has persisted in the year post-diagnosis. The anti-inflammatory and immunosuppressive effects of persistent exposure are of concern, particularly in the context of infectious diseases (557). Following growing evidence of NSAID toxicity, the proportion with long-term (≥90 days) prescribing in the year post-diagnosis declined but remained substantial (27% in 2017). Persistent corticosteroid and NSAID prescribing across the patient life-course was reported, which has implications for masking poor RA disease control as well as risk of toxicity, and suggest an unmet need for pain control. While bone and gastrointestinal prophylactic therapy improved, it remained sub-optimal, which is increasingly important to address given the rising RA prevalence in older patients (as identified in Chapter 7) who are more commonly comorbid. Further, the rise in DMARD prescribing slowed and was predominantly limited to new cases, suggesting the importance of medication review across the life-course. If optimal DMARD therapy is reached in a rheumatology clinic, pharmacologic management should transfer to GPs and yet GP DMARD prescribing did not increase across the life-

course. This study met the fourth objective of the thesis in describing real-world disease management.

The study presented in Chapter 9 suggested that rheumatologists should review corticosteroid prescribing where it is initiated upon commencement of DMARD therapy, to address any issues in tapering and withdrawal. The study suggested the need for improved primary-secondary care co-management of RA, pain and wider comorbidities, and the importance of having an informed discussion with patients regarding co-protectants.

10.2.5 Hypothesis: Improving Understanding of Disease Epidemiology and Management using EHR data

The studies in this thesis explored the hypothesis that EHR-based research can provide information on disease epidemiology and management, relevant to clinicians and decision-makers. The studies reviewed the existing EHR-based studies of disease epidemiology and management in inflammatory arthritis and met the need for information on the incidence and prevalence, timeliness of diagnosis and pharmacologic management in inflammatory arthritis, using EHR data.

This thesis provides information required by clinicians and decision-makers to improve patient management:

- The systematic review of EHR-based studies in gout, and scoping reviews of EHR-based studies in AS and RA, evaluated current understanding and synthesised the results of studies to contribute a review of the real-world evidence.
- The investigation of AS incidence and prevalence addressed the unmet need for this, which was highlighted in a national review as being a limiting factor for understanding of the disease burden (190). The rising proportion of patients with a single, but not ≥2, instances of AS code recording may suggest improvements in data recording or else the impact of efforts to raise awareness of AS, which could inform further efforts to improve the timeliness of specialist referral for AS and other diseases. The study provided information on the spatial and demographic trends in AS, showing the growing prevalence in women and patients aged ≥60, which informs efforts to estimate the health and economic impact of an ageing population.

- The examination of RA incidence and prevalence updated the understanding of these following an apparent change in coding practice in recent years and identified growing prevalence in patients aged ≥70, informing efforts to evaluate and manage an ageing population. By identifying changing coding practices, the study also informed evaluations of other GP EHR RA studies across the period of change and informed future efforts to improve data quality by suggesting the impact of payment incentives on diagnostic coding practice.
- The investigation of time to diagnosis in AS highlighted persistent delay in rheumatology referral and diagnosis, informing policymakers of the need to facilitate pathways for early rheumatology referral of IBP cases in order to improve the prescribing of DMARDs in early AS and in turn reduce the personal and societal cost of AS.
- The evaluation of RA medication and prophylaxis prescribing showed the stalling of improvement in DMARD and prophylaxis prescribing and in reduction of NSAID prescribing, and persistent corticosteroid prescribing across the life-course. It also identified where there is variation with age, sex and socio-economic status. This study highlighted the need for efforts to optimise prescribing and prophylaxis, and suggested the importance of incentivising medication and pain review in the extant RA population. The study also suggested the impact of evidence and guidelines, as NSAID prescribing seemed to decline following reports of toxicity while corticosteroid and NSAID prophylaxis initially increased following the publication of national guidelines. Such understanding of the impact of guidelines may inform further efforts to educate healthcare professionals and incentivise the use of evidence-based care pathways.

This thesis also raises considerations for future EHR-based studies on using EHR data in research:

- The systematic review in gout management (Chapter 4) and the investigation of RA management (Chapter 9) both reported on the impact of the specificity of the diagnostic definition on prescribing estimates, while Chapters 6 and 7 reported on the impact of diagnostic definition specificity on incidence estimates. These chapters raised the importance of considering the diagnostic specificity in EHR-based studies.
- In Chapter 4, areas were identified, and considerations were raised, for further improving the reporting and quality assessment of EHR-based studies of gout management, which apply to all EHR-based studies.

- The review of research themes in existing EHR-based studies in AS and RA (Chapters 5, 6) informs understanding of the ways in which EHR data can be used to derive real-world evidence of any disease.
- The investigation of AS incidence and prevalence (Chapter 7) identified change in the proportion of patients with a single AS code, suggesting the importance of conducting a sensitivity analysis of confirmed diagnoses in order to distinguish practice changes that might affect the sensitivity or specificity of diagnostic definitions across the study period. The sensitivity analyses, which required additional confirmatory coding within >7 and ≥180 day time windows, also suggested that a >7 day window suffices for assessing confirmed AS in GP EHR data.
- The study of RA incidence and prevalence (Chapter 7) identified an apparent change in RA coding practice and highlighted the importance of considering such changes across a study period that affect data recording and in turn the sensitivity or specificity of study variables. The study showed that sensitivity analyses offer an approach to identifying any such change in EHR data recording.
- In the examination of time to diagnosis in AS (Chapter 8), the greater diagnostic delay in sensitivity analyses with ≥2 and ≥3 UTS years prior to diagnosis, showed the importance of accounting for data quality in EHR-based studies, and considering the impact of follow-up duration in longitudinal analysis.
- Chapter 6 highlighted the lack of comprehensive reporting of medication definitions in RA EHR studies, and the investigation of prescribing in RA with and without BNF chapter constraints in Chapter 9 suggested that these are not applied in existing EHR-based studies. This finding, in corroboration with the systematic review in gout management (Chapter 4), raised a need for more comprehensive reporting of EHR data handling – potentially through supplementary material.

The thesis therefore addressed the underlying hypothesis and presented evidence suggesting that EHR data can be used to develop understanding on disease epidemiology and management, which informs clinicians and decision-makers.

10.3 Strengths and Limitations

Many of the strengths and limitations were discussed in the relevant chapters; those common across the EHR data studies (Chapters 7-9) are discussed here.

A strength across the studies was the comprehensive reporting of EHR data handling, following the RECORD statement (345). The study methodologies were described, including the definition of population denominators, and code-lists were provided, to aid reproducibility and interpretation of the studies.

The studies were strengthened by the selected dataset and cohort selection criteria:

- The dataset was GP practice EHR data representative of the UK, which suited the study objectives of reporting national incidence and prevalence estimates and the diagnosis and management of AS and RA. In the UK, these inflammatory arthritides are commonly managed between primary and secondary care (shared care).
- The comprehensive recording of prescriptions in UK GP EHRs, and increasing recording of back-pain symptoms (Chapter 8), RA diagnoses (Chapter 7, following payment incentives introduced in 2013 (20)) and rheumatology referrals (Chapter 8) over time as identified have improved the data quality in this thesis. The apparent change in RA diagnostic coding following the introduction of payment incentives, and the greater diagnostic delay evident in AS patient records with a longer duration of quality data (UTS) follow-up, highlight the importance of considering data quality across the study period in any EHR-based study.
- Data extraction occurred 4 months post-study follow-up to allow for retrospective coding and GP practice data upload.
- The large population-based cohort size was selected from a representative population, inclusion of patients irrespective of disease severity, and excluding low quality records as is appropriate for observational studies (89, 95, 96).
- Patient exclusion criteria were minimised in order to maintain external validity.
- Patients required a year of quality registration prior to follow-up, enabling discernment of incident diagnoses from any prevalent cases recorded as incident diagnosis instead of medical history during GP registration. This duration is sufficient for chronic diseases such as AS and RA, and a prior study of RA incidence reported no significant difference when either one or three years of prior registration were required (261, 396).

The studies were strengthened through sensitivity analyses:

- These were performed to improve the specificity of the definitions of AS and RA, excluding patients with suspected but subsequently dismissed diagnoses, as inferred from a lack of confirmatory diagnosis or subsequent DMARD prescribing.

- These analyses ended in 2016, providing >16 months of follow-up for the additional coding and prescribing to occur. This also allowed time for AS or RA diagnosed by rheumatologists to be coded in GP EHRs, and for GPs to become involved in DMARD prescribing after it is initiated in secondary care.
- Tate et al. (2017) reported higher diagnostic coding variation in “poor quality” records (510), which these sensitivity analyses may have helped to address.
- The sensitivity of the diagnostic definition in sensitivity analysis RA2, using prescribed DMARDs, was enhanced by the comprehensive recording of prescriptions in UK GP EHRs, which ensures that all GP-prescribed DMARDs will have been identified.
- Comparison between the primary analyses and these sensitivity analyses informed the consideration of potential causes of reported changes across the study period. For example, the rise in RA incidence pre-2013 (Chapter 7) was steepest in sensitivity analysis RA2 (compared with the primary analysis and sensitivity analysis RA1), when in Chapter 9 DMARD prescribing was shown to have increased. Further, sensitivity analysis RA2 showed the least change in incidence post-2013 when, as shown in Chapter 9, DMARD prescribing was unaffected by payment incentives (unlike RA coding).
- These sensitivity analyses confirmed the study findings as showed consistent results.

There were limitations in the study dataset, cohort selection criteria and study definitions:

- While data was utilised from a representative sample of UK GP practices, the epidemiology, diagnosis and management of disease may differ between settings and countries. However, the finding of rising AS and RA prevalence in older patients may be generalisable to countries with similar care pathways and population demographics. The diagnostic delay in AS is not UK-specific and the suggested need to support the recognition of IBP features and prompt early rheumatology referral may apply to other countries including the USA and India (520, 561). The finding of persistent corticosteroid prescribing in RA patients has been also recently been shown in the USA across the first year post-diagnosis, with 29.2% of patients prescribed corticosteroids 10-12 months-post diagnosis .
- In the UK, while GPs play an important role in the management of AS and RA, aspects of care performed by secondary care may have been missed in the GP EHR data. Diagnoses may be made in secondary care and DMARD initiation, intra-muscular and intra-articular corticosteroid injections and biologic prescribing is predominantly led by secondary care. The potential impact of this

on the analyses of time to diagnosis (Chapter 8) and pharmacologic prescribing (sensitivity analysis RA2 in Chapter 7, and Chapter 9) are discussed in the strengths and limitations sections of Chapters 7-9.

- There are aspects of physician, patient and carer decision-making that are not recorded using EHRs. This can limit interpretations regarding disease trajectories, medication effectiveness and patient management. The comparison of prescribing trends with national guidelines on gout and RA management (Chapters 4, 9) was constrained by not assessing all of the important considerations that could have influenced prescribing. Interpretations of the appropriateness of prescribing may be confounded by unmeasured factors that include disease severity, presence or absence of risk factors, and a patient's psychological needs, wishes, frailty, comorbidities and concomitant medication. However, general patterns in prescribing can be described using EHR data, highlighting areas for further investigation. Further, the application of data mining and artificial intelligence approaches may aid in identifying some of the confounding factors; for example, algorithms may predict frailty, disease severity, disease activity and patient risk of disease (562-565).
- EHR systems may shape coding practices, with some interfaces designed to facilitate clinical coding through predictive or automated means, so that estimates derived from GP Practices using different EHR systems than those contributing to the CPRD may vary (66). However, a recent study shows that antibiotic prescribing and infections rates in the CPRD GOLD are comparable to that recorded by GP practices using Emis, the main EHR system supplier in the UK, suggesting that there are similarities in recording practice (566).
- The declining coverage of CPRD GOLD from 2013, with regional variation, was identified in Chapter 7. This may be attributable to GP system migration and increased awareness of public concern regarding secondary use of EHR data which was expressed in regard to the UK care.data programme (567). This may affect external validity, although the proportion of patients remained significant and there was no notable change in study outcomes in 2013 beyond those attributable to RA coding incentives.
- In investigations of regional and demographic variation, patient exclusions based on regional and demographic factors were applied (e.g. patients aged >99). However these patient exclusions were not applied in other analyses.
- Medications were identified based on drug terms rather than by BNF chapters, meaning that the drug lists would need manually updating in future studies as new drugs are developed and added to existing chapters. However this ensured that non-relevant drugs in the same chapters were not included and relevant medication were identified regardless of their BNF chapter. In Chapter 9 it was

determined that chapter constraints substantially reduce prescription counts below those reported in other studies.

The study limitations included those common to EHR-based studies where external factors (e.g. local practice, EHR system changes, national policy, coding systems) shape information recording (485, 510).

- Besides medication prescriptions and demographics, the recording of aspects of health and healthcare delivery may be incomplete or not systematic, and the reasons for this are not always known to secondary data users. Improvements in routine data collection and data quality standardisation would facilitate the secondary use of data in deriving real-world evidence. However, this is difficult to balance with enabling flexibility in data recording based on differing clinical or professional approaches and patient needs. Ultimately, while facilitating research remains a secondary use of EHRs, measures to consider data quality (such as sensitivity analyses) will be a necessity in EHR-based research. In data quality considerations, the diversity of research goals constrains efforts to standardise methodologic approaches; consequently the comprehensive reporting of data handling methodology is crucial for transparency in EHR-based research.
- The diagnostic codes used in defining AS and RA may have been missing, or recorded after a temporal lag, in GP EHRs where the diagnosis is made in secondary care. Delayed coding is particularly relevant to studying incidence and prescribing in the year post-diagnosis, although the reported epidemiology and management of disease is reflective of the GP perspective.
- Study outcomes in RA pre- and post- 2013 should be compared with caution due to changes in coding practices affecting cohort comparability. The introduction of payment incentives may have increased the sensitivity of the RA diagnostic definition as more RA patients received a diagnostic code. The recording of ≥2 RA codes may previously have been an indicator of severe disease cases with frequent visits in which the additional coding could occur, meaning that the cohort in sensitivity analysis RA1 may have represented more severe cases pre-2013 compared with post-2013.
- Rising DMARD prescribing may have increased the sensitivity of the RA definition across the study period in sensitivity analysis RA2 (where prescribed DMARD was required).
- Sensitivity analysis RA2 might have excluded patients receiving biologic DMARDs in secondary care, however such treatment tends to be prescribed following or alongside non-biologic DMARDs and only in severe cases (568).

Similarly, should patients with RA not receive DMARDs, this would reduce the sensitivity of sensitivity analysis RA2.

Processing and data selection by research database owners also shaped the data analysed, affecting the real-world representation and external validity:

- The CPRD provide training in coding to GP practices and review the data to improve data quality for such epidemiological studies. This may introduce variation compared with other databases, and mean that caution should be applied in interpreting, based on these studies, the coding practice of GP practices that do not contribute to the CPRD. The analysis of prescribing data should not be affected however, as automation of prescription coding is widespread.
- The CPRD does not provide historic GP practice registration details so that where patients migrate between GP practices that contribute to the CPRD, only their current registration contributes to the calculation of the duration of their UTS registration. This could exclude relevant patients with the requisite ≥1 year prior UTS registration from follow-up where they have migrated between UTS GP practices, until they had 1 year of UTS registration in the current practice.

10.4 Future Directions

There are opportunities for research that build directly on the studies in this thesis:

- The considerations raised in Chapter 4, for further improving CoR and RoB assessment in EHR-based studies, could inform the development of revised CoR and RoB checklists and tools.
- Suggestions for further research in RA were made in Chapter 6, based on gaps in the existing EHR-based literature.
- The increasing prevalence in AS and RA among older patients (Chapter 7) suggests the need for research into survival trends, especially given the high mortality in AS described in Chapter 2. As cardiovascular mortality is the main contributor to mortality in AS patients, any changes in prescribing of cardiovascular prophylactics and inflammation-reducing anti-TNF medication could also be investigated (177, 569, 570). Linked mortality data from the ONS would facilitate this investigation (105).
- In future years, an updated investigation of the time to diagnosis in AS (Chapter 8) could examine the impact of national guidelines published in 2019 that prompt early referral in AS (223).

- The factors contributing to persistent suboptimal prescribing of medication and prophylaxis for gout and RA in the real-world setting (Chapters 4, 9) require investigation in order to identify appropriate areas for intervention. Variation in NSAID prescribing by GP practice and socio-economic deprivation level was identified in Chapter 9, suggesting that these may be of consideration in any future interventions targeting NSAID prescribing. As corticosteroid and NSAID prescribing was lower among DMARD-naïve RA patients, further study could investigate the impact of DMARD prescribing on long-term corticosteroid and NSAID prescribing.
- Investigation into the comparison of prophylaxis co-prescribing between RA and non-RA patients, between patients with high- and low- comorbidity risk, and between prescribing by rheumatologists and GPs in the UK, could be informative. Higher prophylaxis co-prescribing by rheumatologists has been reported in Canada (301).
- The impact of the RA medication prophylaxis prescribing changes, identified in Chapter 9, on health outcomes and comorbidity and mortality, could be investigated.
- EHR data would facilitate investigation of regional and demographic variation in prescribing and adherence also.

Building on the theme of investigating the epidemiology and management of inflammatory arthritis, a next step would be to explore the substantial comorbidities in AS by using EHR data. The common EAMs including uveitis, psoriasis and IBD, and comorbidities including hypertension, CVD, osteoporosis and depression, were described in Chapter 2. A descriptive study using the GP EHR data described in this thesis could investigate trends in the incidence and prevalence of these over time, exploring the pattern of occurrence in terms of the order of manifestation, in relation to the diagnosis of AS and other comorbidities. Investigation of trends in the rate of comorbidity accumulation could suggest whether this has declined following improved biologic uptake in early AS, e.g. after the BSR publication in 2005 as described in Chapter 7 (240). Trends in comorbidity screening rates could also be investigated, with stratification by patient risk. The timing of EAM diagnoses in relation to first recorded back pain symptoms could suggest approaches to reduce diagnostic delay. In this study, BMI and socio-economic deprivation would be important factors to account for, and could be calculated from information that is commonly recorded in EHRs. In mapping the trajectory of comorbidities, the application of process mining techniques may be useful (571).

The strengths of EHR data identified in this thesis also informs the direction of EHR-based research:

- This thesis identified that substantial societal and health system costs (Chapter 5) and suboptimal pharmacologic prescribing (Chapters 4, 9) in inflammatory arthritis can be identified using EHR data. Consequently, EHR data would inform a health economics analysis of the costs associated with sub-optimal pharmacologic management in inflammatory arthritis, compared with guideline-recommended practices. These costs could include medication, prescription exemption status due to unemployment, consultations and health service utilisation, surgery and hospitalisations. In a linked data analysis, these costs could include absenteeism and work disability.
- This thesis also suggested that the impact of management guideline publications can be investigated using EHR data. The apparent increase in suspected AS from 2007 (Chapter 7) suggests that heightened disease awareness may be a positive impact of the publication of management guidelines (240).
- This thesis identified methodologic considerations for future EHR-based research. The impact of diagnostic definition specificity was raised. The RA studies highlighted the importance of incorporating diagnostic codes whose use is incentivised in diagnostic definitions. Future studies should consider the code-lists and EHR phenotyping algorithms shared or validated in this and other previous studies, making use of code-list repositories and open-source platforms (572, 573).

The unprecedented growth in routine data collection, computing and processing power, and advances in artificial intelligence techniques (e.g. machine learning, natural language processing and deep learning), was described in Chapter 2. There is increasing availability of a growing breadth of wearable devices and mobile applications generating health and fitness data, such as electrodes in a smartwatch than can detect heart rate variability and pills that emit a signal to a wearable sensor patch when ingested (574, 575). Many EHR systems including Epic and SystmOne enable the integration of such apps with EHRs, and automated workflows are starting to be developed whereby patient-generated readings outside of specified parameters can trigger notifications or symptom profiles, which inform care pathways and prescription adjustments (576, 577). This growth in data, methodology and app integration is uncovering untapped opportunities for EHR-based research:

- Research that integrates such patient-generated data in EHRs and facilitates remote monitoring would support timely care interventions.

- As devices are increasingly networked with sensors in everyday objects, the emergence of the ‘Internet of Things’ facilitates research in tracking the environmental influences on disease trajectories and the impact of physical activity, including falls (578).
- Data-driven prediction and forecasting can inform national-level risk assessment and the personalisation of care pathways that even factor in multi-morbidity. Machine learning can help researchers to process the otherwise overwhelming extent and variety of data that is becoming available, by identifying patterns that would not otherwise be discernible and informing the generation of novel hypotheses (579). Applications that engage natural language processing and cognitive computing, such as IBM Watson, can help to unlock the utility of clinical notes and letters, diarised entries in health apps, and social media data (580, 581). Deep learning is exposing opportunities for research, combining imaging data with contextual information from EHRs, to improve the speed and accuracy of diagnosis and clinical decision making; one such study, through clinical workflow integration, reduced the time to diagnosis of intracranial haemorrhage in routine practice by 96% (582, 583).

The opportunities for active interventions triggered by digital signals are manifold, from early diagnosis of depression informed by voice analysis, through chronic asthma management based on tracking the frequency and location of inhaler use, to patient education and medication adherence reminders, digital cognitive behavioural therapy for insomnia, and triggering emergency response for myocardial infarctions (584-587).

There are challenges to overcome regarding this growth in data. These include ensuring that the integration of data into clinical workflow is meaningful, that devices are validated and regulated, that app accessibility does not exacerbate health inequalities, as well as addressing the high drop-off rate that is common in activity tracking device usage (588, 589). This is in addition to ethical concerns that will become increasingly vociferous as data capture becomes pervasive. These include concerns regarding privacy, autonomy, and ensuring appropriate health literacy and informed consent. There are issues of public trust and transparency in data use, especially when using commercial applications such as IBM Watson and Google DeepMind in research (580, 590). The development in recent years, of regulations regarding the use of personal data (e.g. GDPR (77)), are helping to provide clarity on the legislative stance on appropriate data handling methodologies, as well as a formalised language for raising privacy concerns. Approaches such as the Observational Medical Outcomes Partnership Common Data Model, to design applications and EHRs using a common format from which summary statistics can be derived without sharing the raw data (591), may facilitate the availability of such data

for research while addressing some of these concerns. The current response to SARS-CoV-2 has, in some instances, seen a focused 'opening up' of EHR data which has facilitated an explosion of Covid-19-related research: this may be influential in shaping the evolution of the mechanisms, principles and regulations for post-pandemic EHR-based research (592-595).

In order to fully realise the benefits of EHR-based research, a future direction must be the embedding of EHR-based research in EHRs. In this way, EHRs can support clinicians in translating the results of EHR-based studies, especially those that supplement randomised clinical trials, into personalised clinical decision-making based on an individual patient's risk profile (596). This is an important step in a learning healthcare system and offers exciting opportunities for research (70). Future research could incorporate the findings described in this thesis into EHRs:

- The analyses conducted to define diagnostic delay in AS (Chapter 8), and suboptimal prescribing and prophylaxis in gout and RA (Chapters 4, 9) could inform report templates to be run on EHR systems by health organisations or clinical commissioning groups to identify local training needs. Such reports could similarly be developed to evaluate comorbidity screening if future research indicate areas of suboptimal practice.
- As discussed in Chapter 9, when initiating DMARDs in RA there are a number of corticosteroid tapering strategies. A future direction to compare these could be to identify and enrol RA patients upon DMARD initiation through an EHR notification, with assignment to one of the approaches being guided and recorded in the EHR, with subsequent follow-up in the EHRs informing research.

The embedding of research into EHRs may help to drive future efforts to improve EHR data quality, both at the ground level and through policy changes. In addition, the future integration of genomic biomarker data in EHRs would facilitate the embedding of multi-omic research in EHRs, and improve EHR data quality where genetic variants can inform imputation of missing data (597). Embedded analytics could provide beneficial information to EHR users in cases where the relevant data has been recorded and thus encourage improved data recording. Informed by real-world evidence from EHR-based studies, policy-makers could incentivise data recording in order to realise effective healthcare through the recording of factors that are inputs for research-designed clinical decision support tools.

This thesis also highlights that a future direction in research is to address the issues in EHR-based research:

- As described in Chapter 2, assessing and addressing issues of data quality is a complex undertaking (71, 598) and methodologies to assess and account for this require further research.
- Standardised approaches to report on data quality and data handling are required, which can be informed by the considerations raised in Chapter 4. This is important given the finding in Chapters 4-9 that estimations may be affected by the specificity of the study variables used in EHR-based research.
- The studies in Chapters 7-9 used GP data and were limited by the potential of missing diagnoses or prescriptions made in secondary care. The extent to which data flow, both between the settings that deliver healthcare, and to research databases, is constrained, will continue to hinder efforts for whole systems research across all health settings at a representative, population-based level (48, 599). The linkage of EHR databases on international scales would similarly facilitate research, particularly in less common diseases or latent adverse events. As well as addressing the technical issues in linkage, a societal consensus on the secondary use of health data and data protection, and trust in secure infrastructure, will be required (600). This consensus would likely need to evolve as data collection and linkage becomes more pervasive and methodologic capability advances to unlock unforeseen potential.
- The study of RA management was limited by uncertainty in whether prescriptions were made in the course of RA management. EHR functionality exists to attribute medications to a symptom or diagnosis upon prescription yet this is not widely used - if this practice was adopted it could assist future research. Propensity score matching methods could also aid future such studies.

Contributions to clinical practice made by EHR-based research may, over time, encourage the development of solutions to many of the current issues in the research domain. In the meantime, interpretation of all EHR-based studies should be contextualised through understanding of the primary purpose of EHRs and the data provenance. Collaboration between EHR users, academia, regulatory bodies, patients and policy-makers will be essential in addressing the ongoing issues in EHR-based research.

10.5 Conclusions

This thesis appraised existing research and contributed to the understanding of the EHR-based studies conducted in the domain of inflammatory arthritis. A systematic literature review and thematic scoping reviews identified the range of data sources and themes investigated in EHR-based research, quantified the impact of the specificity of diagnostic definitions on research outcomes, and raised considerations for improving for the reporting of EHR data handling and assessment of risk of bias in EHR-based studies. In doing so, the reviews summarised EHR-based research on epidemiology and management in inflammatory arthritides, and provided information for the analysis and reporting of future such studies.

The UK GP EHR data studies in this thesis provided novel information on the incidence and prevalence of AS and RA, and evidence of diagnostic delay in AS and prescribing patterns in RA, over the last two decades. Initial declines in AS incidence have stabilised in recent years and the rising prevalence of AS and RA in the older population raises important implications for the health burden of an ageing population. The persisting, and trend to worsening, delay in rheumatology referral and diagnosis of AS, especially in women, suggest the need for education on IBP in non-rheumatologists, if the benefits offered by DMARD prescribing in early AS are to be realised. While modern strategies have improved DMARD prescribing in RA, only through addressing the substantial long-term prescribing of corticosteroids and NSAIDs, and improving prophylaxis prescribing, can appropriate pharmacologic management of RA be attained. This evidence of epidemiology and management was derived from a dataset representative of the UK population and is relevant to clinicians and decision-makers.

This thesis concludes that the presented evidence, from a systematic review, thematic reviews and EHR data studies, supports the hypothesis that important and critical information on disease epidemiology and management can be derived using EHR data.

References

1. Dinov, I.D. Volume and value of big healthcare data. *Journal of Medical Statistics and Informatics.* 2016, **4**.
2. Reinsel, D., Gantz, J. and Rydning, J. *Data Age 2025: The Evolution of Data to Life-Critical (IDC White Paper)*. Seagate. 2017.
3. Cottle, M., Hoover, W., Kanwal, S., Kohn, M., Strome, T. et al. Transforming Health Care Through Big Data Strategies for leveraging big data in the health care industry. *Institute for Health Technology Transformation.* 2013.
4. Arndt, B.G., Beasley, J.W., Watkinson, M.D., Temte, J.L., Tuan, W.-J. et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. 2017, **15**(5), pp.419-426.
5. US 111th Congress. *The American Recovery and Reinvestment Act of 2009.* 2009.
6. Her Majesty's Treasury. *The plan for growth.* London, UK, 2011.
7. Raghupathi, W. and Raghupathi, V. Big data analytics in healthcare: promise and potential. *Health Information Science and Systems.* 2014, **2**(1), pp.1-10.
8. Ngiam, K.Y. and Khor, W. Big data and machine learning algorithms for health-care delivery. *The Lancet Oncology.* 2019, **20**(5), pp.e262-e273.
9. Huang, J.Y.J.E. Representativeness is not representative: addressing major inferential threats in the UK Biobank and other big data repositories. 2021, **32**(2), pp.189-193.
10. Sherman, R.E., Anderson, S.A., Dal Pan, G.J., Gray, G.W., Gross, T. et al. Real-World Evidence-What Is It and What Can It Tell Us? *New England Journal of Medicine.* 2016, **375**(23), pp.2293-2297.
11. Bollier, D. and Firestone, C.M. *The promise and peril of big data.* Aspen Institute, Communications and Society Program Washington, DC, 2010.
12. Katkade, V.B., Sanders, K.N. and Zou, K.H. Real world data: an opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. *Journal of Multidisciplinary Healthcare.* 2018, **11**, p.295.
13. Rotar, A.M., Van Den Berg, M.J., Schäfer, W., Kringos, D.S. and Klazinga, N.S. Shared decision making between patient and GP about referrals from primary care: Does gatekeeping make a difference? *PLoS One.* 2018, **13**(6), p.e0198729.
14. Johnson, O.A., Fraser, H.S.F., Wyatt, J.C. and Walley, J.D. Electronic health records in the UK and USA. *The Lancet.* 2014, **384**(9947), p.954.
15. Hetland, M.L. DANBIO—powerful research database and electronic patient record. *Rheumatology (Oxford).* 2011, **50**(1), pp.69-77.
16. Benson, T. Why general practitioners use computers and hospital doctors do not—Part 1: incentives. *BMJ.* 2002, **325**(7372), pp.1086-1089.
17. McMillan, B., Eastham, R., Brown, B., Fitton, R. and Dickinson, D. Primary Care Patient Records in the United Kingdom: Past, Present, and Future Research Priorities. *Journal of Medical Internet Research.* 2018, **20**(12), p.e11293.
18. Wallace, P., Delaney, B. and Sullivan, F. Unlocking the research potential of the GP electronic care record. *British Journal of General Practice.* 2013.
19. McLintock, K., Russell, A.M., Alderson, S.L., West, R., House, A. et al. The effects of financial incentives for case finding for depression in patients with diabetes and coronary heart disease: interrupted time series analysis. *BMJ Open.* 2014, **4**(8), p.e005178.
20. NHS Digital. *Quality and Outcomes Framework - 2012-13.* [Online]. 2012. [Accessed 10 February 2020]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-data/quality-and-outcomes-framework-2012-13

21. National Institute for Health and Care Excellence. *Improving health and social care through evidence-based guidance.* [Online]. 2021. [Accessed 05 February 2021]. Available from: https://www.nice.org.uk/
22. Strong, K., Mathers, C., Leeder, S. and Beaglehole, R. Preventing chronic diseases: how many lives can we save? *The Lancet.* 2005, **366**(9496), pp.1578-1582.
23. Parker, L., Moran, G.M., Roberts, L.M., Calvert, M. and McCahon, D. The burden of common chronic disease on health-related quality of life in an elderly community-dwelling population in the UK. *Family Practice.* 2014, **31**(5), pp.557-563.
24. Vos, T., Flaxman, A.D., Naghavi, M., Lozano, R., Michaud, C. et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet.* 2012, **380**(9859), pp.2163-2196.
25. Office for National Statistics. *General health (General Lifestyle Survey Overview - a report on the 2011 General Lifestyle Survey).* London, UK, 2013.
26. Grassi, W., De Angelis, R., Lamanna, G. and Cervini, C. The clinical features of rheumatoid arthritis. *European Journal of Radiology.* 1998, **27**, pp.S18-S24.
27. Singh, J.A., Reddy, S.G. and Kundukulam, J. Risk factors for gout and prevention: a systematic review of the literature. *Curr Opin Rheumatol.* 2011, **23**(2), pp.192-202.
28. Siegel, R.J., Bridges Jr, S.L. and Ahmed, S. HLA-C: An Accomplice in Rheumatic Diseases. *ACR Open Rheumatology.* 2019, **1**(9), pp.571-579.
29. Combe, B., Landewe, R., Daien, C.I., Hua, C., Aletaha, D. et al. 2016 update of the EULAR recommendations for the management of early arthritis. *Annals of the Rheumatic Diseases.* 2017, **76**(6), pp.948-959.
30. Feldtkeller, E., Khan, M., van der Heijde, D., van der Linden, S. and Braun, J. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. *Rheumatology International.* 2003, **23**(2), pp.61-66.
31. van der Heijde, D., Ramiro, S., Landewé, R., Baraliakos, X., Van den Bosch, F. et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Annals of the Rheumatic Diseases.* 2017, **76**(6), pp.978-991.
32. Smolen, J.S., Landewé, R., Bijlsma, J., Burmester, G., Chatzidionysiou, K. et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Annals of the Rheumatic Diseases.* 2017, **76**(6), pp.960-977.
33. Hui, M., Carr, A., Cameron, S., Davenport, G., Doherty, M. et al. The British Society for Rheumatology Guideline for the Management of Gout. *Rheumatology (Oxford).* 2017, **56**(7), pp.1056-1059.
34. Kuo, C.F., Grainge, M.J., Mallen, C., Zhang, W. and Doherty, M. Comorbidities in patients with gout prior to and following diagnosis: case-control study. *Annals of the Rheumatic Diseases.* 2014, **75**(1), pp.210-217.
35. Boonen, A., Chorus, A., Miedema, H., Van der Heijde, D., Landewe, R. et al. Withdrawal from labour force due to work disability in patients with ankylosing spondylitis. *Annals of the Rheumatic Diseases.* 2001, **60**(11), pp.1033-1039.
36. Aviña-Zubieta, J.A., Choi, H.K., Sadatsafavi, M., Etminan, M., Esdaile, J.M. et al. Risk of cardiovascular mortality in patients with rheumatoid arthritis: a meta-analysis of observational studies. *Arthritis Care & Research.* 2008, **59**(12), pp.1690-1697.
37. Parsons, S. and Symmons, D.P.M. The burden of musculoskeletal conditions. *Medicine.* 2014, **42**(4), pp.190-192.
38. Hoerbst, A. and Ammenwerth, E. Electronic health records. *Methods Inf Med.* 2010, **49**(4), pp.320-336.
39. Hall, G. Embedding technology in health and social care: PPM+ and the Leeds Care Record. In: *The Kings Fund*, 2017.

40. Kierkegaard, P. eHealth in Denmark: a case study. *J Med Syst.* 2013, **37**(6), p.9991.
41. Walsh, S.H. The clinician's perspective on electronic health records and how they can affect patient care. *BMJ.* 2004, **328**(7449), pp.1184-1187.
42. Coiera, E. When conversation is better than computation. *Journal of the American Medical Informatics Association.* 2000, **7**(3), pp.277-286.
43. Xierali, I.M., Hsiao, C.-J., Puffer, J.C., Green, L.A., Rinaldo, J.C. et al. The rise of electronic health record adoption among family physicians. *The Annals of Family Medicine.* 2013, **11**(1), pp.14-19.
44. Kim, Y.-G., Jung, K., Park, Y.-T., Shin, D., Cho, S.Y. et al. Rate of electronic health record adoption in South Korea: a nation-wide survey. *International Journal of Medical Informatics.* 2017, **101**, pp.100-107.
45. Ouhbi, S., Idri, A., Fernández-Alemán, J.L., Toval, A. and Benjelloun, H. Applying ISO/IEC 25010 on Mobile Personal Health Records. In: *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies-Volume 5*: SCITEPRESS-Science and Technology Publications, Lda, 2015, pp.405-412.
46. Coiera, E. *Guide to health informatics.* CRC press, 2015.
47. Holmgren, A.J., Adler-Milstein, J. and McCullough, J. Are all certified EHRs created equal? Assessing the relationship between EHR vendor and hospital meaningful use performance. *Journal of the American Medical Informatics Association.* 2018, **25**(6), pp.654-660.
48. Crossfield, S.S. and Clamp, S.E. *Centralised Electronic Health Records Research Across Health Organisation Types*. Springer, 2013.
49. Crossfield, S.S. and Bates, C. Centralised Health Records: Closing the Gap between Public and Personal Health In: *International Conference on Urban Health*, *Manchester*. https://www.icuh2014.com/Resources/Abstract-book-FRIDAY-ORAL-new.pdf, 2014.
50. Hecht, J. The future of electronic health records. *Nature.* 2019, **573**(7775), pp.S114-S114.
51. Office of the National Coordinator for Health Information Technology. 2016 Report to Congress on Health IT Progress: Examining the HITECH Era and the Future of Health IT. 2016.
52. Nordo, A.H., Levaux, H.P., Becnel, L.B., Galvez, J., Rao, P. et al. Use of EHRs data for clinical research: Historical progress and current applications. *Learning Health Systems.* 2019, **3**(1), p.e10076.
53. Kush, R. Interoperability review: EHRs for clinical research. *AMIA.* 2012, **2**.
54. Payne, T.H., Detmer, D.E., Wyatt, J.C. and Buchan, I.E. National-scale clinical information exchange in the United Kingdom: lessons for the United States. *Journal of the American Medical Informatics Association.* 2011, **18**(1), pp.91-98.
55. Department of Health. *The NHS quality, innovation, productivity and prevention challenge: an introduction for clinicians*. DH London. 2010.
56. Department of Health. *The power of information: Putting all of us in control of the health and care information we need.* London: Crown Copywright, 2012.
57. Wang, Y., Zhao, Y., Dang, W., Zheng, J. and Dong, H. The Evolution of Publication Hotspots in Electronic Health Records from 1957 to 2016 and Differences Among Six Countries. *Big Data.* 2020, **8**(2), pp.89-106.
58. Drozd, M., Cubbon, R., Gierula, J., Jamil, H., Crossfield, S. et al. 54 Calcium Supplementation in Patients with Chronic Heart Failure: Is it Safe? *Heart.* 2014, **100**(Suppl 3), pp.A31.31-A31.
59. Hippisley-Cox, J., Coupland, C. and Brindle, P. The performance of seven QPrediction risk scores in an independent external sample of patients from general practice: a validation study. *BMJ Open.* 2014, **4**(8), p.e005809.
60. Cottrell, E., Crabtree, V., Edwards, J.J. and Roddy, E. Improvement in the management of gout is vital and overdue: an audit from a UK primary care medical practice. *BMC Fam Pract.* 2013, **14**(170).

61. Kuo, C.F., Grainge, M.J., Mallen, C., Zhang, W. and Doherty, M. Impact of gout on the risk of atrial fibrillation. *Rheumatology (Oxford).* 2016, **55**(4), pp.721-728.
62. Dubreuil, M., Peloquin, C., Zhang, Y., Choi, H.K., Inman, R.D. et al. Validity of ankylosing spondylitis diagnoses in The Health Improvement Network. *Pharmacoepidemiology and Drug Safety.* 2016, **25**(4), pp.399-404.
63. Martin, L., Hutchens, M., Hawkins, C. and Radnov, A. How much do clinical trials cost? *Nature Reviews Drug Discovery.* 2017.
64. Kristman, V., Manno, M. and Côté, P. Loss to follow-up in cohort studies: how much is too much? *European Journal of Epidemiology.* 2004, **19**(8), pp.751-760.
65. Rea, F. *Monitoring and assessing diagnostic-therapeutic paths with healthcare utilization databases: experiences, concerns and challenges*. thesis, Bicocca Universita, 2020.
66. Crossfield, S., Bates, C. and Parry, J. ResearchOne Database Protocol. *Copyright ResearchOne.* 2012.
67. Delvaux, N., Piessens, V., De Burghgraeve, T., Mamouris, P., Vaes, B. et al. Clinical decision support improves the appropriateness of laboratory test ordering in primary care without increasing diagnostic error: the ELMO cluster randomized trial. *Implementation Science.* 2020, **15**(1), pp.1-10.
68. Ashley, L., Jones, H., Thomas, J., Forman, D., Newsham, A. et al. Integrating cancer survivors' experiences into UK cancer registries: design and development of the ePOCS system (electronic Patient-reported Outcomes from Cancer Survivors). *British Journal of Cancer.* 2011, **105**(1), pp.S74-S81.
69. Delaney, B.C., Curcin, V., Andreasson, A., Arvanitis, T.N., Bastiaens, H. et al. Translational medicine and patient safety in Europe: TRANSFoRm—architecture for the learning health system in Europe. *BioMed Research International.* 2015, **2015**.
70. Institute of Medicine. Observational Studies in a Learning Health System: Workshop Summary. *Washington (DC): National Academies Press (US).* 2013.
71. Weiskopf, N.G. and Weng, C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc.* 2013, **20**(1), pp.144-151.
72. Wells, B.J., Chagin, K.M., Nowacki, A.S. and Kattan, M.W. Strategies for handling missing data in electronic health record derived data. *EGEMS.* 2013, **1**(3).
73. Herrett, E., Thomas, S.L., Schoonen, W.M., Smeeth, L. and Hall, A.J. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol.* 2010, **69**(1), pp.4-14.
74. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence.* 2019, **1**(5), pp.206-215.
75. Gandomi, A. and Haider, M. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management.* 2015, **35**(2), pp.137-144.
76. US Department of Health and Human Services. *Health insurance portability and accountability act (HIPAA).* 1996.
77. The European Parliament and The European Council. General Data Protection Regulation. *Official Journal of the European Union.* 2016, **2014**, pp.20-30.
78. UK Parliament. *Data Protection Act 2018.* [Online]. 2018. [Accessed 13 February 2020]. Available from: http://www.legislation.gov.uk/ukpga/2018/12/contents/enacted
79. Department of Health. *The Common Law Duty of Confidentiality.* [Online]. 2020. [Accessed 16 April 2020]. Available from: https://www.health-ni.gov.uk/articles/common-law-duty-confidentiality
80. UK Government. *National Health Service Act 2006: Section 251.* [Online]. 2006. [Accessed 07 February 2019]. Available from: https://www.legislation.gov.uk/ukpga/2006/41/section/251

81. Fundamental Rights Agency. Handbook of European Data Protection Law. 2018.
82. Information Commissioner's Office. *Anonymisation Code of Practice.* [Online]. 2014. [Accessed 02 May 2018]. Available from: http://ico.org.uk/for_organisations/data_protection/topic_guides/anonymisation
83. Freeman, G. *The Care Quality Commission and National Data Guardian for Health and Care's Independent Reviews into Data Security, Consent and Opt-Outs: Written statement - HCWS62.* House of Commons, 2016.
84. Sterckx, S., Rakic, V., Cockbain, J. and Borry, P. "You hoped we would sleep walk into accepting the collection of our data": controversies surrounding the UK care. data scheme and their wider relevance for biomedical research. *Medicine, Health Care and Philosophy.* 2016, **19**(2), pp.177-190.
85. Weng, C., Appelbaum, P., Hripcsak, G., Kronish, I., Busacca, L. et al. Using EHRs to integrate research with patient care: promises and challenges. *Journal of the American Medical Informatics Association.* 2012, **19**(5), pp.684-687.
86. Gil, M., Rodríguez-Miguel, A., Montoya-Catalá, H., González-González, R., Álvarez-Gutiérrez, A. et al. Validation study of colorectal cancer diagnosis in the Spanish primary care database, BIFAP. *Pharmacoepidemiology and Drug Safety.* 2019, **28**(2), pp.209-216.
87. Horton, D.B., Bhullar, H., Carty, L., Cunningham, F., Ogdie, A. et al. Electronic Health Record Databases. In: Strom, B.L. et al. eds. *Pharmacoepidemiology, 6th Edition*. 2019, pp.241-289.
88. Johnson, A.E., Pollard, T.J., Shen, L., Lehman, L.W., Feng, M. et al. MIMIC-III, a freely accessible critical care database. *Sci Data.* 2016, **3**, p.160035.
89. Herrett, E., Gallagher, A.M., Bhaskaran, K., Forbes, H., Mathur, R. et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *International Journal of Epidemiology.* 2015, **44**(3), pp.827-836.
90. El Fadly, A., Rance, B., Lucas, N., Mead, C., Chatellier, G. et al. Integrating clinical research with the Healthcare Enterprise: from the RE-USE project to the EHR4CR platform. *Journal of Biomedical Informatics.* 2011, **44**, pp.S94-S102.
91. Chen, C., Garrido, T., Chock, D., Okawa, G. and Liang, L. The Kaiser Permanente Electronic Health Record: transforming and streamlining modalities of care. *Health Affairs.* 2009, **28**(2), pp.323-333.
92. Harno, K. and Ruotsalainen, P. Sharable EHR systems in Finland. *Studies in Health Technology and Informatics.* 2006, **121**, pp.364-370.
93. NHS Digital. *Hospital Episode Statistics.* [Online]. 2017. [Accessed 25 May 2017]. Available from: http://content.digital.nhs.uk/hes
94. Gulliford, M.C., Sun, X., Anjuman, T., Yelland, E. and Murray-Thomas, T. Comparison of antibiotic prescribing records in two UK primary care electronic health record systems: cohort study using CPRD GOLD and CPRD Aurum databases. *BMJ Open.* 2020, **10**(6), p.e038767.
95. Mathur, R., Bhaskaran, K., Chaturvedi, N., Leon, D.A., vanStaa, T. et al. Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *Journal of Public Health.* 2013, **36**(4), pp.684-692.
96. Bhaskaran, K., Forbes, H.J., Douglas, I., Leon, D.A. and Smeeth, L. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). *BMJ Open.* 2013, **3**(9), p.e003389.
97. Au, S. *CPRD Charges*, 14.09.2015, 2015.
98. Booth, H.P., Prevost, A.T. and Gulliford, M.C. Validity of smoking prevalence estimates from primary care electronic health records compared with national population survey data for England, 2007 to 2011. *Pharmacoepidemiology and Drug Safety.* 2013, **22**(12), pp.1357-1361.
99. Nissen, F., Morales, D.R., Mullerova, H., Smeeth, L., Douglas, I.J. et al. Validation of asthma recording in the Clinical Practice Research Datalink (CPRD). *BMJ Open.* 2017, **7**(8), p.e017474.
100. Herrett, E., Shah, A.D., Boggon, R., Denaxas, S., Smeeth, L. et al. Completeness and diagnostic validity of recording acute myocardial infarction

events in primary care, hospital care, disease registry, and national mortality records: cohort study. *BMJ.* 2013, **346**.
101. Blak, B., Thompson, M., Dattani, H. and Bourke, A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. *Journal of Innovation in Health Informatics.* 2011, **19**(4), pp.251-255.
102. Cegedim. *How do you fix the quality of your GP practice data?* [Online]. 2020. [Accessed 26 January 2020]. Available from: https://info.visionhealth.co.uk/5k-challenge
103. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P. 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 Medicine.* 2015, **12**(3), p.e1001779.
104. Office for National Statistics. *Deaths: Deaths broken down by age, sex, area and cause of death.* [Online]. 2021. [Accessed 27 January 2021]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths
105. Office for National Statistics. *Office for National Statistics: User guide to mortality statistics, July 2017.* [Online]. 2017. [Accessed 15 December 2017]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/methodologies/userguidetomortalitystatisticsjuly2017
106. Khanna, D., Fitzgerald, J.D., Khanna, P.P., Bae, S., Singh, M.K. et al. 2012 American College of Rheumatology guidelines for management of gout. Part 1: systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. *Arthritis Care & Research.* 2012, **64**(10), pp.1431-1446.
107. Smith, E., Hoy, D., Cross, M., Merriman, T.R., Vos, T. et al. The global burden of gout: estimates from the Global Burden of Disease 2010 study. *Annals of the Rheumatic Diseases.* 2014, **73**(8), pp.1470-1476.
108. Grassi, W. and De Angelis, R. Clinical features of gout. *Reumatismo.* 2011, pp.238-245.
109. Underwood, M. Gout. *BMJ Clinical Evidence.* 2015, **1120**.
110. Khanna, D., Khanna, P.P., Fitzgerald, J.D., Singh, M.K., Bae, S. et al. 2012 American College of Rheumatology guidelines for management of gout. Part 2: therapy and antiinflammatory prophylaxis of acute gouty arthritis. *Arthritis Care & Research.* 2012, **64**(10), pp.1447-1461.
111. Bardin, T. and Richette, P. Definition of hyperuricemia and gouty conditions. *Curr Opin Rheumatol.* 2014, **26**(2), pp.186-191.
112. Loeb, J.N. The influence of temperature on the solubility of monosodium urate. *Arthritis & Rheumatism.* 1972, **15**(2), pp.189-192.
113. Roddy, E., Zhang, W. and Doherty, M. The changing epidemiology of gout. *Nat Clin Pract Rheumatol.* 2007, **3**(8), pp.443-449.
114. Kuo, C.F., Grainge, M.J., Zhang, W. and Doherty, M. Global epidemiology of gout: prevalence, incidence and risk factors. *Nat Rev Rheumatol.* 2015, **11**(11), pp.649-662.
115. Jackson, G., Wright, C., Thornley, S., Taylor, W.J., Te Karu, L. et al. Potential unmet need for gout diagnosis and treatment: capture–recapture analysis of a national administrative dataset. *Rheumatology (Oxford).* 2012, **51**(10), pp.1820-1824.
116. Kuo, C.F., Grainge, M.J., Mallen, C., Zhang, W. and Doherty, M. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. *Annals of the Rheumatic Diseases.* 2014, **74**(4), pp.661-667.
117. Rothenbacher, D., Primatesta, P., Ferreira, A., Cea-Soriano, L. and Rodriguez, L.A. Frequency and risk factors of gout flares in a large population-based cohort of incident gout. *Rheumatology (Oxford).* 2011, **50**(5), pp.973-981.
118. Roddy, E. and Doherty, M. Gout. Epidemiology of gout. *Arthritis Research & Therapy.* 2010, **12**(6), pp.1-11.

119. Arromdee, E., Michet, C.J., Crowson, C.S., O'Fallon, W.M. and Gabriel, S.E. Epidemiology of gout: is the incidence rising? *J Rheumatol.* 2002, **29**(11), pp.2403-2406.
120. Major, T.J., Dalbeth, N., Stahl, E.A. and Merriman, T.R. An update on the genetics of hyperuricaemia and gout. *Nature Reviews Rheumatology.* 2018, **14**(6), pp.341-353.
121. Krishnan, E., Lessov-Schlaggar, C.N., Krasnow, R.E. and Swan, G.E. Nature versus nurture in gout: a twin study. *The American Journal of Medicine.* 2012, **125**(5), pp.499-504.
122. Jordan, K.M., Cameron, J.S., Snaith, M., Zhang, W., Doherty, M. et al. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of gout. *Rheumatology (Oxford).* 2007, **46**(8), pp.1372-1374.
123. Elliot, A.J., Cross, K.W. and Fleming, D.M. Seasonality and trends in the incidence and prevalence of gout in England and Wales 1994-2007. *Annals of the Rheumatic Diseases.* 2009, **68**(11), pp.1728-1733.
124. Huang, Q.R., Zhenxing, Q., Zhang, S. and Chow, C.M. Clinical Patterns of Obstructive Sleep Apnea and Its Comorbid Conditions: A data mining approach. *Journal of Clinical Sleep Medicine.* 2008, **4**(6), pp.543-550.
125. Ben Salem, C., Slim, R., Fathallah, N. and Hmouda, H. Drug-induced hyperuricaemia and gout. *Rheumatology (Oxford).* 2017, **56**(5), pp.679-688.
126. Mikuls, T.R., Farrar, J.T., Bilker, W.B., Fernandes, S., Schumacher, H.R., Jr. et al. Gout epidemiology: results from the UK General Practice Research Database, 1990-1999. *Annals of the Rheumatic Diseases.* 2005, **64**(2), pp.267-272.
127. Rho, Y.H., Lu, N., Peloquin, C.E., Man, A., Zhu, Y. et al. Independent impact of gout on the risk of diabetes mellitus among women and men: a population-based, BMI-matched cohort study. *Annals of the Rheumatic Diseases.* 2014, **75**(1), pp.91-95.
128. Scherer, M., Hansen, H., Gensichen, J., Mergenthal, K., Riedel-Heller, S. et al. Association between multimorbidity patterns and chronic pain in elderly primary care patients: a cross-sectional observational study. *BMC Fam Pract.* 2016, **17**, p.8.
129. Zhang, W., Doherty, M., Bardin, T., Pascual, E., Barskova, V. et al. EULAR evidence based recommendations for gout. Part II: Management. Report of a task force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Annals of the Rheumatic Diseases.* 2006, **65**(10), pp.1312-1324.
130. Garbe, E., Suissa, S. and LeLorier, J. Exposure to allopurinol and the risk of cataract extraction in elderly patients. *Archives of Ophthalmology.* 1998, **116**(12), pp.1652-1656.
131. Alonso, A., Rodriguez, L.A., Logroscino, G. and Hernan, M.A. Gout and risk of Parkinson disease: a prospective study. *Neurology.* 2007, **69**(17), pp.1696-1700.
132. National Institute for Health and Care Excellence. *Gout.* [Online]. 2018. [Accessed 09 December 2020]. Available from: https://cks.nice.org.uk/topics/gout/diagnosis/
133. Sturrock, R. Gout: Easy to misdiagnose. *BMJ.* 2000.
134. Zhang, W., Doherty, M., Pascual, E., Bardin, T., Barskova, V. et al. EULAR evidence based recommendations for gout. Part I: Diagnosis. Report of a task force of the Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Annals of the Rheumatic Diseases.* 2006, **65**(10), pp.1301-1311.
135. Richette, P., Doherty, M., Pascual, E., Barskova, V., Becce, F. et al. SAT0531 Updated EULAR Evidence-Based Recommendations for the Management of Gout. *Annals of the Rheumatic Diseases.* 2014, **73**(Suppl 2), pp.783.781-783.

136. Schlesinger, N., Detry, M.A., Holland, B.K., Baker, D.G., Beutler, A.M. et al. Local ice therapy during bouts of acute gouty arthritis. *The Journal of Rheumatology.* 2002, **29**(2), pp.331-334.
137. Annemans, L., Spaepen, E., Gaskin, M., Bonnemaire, M., Malier, V. et al. Gout in the UK and Germany: prevalence, comorbidities and management in general practice 2000-2005. *Annals of the Rheumatic Diseases.* 2008, **67**(7), pp.960-966.
138. Doghramji, P.P. Managing your patient with gout: A review of treatment options. *Postgraduate Medicine.* 2011, **123**(3), pp.56-71.
139. Richette, P., Doherty, M., Pascual, E., Barskova, V., Becce, F. et al. 2016 updated EULAR evidence-based recommendations for the management of gout. *Annals of the Rheumatic Diseases.* 2016, **76**(1), pp.29-42.
140. Roddy, E., Zhang, W. and Doherty, M. Concordance of the management of chronic gout in a UK primary-care population with the EULAR gout recommendations. *Annals of the Rheumatic Diseases.* 2007, **66**(10), pp.1311-1315.
141. Dalbeth, N., House, M.E., Horne, A., Petrie, K.J., McQueen, F.M. et al. Prescription and dosing of urate-lowering therapy, rather than patient behaviours, are the key modifiable factors associated with targeting serum urate in gout. *BMC Musculoskelet Disord.* 2012, **13**, p.5.
142. Mikuls, T.R., MacLean, C.H., Olivieri, J., Patino, F., Allison, J.J. et al. Quality of care indicators for gout management. *Arthritis & Rheumatism.* 2004, **50**(3), pp.937-943.
143. Lieberman, J.A., 3rd. Treatment and prophylaxis of gout flare in the clinic: an office-based approach to gout management. *Postgraduate Medicine.* 2011, **123**(6), pp.151-165.
144. Chandratre, P., Mallen, C., Roddy, E., Liddle, J. and Richardson, J. "You want to get on with the rest of your life": a qualitative study of health-related quality of life in gout *Clinical Rheumatology.* 2016, **35**(5), p.9.
145. Corbett, E., Pentony, P. and McGill, N. AB0806 How Often Do Gout Patients Reach Target in A Rheumatology Practice? *Annals of the Rheumatic Diseases.* 2016.
146. Kuo, C.F., Grainge, M.J., Mallen, C., Zhang, W. and Doherty, M. Effect of allopurinol on all-cause mortality in adults with incident gout: propensity score-matched landmark analysis. *Rheumatology (Oxford).* 2015, **54**(12), pp.2145-2150.
147. Sarawate, C.A., Brewer, K.K., Yang, W., Patel, P.A., Schumacher, H.R. et al. Gout medication treatment patterns and adherence to standards of care from a managed care perspective. *Mayo Clinic Proceedings.* 2006, **81**(7), pp.925-934.
148. Fanouriakis, A., Karantanas, A., Fragouli, E., Repa, A. and Sidiropoulos, P. Control of flares and relief of urate burden with canakinumab and targeted urate-lowering therapy in tophaceous gouty arthritis. *Rheumatology (Oxford).* 2014, **53**(4), pp.764-766.
149. Harrold, L.R., Andrade, S.E., Briesacher, B., Raebel, M.A., Fouayzi, H. et al. The dynamics of chronic gout treatment: medication gaps and return to therapy. *The American Journal of Medicine.* 2010, **123**(1), pp.54-59.
150. Mikuls, T.R., Farrar, J.T., Bilker, W.B., Fernandes, S. and Saag, K.G. Suboptimal physician adherence to quality indicators for the management of gout and asymptomatic hyperuricaemia: results from the UK General Practice Research Database (GPRD). *Rheumatology (Oxford).* 2005, **44**(8), pp.1038-1042.
151. Cottrell, E., Crabtree, V., Edwards, J.J. and Roddy, E. Improvement in the management of gout is vital and overdue: an audit from a UK primary care medical practice. *BMC Fam Pract.* 2013, **14**, p.170.
152. Kuo, C.F., Grainge, M.J., Mallen, C., Zhang, W. and Doherty, M. Eligibility for and prescription of urate-lowering treatment in patients with incident gout in England. *JAMA.* 2014, **312**(24), pp.2684-2686.

153. Al-Allaf, A.-W., Mohiaddin, H. and Al-Allaf, O. AB0799 Gout Management: Audit for Best Practice in Gout. *Annals of the Rheumatic Diseases.* 2016.
154. Rees, F., Jenkins, W. and Doherty, M. Patients with gout adhere to curative treatment if informed appropriately: proof-of-concept observational study. *Annals of the Rheumatic Diseases.* 2013, **72**(6), pp.826-830.
155. Wright, E., Darer, J., Tang, X., Thompson, J., Tusing, L. et al. Sharing Physician Notes Through an Electronic Portal is Associated With Improved Medication Adherence: Quasi-Experimental Study. *J Med Internet Res.* 2015, **17**(10), p.e226.
156. Rashid, N., Levy, G.D., Wu, Y.L., Zheng, C.Y., Koblick, R. et al. Patient and clinical characteristics associated with gout flares in an integrated healthcare system. *Rheumatol Int.* 2015, **35**(11), pp.1799-1807.
157. Roddy, E., Mallen, C.D., Hider, S.L. and Jordan, K.P. Prescription and comorbidity screening following consultation for acute gout in primary care. *Rheumatology (Oxford).* 2010, **49**(1), pp.105-111.
158. Clarson, L.E., Hider, S.L., Belcher, J., Heneghan, C., Roddy, E. et al. Increased risk of vascular disease associated with gout: a retrospective, matched cohort study in the UK clinical practice research datalink. *Annals of the Rheumatic Diseases.* 2015, **74**(4), pp.642-647.
159. Benis, A., Jaffe, D., Flores, N., Gabay, H., Morlock, R. et al. FRI0564 Serum Uric Acid Testing Practices over Five Years among Incident Gout Cases. *Annals of the Rheumatic Diseases.* 2016.
160. Wijnands, J.M., van Durme, C.M., Driessen, J.H., Boonen, A., Klop, C. et al. Individuals With Type 2 Diabetes Mellitus Are at an Increased Risk of Gout But This Is Not Due to Diabetes: A Population-Based Cohort Study. *Medicine (Baltimore).* 2015, **94**(32), p.e1358.
161. Bruderer, S.G., Bodmer, M., Jick, S.S. and Meier, C.R. Poorly controlled type 2 diabetes mellitus is associated with a decreased risk of incident gout: a population-based case-control study. *Annals of the Rheumatic Diseases.* 2015, **74**(9), pp.1651-1658.
162. Bruderer, S.G., Bodmer, M., Jick, S.S. and Meier, C.R. Association of hormone therapy and incident gout: population-based case-control study. *Menopause.* 2015, **22**(12), p.8.
163. Braun, J. and Sieper, J. Ankylosing spondylitis. *The Lancet.* 2007, **369**(9570), pp.1379-1390.
164. Martin, T.M., Smith, J.R. and Rosenbaum, J.T. Anterior uveitis: current concepts of pathogenesis and interactions with the spondyloarthropathies. *Curr Opin Rheumatol.* 2002, **14**(4), pp.337-341.
165. Roux, C. Osteoporosis in inflammatory joint diseases. *Osteoporosis International.* 2011, **22**(2), pp.421-433.
166. Schlosstein, L., Terasaki, P.I., Bluestone, R. and Pearson, C.M. High association of an HL-A antigen, W27, with ankylosing spondylitis. *New England Journal of Medicine.* 1973, **288**(14), pp.704-706.
167. Mielants, H., Veys, E., Joos, R., Noens, L., Cuvelier, C. et al. HLA antigens in seronegative spondylarthropathies. Reactive arthritis and arthritis in ankylosing spondylitis: relation to gut inflammation. *The Journal of Rheumatology.* 1987, **14**(3), pp.466-471.
168. Zhu, W., He, X., Cheng, K., Zhang, L., Chen, D. et al. Ankylosing spondylitis: etiology, pathogenesis, and treatments. *Bone Research.* 2019, **7**(1), pp.1-16.
169. Van der Heijde, D., Landewe, R., Einstein, S., Ory, P., Vosse, D. et al. Radiographic progression of ankylosing spondylitis after up to two years of treatment with etanercept. *Arthritis & Rheumatism.* 2008, **58**(5), pp.1324-1331.
170. Gran, J. and Skomsvoll, J. The outcome of ankylosing spondylitis: a study of 100 patients. *Rheumatology (Oxford).* 1997, **36**(7), pp.766-771.
171. Sieper, J., Braun, J., Rudwaleit, M., Boonen, A. and Zink, A. Ankylosing spondylitis: an overview. *Annals of the Rheumatic Diseases.* 2002, **61**(suppl 3), pp.iii8-iii18.

172. Boonen, A., Brinkhuizen, T., Landewé, R., van der Heijde, D. and Severens, J.L. Impact of ankylosing spondylitis on sick leave, presenteeism and unpaid productivity, and estimation of the societal cost. *Annals of the Rheumatic Diseases.* 2010, **69**(6), pp.1123-1128.
173. Bostan, E.E., Borman, P., Bodur, H. and Barca, N. Functional disability and quality of life in patients with ankylosing spondylitis. *Rheumatology International.* 2003, **23**(3), pp.121-126.
174. Maksymowych, W.P., Gooch, K.L., Wong, R.L., Kupper, H. and van der Heijde, D. Impact of age, sex, physical function, health-related quality of life, and treatment with adalimumab on work status and work productivity of patients with ankylosing spondylitis. *The Journal of Rheumatology.* 2010, **37**(2), pp.385-392.
175. Mathieu, S., Pereira, B. and Soubrier, M. Cardiovascular events in ankylosing spondylitis: an updated meta-analysis. In: *Seminars in Arthritis and Rheumatism*: Elsevier, 2015, pp.551-555.
176. Bakland, G., Gran, J.T. and Nossent, J.C. Increased mortality in ankylosing spondylitis is related to disease activity. *Annals of the Rheumatic Diseases.* 2011, **70**(11), pp.1921-1925.
177. Exarchou, S., Lie, E., Lindström, U., Askling, J., Forsblad-d'Elia, H. et al. Mortality in ankylosing spondylitis: results from a nationwide population-based study. *Ann Rheum Dis.* 2015, **75**(8), pp.1466-1472.
178. Rudwaleit, M., Listing, J., Brandt, J., Braun, J. and Sieper, J. Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor α blockers in ankylosing spondylitis. *Annals of the Rheumatic Diseases.* 2004, **63**(6), pp.665-670.
179. Yi, E., Ahuja, A., Rajput, T., George, A.T. and Park, Y. Clinical, Economic, and Humanistic Burden Associated With Delayed Diagnosis of Axial Spondyloarthritis: A Systematic Review. *Rheumatology Therapy.* 2020, pp.1-23.
180. Stolwijk, C., Essers, I., van Tubergen, A., Boonen, A., Bazelier, M.T. et al. The epidemiology of extra-articular manifestations in ankylosing spondylitis: a population-based matched cohort study. *Annals of the Rheumatic Diseases.* 2015, **74**(7), pp.1373-1378.
181. Wang, R. and Ward, M.M. Epidemiology of axial spondyloarthritis: an update. *Curr Opin Rheumatol.* 2018, **30**(2), p.137.
182. Bakland, G., Nossent, H.C. and Gran, J.T. Incidence and prevalence of ankylosing spondylitis in Northern Norway. *Arthritis Care & Research.* 2005, **53**(6), pp.850-855.
183. Wright, K.A., Crowson, C.S., Michet, C.J. and Matteson, E.L. Time trends in incidence, clinical features, and cardiovascular disease in ankylosing spondylitis over three decades: a population-based study. *Arthritis Care & Research.* 2015, **67**(6), pp.836-841.
184. Carbone, L.D., Cooper, C., Michet, C.J., Atkinson, E.J., Michael O'Fallon, W. et al. Ankylosing spondylitis in Rochester, Minnesota, 1935–1989. Is the epidemiology changing? *Arthritis & Rheumatism.* 1992, **35**(12), pp.1476-1482.
185. Park, J.-S., Hong, J.-Y., Park, Y.-S., Han, K. and Suh, S.-W. Trends in the prevalence and incidence of ankylosing spondylitis in South Korea, 2010–2015 and estimated differences according to income status. *Scientific Reports.* 2018, **8**(1), pp.1-6.
186. Brown, M., Jepson, A., Young, A., Whittle, H., Greenwood, B. et al. Ankylosing spondylitis in West Africans—evidence for a non-HLA-B27 protective effect. *Annals of the Rheumatic Diseases.* 1997, **56**(1), pp.68-70.
187. Reveille, J.D., Hirsch, R., Dillon, C.F., Carroll, M.D. and Weisman, M.H. The prevalence of HLA–B27 in the US: data from the US National Health and Nutrition Examination Survey, 2009. *Arthritis & Rheumatism.* 2012, **64**(5), pp.1407-1411.
188. Taurog, J.D., Chhabra, A. and Colbert, R.A. Ankylosing spondylitis and axial spondyloarthritis. *New England Journal of Medicine.* 2016, **374**(26), pp.2563-2574.

189. Cooksey, R., Husain, M.J., Brophy, S., Davies, H., Rahman, M.A. et al. The Cost of Ankylosing Spondylitis in the UK Using Linked Routine and Patient-Reported Survey Data. *PLoS One.* 2015, **10**(7), p.e0126105.
190. McLeod, C., Bagust, A., Boland, A., Dagenais, P., Dickson, R. et al. Adalimumab, etanercept and infliximab for the treatment of ankylosing spondylitis: a systematic review and economic evaluation. In: *NIHR Health Technology Assessment programme: Executive Summaries*. NIHR Journals Library, 2007.
191. Asquith, M., Sternes, P.R., Costello, M.E., Karstens, L., Diamond, S. et al. HLA alleles associated with risk of ankylosing spondylitis and rheumatoid arthritis influence the gut microbiome. *Arthritis & Rheumatology.* 2019, **71**(10), pp.1642-1650.
192. Cortes, A., Hadler, J., Pointon, J.P., Robinson, P.C., Karaderi, T. et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nature Genetics.* 2013, **45**(7), p.730.
193. Hanson, A. and Brown, M.A. Genetics and the causes of ankylosing spondylitis. *Rheumatic Disease Clinics of North America.* 2017, **43**(3), pp.401-414.
194. Mielants, H., Veys, E., Cuvelier, C., De Vos, M. and Botelberghe, L. HLA-B27 related arthritis and bowel inflammation. Part 2. Ileocolonoscopy and bowel histology in patients with HLA-B27 related arthritis. *The Journal of Rheumatology.* 1985, **12**(2), pp.294-298.
195. Ciccia, F., Accardo-Palumbo, A., Alessandro, R., Rizzo, A., Principe, S. et al. Interleukin-22 and interleukin-22–producing NKp44+ natural killer cells in subclinical gut inflammation in ankylosing spondylitis. *Arthritis & Rheumatism.* 2012, **64**(6), pp.1869-1878.
196. Martinez-Gonzalez, O., Cantero-Hinojosa, J., Paule-Sastre, P., Gomez-Magan, J. and Salvatierra-Rios, D. Intestinal permeability in patients with ankylosing spondylitis and their healthy relatives. *British Journal of Rheumatology.* 1994.
197. Costello, M.-E., Elewaut, D., Kenna, T.J. and Brown, M.A. Microbes, the gut and ankylosing spondylitis. *Arthritis Research & Therapy.* 2013, **15**(3), pp.1-12.
198. Gooren, L.J., Giltay, E.J., van Schaardenburg, D. and Dijkmans, B.A. Gonadal and adrenal sex steroids in ankylosing spondylitis. *Rheumatic Disease Clinics of North America.* 2000, **26**(4), pp.969-987.
199. Gracey, E., Yao, Y., Green, B., Qaiyum, Z., Baglaenko, Y. et al. Sexual dimorphism in the Th17 signature of ankylosing spondylitis. *Arthritis & Rheumatology.* 2016, **68**(3), pp.679-689.
200. Cua, D.J. and Tato, C.M. Innate IL-17-producing cells: the sentinels of the immune system. *Nature Reviews Immunology.* 2010, **10**(7), pp.479-489.
201. van der Horst-Bruinsma, I.E., Zack, D.J., Szumski, A. and Koenig, A.S. Female patients with ankylosing spondylitis: analysis of the impact of gender across treatment studies. *Annals of the Rheumatic Diseases.* 2013, **72**(7), pp.1221-1224.
202. Linden, S.V.D., Valkenburg, H.A. and Cats, A. Evaluation of diagnostic criteria for ankylosing spondylitis. *Arthritis & Rheumatism.* 1984, **27**(4), pp.361-368.
203. Lee, W., Reveille, J.D., Davis, J.C., Learch, T.J., Ward, M.M. et al. Are there gender differences in severity of ankylosing spondylitis? Results from the PSOAS cohort. *Annals of the Rheumatic Diseases.* 2007, **66**(5), pp.633-638.
204. Alamanos, Y., Papadopoulos, N., Voulgari, P., Karakatsanis, A., Siozos, C. et al. Epidemiology of ankylosing spondylitis in Northwest Greece, 1983–2002. *Rheumatology (Oxford).* 2004, **43**(5), pp.615-618.
205. Rudwaleit, M., Khan, M.A. and Sieper, J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? *Arthritis & Rheumatism.* 2005, **52**(4), pp.1000-1008.
206. Poddubnyy, D., Haibel, H., Listing, J., Märker-Hermann, E., Zeidler, H. et al. Baseline radiographic damage, elevated acute-phase reactant levels, and cigarette smoking status predict spinal radiographic progression in early axial spondylarthritis. *Arthritis & Rheumatism.* 2012, **64**(5), pp.1388-1398.

207. Rosas, J., Llinares-Tello, F., Senabre-Gallego, J.M., Barber-Vallés, X., Santos-Soler, G. et al. Obesity decreases clinical efficacy and levels of adalimumab in patients with ankylosing spondylitis. *Clin Exp Rheumatol.* 2017, **35**(1), pp.145-148.
208. de Winter, J.J., van Mens, L.J., van der Heijde, D., Landewé, R. and Baeten, D.L. Prevalence of peripheral and extra-articular disease in ankylosing spondylitis versus non-radiographic axial spondyloarthritis: a meta-analysis. *Arthritis Research & Therapy.* 2016, **18**(1), p.196.
209. Zeboulon, N., Dougados, M. and Gossec, L. Prevalence and characteristics of uveitis in the spondyloarthropathies: a systematic literature review. *Annals of the Rheumatic Diseases.* 2008, **67**(7), pp.955-959.
210. Stone, M.A., Mayberry, J.F. and Baker, R. Prevalence and management of inflammatory bowel disease: a cross-sectional study from central England. *European Journal of Gastroenterology and Hepatology.* 2003, **15**(12), pp.1275-1280.
211. Springate, D., Parisi, R., Kontopantelis, E., Reeves, D., Griffiths, C. et al. Incidence, prevalence and mortality of patients with psoriasis: a UK population-based cohort study. *British Journal of Dermatology.* 2017, **176**(3), pp.650-658.
212. Derakhshan, M.H., Dean, L., Jones, G.T., Siebert, S. and Gaffney, K. Predictors of extra-articular manifestations in axial spondyloarthritis and their influence on TNF-inhibitor prescribing patterns: results from the British Society for Rheumatology Biologics Register in Ankylosing Spondylitis. *RMD Open.* 2020, **6**(2), p.e001206.
213. Singh, G., Lawrence, A., Agarwal, V., Misra, R. and Aggarwal, A. Higher prevalence of extra-articular manifestations in ankylosing spondylitis with peripheral arthritis. *Journal of Clinical Rheumatology.* 2008, **14**(5), pp.264-266.
214. Ostensen, M. and Ostensen, H. Ankylosing spondylitis--the female aspect. *The Journal of Rheumatology.* 1998, **25**(1), p.120.
215. Davey-Ranasinghe, N. and Deodhar, A. Osteoporosis and vertebral fractures in ankylosing spondylitis. *Curr Opin Rheumatol.* 2013, **25**(4), pp.509-516.
216. Caplan, A., Fett, N., Rosenbach, M., Werth, V.P. and Micheletti, R.G. Prevention and management of glucocorticoid-induced side effects: a comprehensive review: a review of glucocorticoid pharmacology and bone health. *Journal of the American Academy of Dermatology.* 2017, **76**(1), pp.1-9.
217. Marsico, F., Paolillo, S. and Filardi, P.P. NSAIDs and cardiovascular risk. *Journal of Cardiovascular Medicine.* 2017, **18**, pp.e40-e43.
218. Van Eijk, I., Peters, M., Serne, E., Van der Horst-Bruinsma, I., Dijkmans, B. et al. Microvascular function is impaired in ankylosing spondylitis and improves after tumour necrosis factor α blockade. *Annals of the Rheumatic Diseases.* 2009, **68**(3), pp.362-366.
219. van Sijl, A.M., van Eijk, I.C., Peters, M.J., Serné, E.H., van der Horst-Bruinsma, I.E. et al. Tumour necrosis factor blocking agents and progression of subclinical atherosclerosis in patients with ankylosing spondylitis. *Annals of the Rheumatic Diseases.* 2015, **74**(1), pp.119-123.
220. Bremander, A., Petersson, I.F., Bergman, S. and Englund, M. Population-based estimates of common comorbidities and cardiovascular disease in ankylosing spondylitis. *Arthritis Care & Research.* 2011, **63**(4), pp.550-556.
221. Lehtinen, K. Mortality and causes of death in 398 patients admitted to hospital with ankylosing spondylitis. *Annals of the Rheumatic Diseases.* 1993, **52**(3), pp.174-176.
222. Song, I.-H., Sieper, J. and Rudwaleit, M. Diagnosing early ankylosing spondylitis. *Current Rheumatology Reports.* 2007, **9**(5), pp.367-374.
223. National Institute for Health and Care Excellence. *Ankylosing spondylitis: How do I make a working diagnosis of ankylosing spondylitis?* [Online]. 2019. [Accessed 21 Sep 2020]. Available from: https://cks.nice.org.uk/topics/ankylosing-spondylitis/diagnosis/diagnosis/

224. Rudwaleit, M., Van der Heijde, D., Khan, M., Braun, J. and Sieper, J. How to diagnose axial spondyloarthritis early. *Annals of the Rheumatic Diseases.* 2004, **63**(5), pp.535-543.
225. Cansu, D.Ü., Çalışır, C., Yavaş, U.S., Kaşifoğlu, T. and Korkmaz, C. Predictors of radiographic severity and functional disability in Turkish patients with ankylosing spondylitis. *Clinical Rheumatology.* 2011, **30**(4), pp.557-562.
226. Raychaudhuri, S.P. and Deodhar, A. The classification and diagnostic criteria of ankylosing spondylitis. *Journal of Autoimmunity.* 2014, **48**, pp.128-133.
227. Savigny, P., Kuntze, S., Watson, P., Underwood, M., Ritchie, G. et al. Low back pain: early management of persistent non-specific low back pain. *London: National Collaborating Centre for Primary Care, Royal College of General Practitioners.* 2009, **14**(1), pp.9-13.
228. Hamilton, L., Gilbert, A., Skerrett, J., Dickinson, S. and Gaffney, K. Services for people with ankylosing spondylitis in the UK—a survey of rheumatologists and patients. *Rheumatology (Oxford).* 2011, **50**(11), pp.1991-1998.
229. Feldtkeller, E. Age at disease onset and delayed diagnosis of spondyloarthropathies. *Zeitschrift fur Rheumatologie.* 1999, **58**(1), p.21.
230. Derakhshan, M.H., Pathak, H., Cook, D., Dickinson, S., Siebert, S. et al. Services for spondyloarthritis: a survey of patients and rheumatologists. *Rheumatology (Oxford).* 2018, **57**(6), pp.987-996.
231. Landi, M., Maldonado-Ficco, H., Perez-Alamino, R., Maldonado-Cocco, J.A., Citera, G. et al. Gender differences among patients with primary ankylosing spondylitis and spondylitis associated with psoriasis and inflammatory bowel disease in an iberoamerican spondyloarthritis cohort. *Medicine (Baltimore).* 2016, **95**(51).
232. Braun, J., Van Den Berg, R., Baraliakos, X., Boehm, H., Burgos-Vargas, R. et al. 2010 update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. *Annals of the Rheumatic Diseases.* 2011, **70**(6), pp.896-904.
233. Analay, Y., Ozcan, E., Karan, A., Diracoglu, D. and Aydin, R. The effectiveness of intensive group exercise on patients with ankylosing spondylitis. *Clinical Rehabilitation.* 2003, **17**(6), pp.631-636.
234. Demontis, A., Trainito, S., Del Felice, A. and Masiero, S. Favorable effect of rehabilitation on balance in ankylosing spondylitis: a quasi-randomized controlled clinical trial. *Rheumatology International.* 2016, **36**(3), pp.333-339.
235. Rodriguez-Lozano, C., Juanola, X., Cruz-Martinez, J., Pena-Arrebola, A., Mulero, J. et al. Outcome of an education and home-based exercise programme for patients with ankylosing spondylitis: a nationwide randomized study. *Clin Exp Rheumatol.* 2013, **31**(5), pp.739-748.
236. Zão, A. and Cantista, P. The role of land and aquatic exercise in ankylosing spondylitis: a systematic review. *Rheumatology International.* 2017, **37**(12), pp.1979-1990.
237. Poddubnyy, D., Rudwaleit, M., Haibel, H., Listing, J., Märker-Hermann, E. et al. Effect of non-steroidal anti-inflammatory drugs on radiographic spinal progression in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. *Annals of the Rheumatic Diseases.* 2012, **71**(10), pp.1616-1622.
238. National Institute for Health and Care Excellence. *NSAIDs - prescribing issues.* [Online]. 2019. [Accessed 23 October 2019]. Available from: https://cks.nice.org.uk/nsaids-prescribing-issues#!scenario
239. Henry, D. and McGettigan, P. Epidemiology overview of gastrointestinal and renal toxicity of NSAIDs. *International Journal of Clinical Practice.* 2003, (135), p.43.
240. Keat, A., Barkham, N., Bhalla, A., Gaffney, K., Marzo-Ortega, H. et al. BSR guidelines for prescribing TNF-α blockers in adults with ankylosing spondylitis. Report of a working party of the British Society for Rheumatology. *Rheumatology (Oxford).* 2005, **44**(7), pp.939-947.

241. Mease, P. Emerging immunomodulatory therapies and new treatment paradigms for axial spondyloarthritis. *Current Rheumatology Reports.* 2019, **21**(7), p.35.
242. Braun, J., Zochling, J., Baraliakos, X., Alten, R., Burmester, G. et al. Efficacy of sulfasalazine in patients with inflammatory back pain due to undifferentiated spondyloarthritis and early ankylosing spondylitis: a multicentre randomised controlled trial. *Annals of the Rheumatic Diseases.* 2006, **65**(9), pp.1147-1153.
243. Chen, J., Lin, S. and Liu, C. Sulfasalazine for ankylosing spondylitis. *Cochrane Database of Systematic Reviews.* 2014, (11).
244. Haibel, H., Rudwaleit, M., Braun, J. and Sieper, J. Six months open label trial of leflunomide in active ankylosing spondylitis. *Annals of the Rheumatic Diseases.* 2005, **64**(1), pp.124-126.
245. Van Denderen, J.C., Van der Paardt, M., Nurmohamed, M.T., De Ryck, Y.M., Dijkmans, B.A. et al. Double blind, randomised, placebo controlled study of leflunomide in the treatment of active ankylosing spondylitis. *Annals of the Rheumatic Diseases.* 2005, **64**(12), pp.1761-1764.
246. Gonzalez-Lopez, L., Garcia-Gonzalez, A., Vazquez-Del-Mercado, M., Muñoz-Valle, J.F. and Gamez-Nava, J.I. Efficacy of methotrexate in ankylosing spondylitis: a randomized, double blind, placebo controlled trial. *The Journal of Rheumatology.* 2004, **31**(8), pp.1568-1574.
247. Haibel, H., Brandt, H., Song, I., Brandt, A., Listing, J. et al. No efficacy of subcutaneous methotrexate in active ankylosing spondylitis: a 16-week open-label trial. *Annals of the Rheumatic Diseases.* 2007, **66**(3), pp.419-421.
248. Kubiak, E.N., Moskovich, R., Errico, T.J. and Di Cesare, P.E. Orthopaedic management of ankylosing spondylitis. *Journal of the American Academy of Orthopaedic Surgeons.* 2005, **13**(4), pp.267-278.
249. Haroon, N.N., Paterson, J.M., Li, P. and Haroon, N. Increasing proportion of female patients with ankylosing spondylitis: a population-based study of trends in the incidence and prevalence of AS. *BMJ Open.* 2014, **4**(12), p.e006634.
250. Cakir, N., Pamuk, O.N., Dervis, E., Imeryuz, N., Uslu, H. et al. The prevalences of some rheumatic diseases in western Turkey: Havsa study. *Rheumatol Int.* 2012, **32**(4), pp.895-908.
251. Curtis, J.R., Harrold, L.R., Asgari, M.M., Deodhar, A., Salman, C. et al. Diagnostic Prevalence of Ankylosing Spondylitis Using Computerized Health Care Data, 1996 to 2009: Underrecognition in a US Health Care Setting. *Perm J.* 2016, **20**(4), pp.4-10.
252. Canhao, H., Faustino, A., Martins, F. and Fonseca, J.E. Reuma.pt - the rheumatic diseases portuguese register. *Acta Reumatol Port.* 2011, **36**(1), pp.45-56.
253. Majithia, V. and Geraci, S.A. Rheumatoid arthritis: diagnosis and management. *The American Journal of Medicine.* 2007, **120**(11), pp.936-939.
254. Prete, M., Racanelli, V., Digiglio, L., Vacca, A., Dammacco, F. et al. Extra-articular manifestations of rheumatoid arthritis: an update. *Autoimmunity Reviews.* 2011, **11**(2), pp.123-131.
255. Smolen, J.S., Landewé, R., Breedveld, F.C., Dougados, M., Emery, P. et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Annals of the Rheumatic Diseases.* 2010, **69**(6), pp.964-975.
256. McInnes, I.B. and Schett, G. The pathogenesis of rheumatoid arthritis. *New England Journal of Medicine.* 2011, **365**(23), pp.2205-2219.
257. Filipovic, I., Walker, D., Forster, F. and Curry, A.S. Quantifying the economic burden of productivity loss in rheumatoid arthritis. *Rheumatology (Oxford).* 2011, **50**(6), pp.1083-1090.
258. Meenan, R.F., Kazis, L.E., Anthony, J.M., Wallin, B.A.J.A. and Rheumatology, R.O.J.o.t.A.C.o. The clinical and health status of patients with recent-onset rheumatoid arthritis. 1991, **34**(6), pp.761-765.

259. Widdifield, J., Bernatsky, S., Paterson, J.M., Tomlinson, G., Tu, K. et al. Trends in excess mortality among patients with rheumatoid arthritis in Ontario, Canada. *Arthritis Care & Research.* 2015, **67**(8), pp.1047-1053.
260. Listing, J., Kekow, J., Manger, B., Burmester, G.-R., Pattloch, D. et al. Mortality in rheumatoid arthritis: the impact of disease activity, treatment with glucocorticoids, TNFα inhibitors and rituximab. *Annals of the Rheumatic Diseases.* 2015, **74**(2), pp.415-421.
261. Abhishek, A., Jenkins, W., La-Crette, J., Fernandes, G. and Doherty, M. Long-term persistence and adherence on urate-lowering treatment can be maintained in primary care-5-year follow-up of a proof-of-concept study. *Rheumatology (Oxford).* 2017, **56**(4), pp.529-533.
262. Alamanos, Y. and Drosos, A.A. Epidemiology of adult rheumatoid arthritis. *Autoimmunity Reviews.* 2005, **4**(3), pp.130-136.
263. Albrecht, K., Callhoff, J. and Zink, A. Long-term trends in rheumatology care: Achievements and deficits in 25 years of the German national rheumatology database. *Zeitschrift fur Rheumatologie.* 2019.
264. Garcia Rodriguez, L., Tolosa, L., Ruigomez, A., Johansson, S. and Wallander, M.A. Rheumatoid arthritis in UK primary care: incidence and prior morbidity. *Scandinavian Journal of Rheumatology.* 2009, **38**(3), pp.173-177.
265. Abhishek, A., Doherty, M., Kuo, C.F., Mallen, C.D., Zhang, W. et al. Rheumatoid arthritis is getting less frequent-results of a nationwide population-based cohort study. *Rheumatology (Oxford).* 2017, **56**(5), pp.736-744.
266. Silman, A., Bankhead, C., Rowlingson, B., Brennan, P., Symmons, D. et al. Do new cases of rheumatoid arthritis cluster in time or in space? *Int J Epidemiol.* 1997, **26**(3), pp.628-634.
267. Weyand, C.M., Hicok, K.C., Conn, D.L. and Goronzy, J.J. The influence of HLA-DRB1 genes on disease severity in rheumatoid arthritis. *Annals of Internal Medicine.* 1992, **117**(10), pp.801-806.
268. Gorman, J.D., Lum, R.F., Chen, J.J., Suarez-Almazor, M.E., Thomson, G. et al. Impact of shared epitope genotype and ethnicity on erosive disease: a meta-analysis of 3,240 rheumatoid arthritis patients. *Arthritis & Rheumatism.* 2004, **50**(2), pp.400-412.
269. Sugiyama, D., Nishimura, K., Tamaki, K., Tsuji, G., Nakazawa, T. et al. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. *Annals of the Rheumatic Diseases.* 2010, **69**(01), pp.70-81.
270. Page, C., Francois, C., Goëb, V. and Duverlie, G. Human parvovirus B19 and autoimmune diseases. Review of the literature and pathophysiological hypotheses. *Journal of Clinical Virology.* 2015, **72**, pp.69-74.
271. Kakurina, N., Kadisa, A., Lejnieks, A., Mikazane, H., Kozireva, S. et al. Use of exploratory factor analysis to ascertain the correlation between the activities of rheumatoid arthritis and infection by human parvovirus B19. *Medicina.* 2015, **51**(1), pp.18-24.
272. Hu, Y., Costenbader, K.H., Gao, X., Hu, F.B., Karlson, E.W. et al. Mediterranean diet and incidence of rheumatoid arthritis in women. *Arthritis Care & Research.* 2015, **67**(5), pp.597-606.
273. Joffe, I. and Epstein, S. Osteoporosis associated with rheumatoid arthritis: pathogenesis and management. In: *Seminars in Arthritis and Rheumatism*: Elsevier, 1991, pp.256-272.
274. Han, C., Robinson, D.W., Hackett, M.V., Paramore, L.C., Fraeman, K.H. et al. Cardiovascular disease and risk factors in patients with rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. *The Journal of Rheumatology.* 2006, **33**(11), pp.2167-2172.
275. Panoulas, V.F., Metsios, G.S., Pace, A., John, H., Treharne, G. et al. Hypertension in rheumatoid arthritis. *Rheumatology (Oxford).* 2008, **47**(9), pp.1286-1298.

276. Janssens, H.J., Arts, P.G., Schalk, B.W. and Biermans, M.C. Gout and rheumatoid arthritis, both to keep in mind in cardiovascular risk management: A primary care retrospective cohort study. *Joint Bone Spine.* 2017, **84**(1), pp.59-64.
277. Costello, R.E., Marsden, A., Movahedi, M., Lunt, M., Humphreys, J.H. et al. The effect of glucocorticoid therapy on mortality in patients with rheumatoid arthritis and concomitant type II diabetes: a retrospective cohort study. *BMC Rheumatology.* 2020, **4**(1), pp.1-8.
278. Hippisley-Cox, J., Coupland, C., Vinogradova, Y., Robson, J., May, M. et al. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ.* 2007, **335**(7611), p.136.
279. Lu, M.-C., Guo, H.-R., Lin, M.-C., Livneh, H., Lai, N.-S. et al. Bidirectional associations between rheumatoid arthritis and depression: a nationwide longitudinal study. *Scientific Reports.* 2016, **6**(1), pp.1-7.
280. National Institute for Health and Care Excellence. *NICE Guideline [NG100] Rheumatoid arthritis in adults: management.* [Online]. 2018. [Accessed 30 April 2019]. Available from: https://www.nice.org.uk/guidance/ng100/chapter/Recommendations#investigations
281. Michelsen, B., Kristianslund, E.K., Sexton, J., Hammer, H.B., Fagerli, K.M. et al. Do depression and anxiety reduce the likelihood of remission in rheumatoid arthritis and psoriatic arthritis? Data from the prospective multicentre NOR-DMARD study. *Annals of the Rheumatic Diseases.* 2017, **76**(11), pp.1906-1910.
282. Versini, M., Jeandel, P.-Y., Rosenthal, E. and Shoenfeld, Y. Obesity in autoimmune diseases: not a passive bystander. *Autoimmunity Reviews.* 2014, **13**(9), pp.981-1000.
283. Ajeganova, S., Andersson, M.L., Hafström, I. and Group, B.S. Association of obesity with worse disease severity in rheumatoid arthritis as well as with comorbidities: a long-term followup from disease onset. *Arthritis Care & Research.* 2013, **65**(1), pp.78-87.
284. Shan, J. and Zhang, J. Impact of obesity on the efficacy of different biologic agents in inflammatory diseases: a systematic review and meta-analysis. *Joint Bone Spine.* 2019, **86**(2), pp.173-183.
285. Barnett, K., Mercer, S.W., Norbury, M., Watt, G., Wyke, S. et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet.* 2012, **380**(9836), pp.37-43.
286. Scheelbeek, P.F., Cornelsen, L., Marteau, T.M., Jebb, S.A. and Smith, R.D. Potential impact on prevalence of obesity in the UK of a 20% price increase in high sugar snacks: modelling study. *BMJ.* 2019, **366**, p.l4786.
287. Symmons, D., Turner, G., Webb, R., Asten, P., Barrett, E. et al. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology (Oxford).* 2002, **41**(7), pp.793-800.
288. Aletaha, D., Neogi, T., Silman, A.J., Funovits, J., Felson, D.T. et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis & Rheumatism.* 2010, **62**(9), pp.2569-2581.
289. Allen, A., Carville, S. and McKenna, F. Diagnosis and management of rheumatoid arthritis in adults: summary of updated NICE guidance. *BMJ.* 2018, **362**, p.k3015.
290. Kyburz, D., Gabay, C., Michel, B.A. and Finckh, A. The long-term impact of early treatment of rheumatoid arthritis on radiographic progression: a population-based cohort study. *Rheumatology (Oxford).* 2011, **50**(6), pp.1106-1110.
291. Smolen, J.S., Breedveld, F.C., Burmester, G.R., Bykerk, V., Dougados, M. et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of

an international task force. *Annals of the Rheumatic Diseases.* 2016, **75**(1), pp.3-15.
292. Buer, J.K. A history of the term “DMARD”. *Inflammopharmacology.* 2015, **23**(4), pp.163-171.
293. Singh, J.A., Saag, K.G., Bridges Jr, S.L., Akl, E.A., Bannuru, R.R. et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis & Rheumatology.* 2016, **68**(1), pp.1-26.
294. Edwards, C.J., Campbell, J., van Staa, T. and Arden, N.K. Regional and temporal variation in the treatment of rheumatoid arthritis across the UK: a descriptive register-based cohort study. *BMJ Open.* 2012, **2**(6), p.e001603.
295. Conaghan, P.G. A turbulent decade for NSAIDs: update on current concepts of classification, epidemiology, comparative efficacy, and toxicity. *Rheumatology International.* 2012, **32**(6), pp.1491-1502.
296. Silverstein, F.E., Faich, G., Goldstein, J.L., Simon, L.S., Pincus, T. et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. *JAMA.* 2000, **284**(10), pp.1247-1255.
297. Hoes, J., Jacobs, J., Boers, M., Boumpas, D., Buttgereit, F. et al. EULAR evidence-based recommendations on the management of systemic glucocorticoid therapy in rheumatic diseases. *Annals of the Rheumatic Diseases.* 2007, **66**(12), pp.1560-1567.
298. Boers, M., Verhoeven, A. and van der Linden, S. Combination therapy in early rheumatoid arthritis: the COBRA study. *Nederlands Tijdschrift Voor Geneeskunde.* 1997, **141**(50), pp.2428-2432.
299. Hippisley-Cox, J. and Coupland, C. Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. *BMJ.* 2009, **339**, p.b4229.
300. National Institute for Health and Care Excellence. *Corticosteroids - oral.* [Online]. 2017. [Accessed 23 October 2019]. Available from: https://cks.nice.org.uk/osteoporosis-prevention-of-fragility-fractures#!scenario:1
301. Trijau, S., de Lamotte, G., Pradel, V., Natali, F., Allaria-Lapierre, V. et al. Osteoporosis prevention among chronic glucocorticoid users: results from a public health insurance database. *RMD Open.* 2016, **2**(2), p.e000249.
302. National Institute for Health and Care Excellence. *Osteoporosis - prevention of fragility fractures.* [Online]. 2016. [Accessed 23 October 2019]. Available from: https://cks.nice.org.uk/osteoporosis-prevention-of-fragility-fractures#!scenario:1
303. Nagata, N., Niikura, R., Yamada, A., Sakurai, T., Shimbo, T. et al. Acute middle gastrointestinal bleeding risk associated with NSAIDs, antithrombotic drugs, and PPIs: a multicenter case-control study. *PLoS One.* 2016, **11**(3).
304. Tran, G., Gough, A. and Emery, P. *Yorkshire rheumatology regional guidelines for the monitoring of adult patients on conventional disease modifying drugs, biologic drugs and targeted synthetic drugs: Version 7.* [Online]. 2019. [Accessed 02 July 2019]. Available from: https://www.bradfordhospitals.nhs.uk/wp-content/uploads/2019/07/YORKSHIRE-DMARD-GUIDELINES-2019-FINAL.pdf
305. Lin, C., Karlson, E.W., Canhao, H., Miller, T.A., Dligach, D. et al. Automatic prediction of rheumatoid arthritis disease activity from the electronic medical records. *PLoS One.* 2013, **8**(8), p.e69932.
306. Tennis, P., Bombardier, C., Malcolm, E. and Downey, W. Validity of rheumatoid arthritis diagnoses listed in the Saskatchewan Hospital Separations Database. *Journal of Clinical Epidemiology.* 1993, **46**(7), pp.675-683.
307. Allebeck, P., Ljungström, K. and Allander, E. Rheumatoid arthritis in a medical information system: How valid is the diagnosis? *Scandinavian Journal of Social Medicine.* 1983, **11**(1), pp.27-32.
308. Chiu, H.-Y., Huang, H.-L., Li, C.-H., Chen, H.-A., Yeh, C.-L. et al. Increased risk of chronic kidney disease in rheumatoid arthritis associated with cardiovascular

complications–a national population-based cohort study. *PLoS One.* 2015, **10**(9), p.e0136508.
309. Zhang, J., Chen, L., Delzell, E., Muntner, P., Hillegass, W.B. et al. The association between inflammatory markers, serum lipids and the risk of cardiovascular events in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2014, **73**(7), pp.1301-1308.
310. Li, L.M., Tessier-Cloutier, B., Wang, Y., Bernatsky, S., Vinet, E. et al. Assessing process of care in rheumatoid arthritis at McGill University hospitals. *Journal of Clinical Rheumatology.* 2013, **19**(4), pp.175-179.
311. Desai, S.S., Myles, J.D. and Kaplan, M.J. Suboptimal cardiovascular risk factor identification and management in patients with rheumatoid arthritis: a cohort analysis. *Arthritis Research & Therapy.* 2012, **14**(6), p.R270.
312. Rohr, M.K., Mikuls, T.R., Cohen, S.B., Thorne, J.C. and O'Dell, J.R. Underuse of methotrexate in the treatment of rheumatoid arthritis: a national analysis of prescribing practices in the US. *Arthritis Care & Research.* 2017, **69**(6), pp.794-800.
313. Weycker, D., Elaine, B.Y., Woolley, J.M. and Oster, G. Retrospective study of the costs of care during the first year of therapy with etanercept or infliximab among patients aged≥ 65 years with rheumatoid arthritis. *Clinical Therapeutics.* 2005, **27**(5), pp.646-656.
314. Godot, S., Gottenberg, J.E., Paternotte, S., Pane, I., Combe, B. et al. Safety of surgery after rituximab therapy in 133 patients with rheumatoid arthritis: data from the autoimmunity and rituximab registry. *Arthritis Care & Research.* 2013, **65**(11), pp.1874-1879.
315. Methley, A.M., Campbell, S., Chew-Graham, C., McNally, R. and Cheraghi-Sohi, S.J.B.h.s.r. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. 2014, **14**(1), pp.1-10.
316. *Publish or Perish, Version 5.29.5793.* Harzing, A. W.; Tarma Software Research, 2017.
317. Wallace, B.C., Small, K., Brodley, C.E., Lau, J. and Trikalinos, T.A. Deploying an interactive machine learning system in an evidence-based practice center: abstrackr. In: *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*: ACM, 2012, pp.819-824.
318. Ouzzani, M., Hammady, H., Fedorowicz, Z. and Elmagarmid, A. Rayyan - A web and mobile app for systematic reviews. *Systematic Reviews.* 2016, (5), p.210.
319. Collaboration, C. *Covidence.* [Online]. 2017. [Accessed 07 June 2017]. Available from: https://www.covidence.org
320. Bramer, W.M., Giustini, D., de Jonge, G.B., Holland, L. and Bekhuis, T. De-duplication of database search results for systematic reviews in EndNote. *Journal of the Medical Library Association.* 2016, **104**(3), p.240.
321. Schinasi, L.H., Auchincloss, A.H., Forrest, C.B. and Diez Roux, A.V. Using electronic health record data for environmental and place based population health research: a systematic review. *Ann Epidemiol.* 2018, **28**(7), pp.493-502.
322. Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C. et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Annals of Internal Medicine.* 2007, **147**(8), pp.573-577.
323. Clinical Practice Research Datalink. *CPRD: Benefits for Partner Practices.* [Online]. 2016. [Accessed 29 July 2016]. Available from: https://www.cprd.com/researchpractice/researchgppractice.asp#BenefitsforPartnerPractices
324. Padmanabhan, S. *CPRD GOLD Data Specification Version 1.5.* 2013.
325. NHS Digital. *Technology Reference data Update Distribution* [Online]. 2016. [Accessed 18 May 2016]. Available from: https://isd.digital.nhs.uk/trud3/user/guest/group/0/home

326. National Institute for Health and Care Excellence. *TNF-alpha inhibitors for ankylosing spondylitis and non-radiographic axial spondyloarthritis: Technology appraisal guidance [TA383].* [Online]. 2016. [Accessed 17 August 2017]. Available from: https://www.nice.org.uk/guidance/ta383/chapter/2-Clinical-need-and-practice
327. Gardiner, J., Su, B., Ellis, B. and Soljak, M. Estimating Under-Diagnosis of Rheumatoid Arthritis in Primary Care Data from the UK Clinical Practice Research Datalink. *2015 American College of Rheumatology Annual Meeting.* 2015, (1228).
328. Wallace, K.L., Riedel, A.A., Joseph-Ridge, N. and Wortmann, R. Increasing prevalence of gout and hyperuricemia over 10 years among older adults in a managed care population. *The Journal of Rheumatology.* 2004, **31**(8), pp.1582-1587.
329. Kvien, T.K., Uhlig, T., Ødegård, S. and Heiberg, M.S. Epidemiological aspects of rheumatoid arthritis: the sex ratio. *Annals of the New York Academy of Sciences.* 2006, **1069**(1), pp.212-222.
330. Clinical Practice Research Datalink. *Clinical Practice Research Datalink: Data access.* [Online]. 2021. [Accessed 28 January 2021]. Available from: https://cprd.com/Data-access
331. University of Leeds. *Research Data Management Planning.* [Online]. 2017. [Accessed 07 March 2017]. Available from: https://library.leeds.ac.uk/research-data-management-planning
332. Digital Curation Centre. *DMPOnline.* [Online]. 2017. [Accessed 07 March 2017]. Available from: http://www.dcc.ac.uk/dmponline
333. Keat, A., Gaffney, K., Marzo-Ortega, H., Cornell, T., MacKay, K. et al. Improving the treatment of ankylosing spondylitis in the UK. *Rheumatology (Oxford).* 2011, **50**(11), pp.1936-1939.
334. Pichardo-Lowden, A.R. and Haidet, P.M. Closing the Loop: Optimizing Diabetes Care in the Hospital by Addressing Dispersed Information in Electronic Health Records and Using Clinical Decision Support. *J Diabetes Sci Technol.* 2018, p.1932296818817005.
335. Oderda, G.M., Shiozawa, A., Walsh, M., Hess, K., Brixner, D.I. et al. Physician adherence to ACR gout treatment guidelines: perception versus practice. *Postgraduate Medicine.* 2014, **126**(3), pp.257-267.
336. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. and Group, P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009, **339**, p.b2535.
337. Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care.* [Online]. 2009. [Accessed 23.03.2017]. Available from: https://www.york.ac.uk/crd/guidance/
338. Denison, H.J., Dodds, R., Ntani, G., Cooper, R., Cooper, C. et al. How to get started with a systematic review in epidemiology: an introductory guide for early career researchers. *Archives of Public Health.* 2013, **71**(21).
339. Crossfield, S., Lai, L. and Kingsbury, S. Systematic review of variation in gout medication exposure and its definition and measurement in studies using electronic health records data. *PROSPERO.* 2017.
340. Al Sallakh, M.A., Vasileiou, E., Rodgers, S.E., Lyons, R.A., Sheikh, A. et al. Defining asthma and assessing asthma outcomes using electronic health record data: a systematic scoping review. *European Respiratory Journal.* 2017, **49**(6), p.1700204.
341. Blanch, B., Sweeting, J., Semsarian, C. and Ingles, J. Routinely collected health data to study inherited heart disease: a systematic review (2000–2016). *Open Heart.* 2017, **4**(2), p.e000686.
342. Payet, C., Lifante, J.-C., Carty, M.J., Rabilloud, M. and Duclos, A. Methodological quality of surgical mortality studies using large hospital databases: a systematic review. *Annals of Surgery.* 2017, **265**(6), pp.1113-1118.

343. Critical Appraisal Skills Programme. *CASP Cohort Study Checklist.* [Online]. 2017. [Accessed 20 January 2017]. Available from: http://www.casp-uk.net/checklists
344. Wallace, S.L., Robinson, H., Masi, A.T., Decker, J.L., Mccarty, D.J. et al. Preliminary criteria for the classification of the acute arthritis of primary gout. *Arthritis & Rheumatism.* 1977, **20**(3), pp.895-900.
345. Benchimol, E.I., Smeeth, L., Guttmann, A., Harron, K., Moher, D. et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine.* 2015, **12**(10), p.e1001885.
346. Cochrane Collaboration. *Tool to Assess Risk of Bias in Cohort Studies.* [Online]. 2017. [Accessed 22.01.2018]. Available from: http://methods.cochrane.org/bias/sites/methods.cochrane.org.bias/files/public/uploads/Tool%20to%20Assess%20Risk%20of%20Bias%20in%20Cohort%20Studies.pdf
347. Barry, E., Roberts, S., Oke, J., Vijayaraghavan, S., Normansell, R. et al. Efficacy and effectiveness of screen and treat policies in prevention of type 2 diabetes: systematic review and meta-analysis of screening tests and interventions. *BMJ.* 2017, **356**, p.i6538.
348. Welk, B. and Kwong, J. A review of routinely collected data studies in urology: Methodological considerations, reporting quality, and future directions. *Canadian Urological Association Journal.* 2017, **11**(3-4), p.136.
349. Care Quality Commission. NHS Patient Survey Programme: Survey scoring method. [Online]. 2015. [Accessed 01 October 2018]. Available from: https://www.cqc.org.uk/sites/default/files/20151125_nhspatientsurveys_scoring_methodology.pdf
350. Ramke, J., Palagyi, A., Jordan, V., Petkovic, J. and Gilbert, C.E. Using the STROBE statement to assess reporting in blindness prevalence surveys in low and middle income countries. *PLoS One.* 2017, **12**(5), pp.e0176178-e0176178.
351. Whiting, P.F., Rutjes, A.W., Westwood, M.E., Mallett, S., Deeks, J.J. et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine.* 2011, **155**(8), pp.529-536.
352. Mikuls, T.R., Cheetham, T.C., Levy, G.D. and Rashid, N. Adherence and Outcomes with Urate-Lowering Therapy: A Site-Randomized Trial. *The American Journal of Medicine.* 2019, pp.354-361.
353. Keenan, R.T., O'Brien, W.R., Lee, K., Crittenden, D.B., Fisher, M.C. et al. Prevalence of contraindications and prescription of pharmacologic therapies for gout. *The American Journal of Medicine.* 2011, **124**(2), pp.155-163.
354. Kuo, C.-F., Grainge, M.J., Mallen, C., Zhang, W. and Doherty, M. Eligibility for and prescription of urate-lowering treatment in patients with incident gout in England. *JAMA.* 2014, **312**(24), pp.2684-2686.
355. Scheepers, L.E.J.M., Burden, A.M., Arts, I.C.W., Spaetgens, B., Souverein, P. et al. Medication adherence among gout patients initiated allopurinol: a retrospective cohort study in the Clinical Practice Research Datalink (CPRD). *Rheumatology (Oxford).* 2018, **57**(9), pp.1641-1650.
356. Clarson, L.E., Hider, S.L., Belcher, J., Roddy, E. and Mallen, C.D. Factors influencing allopurinol initiation in primary care. *Ann Fam Med.* 2017, **15**(6), pp.557-560.
357. Roddy, E., Mallen, C.D., Hider, S.L. and Jordan, K.P. Prescription and comorbidity screening following consultation for acute gout in primary care. *Rheumatology (Oxford).* 2010, **49**(1), pp.105-111.
358. Dehlin, M., Ekström, E.H., Petzold, M., Strömberg, U., Telg, G. et al. Factors associated with initiation and persistence of urate-lowering therapy. *Arthritis Research & Therapy.* 2017, **19**(1), p.6.
359. Kuo, C.-F., Grainge, M.J., Mallen, C., Zhang, W. and Doherty, M. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. *Annals of the Rheumatic Diseases.* 2014, **74**(4), pp.661-667.

360. Sigurdardottir, V., Drivelegka, P., Svard, A., Jacobsson, L.T.H. and Dehlin, M. Work disability in gout: a population-based case-control study. *Annals of the Rheumatic Diseases.* 2017, **77**(3), pp.399-404.
361. Jain, R., Dasari, S., Soriano, T., Decherrie, L. and Kerr, L.D. Rheumatologists on the road: a subspecialist's role in caring for the homebound. *Arthritis Care & Research.* 2011, **63**(10), pp.1482-1485.
362. Kapetanovic, M.C., Hameed, M., Turkiewicz, A., Neogi, T., Saxne, T. et al. Prevalence and incidence of gout in southern Sweden from the socioeconomic perspective. *RMD Open.* 2016, **2**(2), p.e000326.
363. Rai, S.K., Avina-Zubieta, J.A., McCormick, N., De Vera, M.A., Shojania, K. et al. The rising prevalence and incidence of gout in British Columbia, Canada: Population-based trends from 2000 to 2012. *Seminars in Arthritis and Rheumatism.* 2017, **46**(4), pp.451-456.
364. Fisher, M.C., Rai, S.K., Lu, N., Zhang, Y.Q. and Choi, H.K. The unclosing premature mortality gap in gout: a general population-based study. *Annals of the Rheumatic Diseases.* 2017, **76**(7), pp.1289-1294.
365. Wahedduddin, S., Singh, J.A., Culhane-Pera, K.A. and Gertner, E. Gout in the Hmong in the United States. *J Clin Rheumatol.* 2010, **16**(6), pp.262-266.
366. Rashid, N., Coburn, B.W., Wu, Y.-L., Cheetham, T.C., Curtis, J.R. et al. Modifiable factors associated with allopurinol adherence and outcomes among patients with gout in an integrated healthcare system. *J Rheumatol.* 2014, **42**(3), pp.504-512.
367. Harrold, L.R., Andrade, S.E., Briesacher, B., Raebel, M.A., Fouayzi, H. et al. The Dynamics of Chronic Gout Treatment: Medication Gaps and Return to Therapy. *Am J Med.* 2010, **123**(1), pp.54-59.
368. Singh, J.A., Hodges, J.S. and Asch, S.M. Opportunities for improving medication use and monitoring in gout. *Annals of the Rheumatic Diseases.* 2008, **68**(8), pp.1265-1270.
369. Zandman-Goddard, G., Amital, H., Shamrayevsky, N., Raz, R., Shalev, V. et al. Rates of adherence and persistence with allopurinol therapy among gout patients in Israel. *Rheumatology (Oxford).* 2013, **52**(6), pp.1126-1131.
370. Mantarro, S., Capogrosso-Sansone, A., Tuccori, M., Blandizzi, C., Montagnani, S. et al. Allopurinol adherence among patients with gout: an Italian general practice database study. *Int J Clin Pract.* 2015, **69**(7), pp.757-765.
371. Coburn, B.W., Michaud, K., Bergman, D.A. and Mikuls, T.R. Allopurinol Dose Escalation and Mortality Among Patients With Gout: A National Propensity-Matched Cohort Study. *Arthritis Rheumatol.* 2018, **70**(8), pp.1298-1307.
372. Coburn, B.W., Bendlin, K.A., Sayles, H., Meza, J., Russell, C.L. et al. Allopurinol Medication Adherence as a Mediator of Optimal Outcomes in Gout Management. *J Clin Rheumatol.* 2017, **23**(6), pp.317-323.
373. Hughes, J.C., Wallace, J.L., Bryant, C.L., Salvig, B.E., Fourakre, T.N. et al. Monitoring of Urate-Lowering Therapy Among US Veterans Following the 2012 American College of Rheumatology Guidelines for Management of Gout. *Ann Pharmacother.* 2017, **51**(4), pp.301-306.
374. Sultan, A.A., Mallen, C., Hayward, R., Muller, S., Whittle, R. et al. Gout and subsequent erectile dysfunction: a population- based cohort study from England. *Arthritis Research & Therapy.* 2017, **19**(1), p.123.
375. Maravic, M., Hincapie, N., Pilet, S., Flipo, R.M. and Lioté, F. Persistent clinical inertia in gout in 2014: An observational French longitudinal patient database study. *Joint Bone Spine.* 2018, **85**(3), pp.311-315.
376. Landgren, A.J., Jacobsson, L.T.H., Lindstrom, U., Sandstrom, T.Z.S., Drivelegka, P. et al. Incidence of and risk factors for nephrolithiasis in patients with gout and the general population, a cohort study. *Arthritis Research & Therapy.* 2017, **19**(1).
377. Pui, K., Gow, P.J. and Dalbeth, N. Efficacy and tolerability of probenecid as urate-lowering therapy in gout; clinical experience in high-prevalence population. *J Rheumatol.* 2013, **40**(6), pp.872-876.

378. Morlock, R., Chevalier, P., Horne, L., Nuevo, J., Storgard, C. et al. Disease Control, Health Resource Use, Healthcare Costs, and Predictors in Gout Patients in the United States, the United Kingdom, Germany, and France: A Retrospective Analysis. *Rheumatol Ther.* 2016, **3**(1), pp.53-75.
379. Ryu, H., Song, R., Kim, H., Kim, J., Lee, E.Y. et al. Clinical risk factors for adverse events in allopurinol users. *J Clin Pharmacol.* 2013, **53**(2), pp.211-216.
380. Hatoum, H., Khanna, D., Lin, S.J., Akhras, K.S., Shiozawa, A. et al. Achieving Serum Urate Goal: A Comparative Effectiveness Study Between Allopurinol and Febuxostat. *Postgraduate Medicine.* 2014, **126**(2), pp.65-75.
381. Singh, J.A., Hodges, J.S., Toscano, J.P. and Asch, S.M. Quality of care for gout in the US needs improvement. *Arthritis & Rheumatism.* 2007, **57**(5), pp.822-829.
382. Janssen, C.A., Jansen, T., Voshaar, M., Vonkeman, H.E. and van de Laar, M. Quality of care in gout: a clinical audit on treating to the target with urate lowering therapy in real-world gout patients. *Rheumatol Int.* 2017, **37**(9), pp.1435-1440.
383. Thueringer, J.T., Doll, N.K. and Gertner, E. Anakinra for the treatment of acute severe gout in critically ill patients. *Seminars in Arthritis and Rheumatism.* 2015, **45**(1), pp.81-85.
384. Crittenden, D.B., Lehmann, R.A., Schneck, L., Keenan, R.T., Shah, B. et al. Colchicine Use Is Associated with Decreased Prevalence of Myocardial Infarction in Patients with Gout. *J Rheumatol.* 2012, **39**(7), pp.1458-1464.
385. Solomon, D.H., Liu, C.C., Kuo, I.H., Zak, A. and Kim, S.C. Effects of colchicine on risk of cardiovascular events and mortality among patients with gout: a cohort study using electronic medical records linked with Medicare claims. *Annals of the Rheumatic Diseases.* 2016, **75**(9), pp.1674-1679.
386. Kwon, O.C., Hong, S., Ghang, B., Kim, Y.G., Lee, C.K. et al. Risk of Colchicine-Associated Myopathy in Gout: Influence of Concomitant Use of Statin. *The American Journal of Medicine.* 2016, **130**(5), pp.583-587.
387. Spaetgens, B., de Vries, F., Driessen, J.H.M., Leufkens, H.G., Souverein, P.C. et al. Risk of infections in patients with gout: a population-based cohort study. *Scientific Reports.* 2017, **7**.
388. Sultan, A.A., Whittle, R., Muller, S., Roddy, E., Mallen, C.D. et al. Risk of fragility fracture among patients with gout and the effect of urate-lowering therapy. *Canadian Medical Association Journal.* 2018, **190**(19), pp.E581-E587.
389. Dennison, E.M., Rubin, K.H., Schwarz, P., Harvey, N.C., Bone, K.W. et al. Is allopurinol use associated with an excess risk of osteoporotic fracture? A National Prescription Registry study. *Arch Osteoporos.* 2015, **10**(36).
390. Kuo, C.-F., Chou, I.-J., See, L.-C., Chen, J.-S., Yu, K.-H. et al. Urate-lowering treatment and risk of total joint replacement in patients with gout. *Rheumatology (Oxford).* 2018, **57**(12), pp.2129-2139.
391. Kuo, C.F., Grainge, M.J., Mallen, C., Zhang, W.Y. and Doherty, M. Effect of allopurinol on all-cause mortality in adults with incident gout: propensity score-matched landmark analysis. *Rheumatology (Oxford).* 2015, **54**(12), pp.2145-2150.
392. Meier, C.R. and Jick, H. Omeprazole, other antiulcer drugs and newly diagnosed gout. *Br J Clin Pharmacol.* 1997, **44**, p.4.
393. Huang, I.J., Liew, J.W., Morcos, M.B., Zuo, S., Crawford, C. et al. Pharmacist-managed titration of urate-lowering therapy to streamline gout management. *Rheumatology International.* 2019, **39**(9), pp.1637-1641.
394. Fu, Y., Yan, M., Yang, H., Ma, X. and Guo, J. Palm-Sized Uric Acid Test Lab Powered by Smartphone for Proactive Gout Management. *IEEE Transactions on Biomedical Circuits and Systems.* 2019, **13**(5), pp.950-956.
395. Doherty, M., Jenkins, W., Richardson, H., Sarmanova, A., Abhishek, A. et al. Efficacy and cost-effectiveness of nurse-led care involving education and engagement of patients and a treat-to-target urate-lowering strategy versus

usual care for gout: a randomised controlled trial. *The Lancet.* 2018, **392**(10156), pp.1403-1412.
396. Lewis, J.D., Bilker, W.B., Weinstein, R.B. and Strom, B.L. The relationship between time since registration and measured incidence rates in the General Practice Research Database. *Pharmacoepidemiology and Drug Safety.* 2005, **14**(7), pp.443-451.
397. Sterne, J.A., Hernan, M.A., Reeves, B.C., Savovic, J., Berkman, N.D. et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ.* 2016, **355**, p.i4919.
398. Wells, G.A., Shea, B., O'Connell, D.a., Peterson, J., Welch, V. et al. *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Oxford. 2000.
399. Greenland, S., Mansournia, M.A. and Altman, D.G. Sparse data bias: a problem hiding in plain sight. *BMJ.* 2016, **352**, p.i1981.
400. Brophy, S., Cooksey, R., Atkinson, M., Zhou, S.M., Husain, M.J. et al. No increased rate of acute myocardial infarction or stroke among patients with ankylosing spondylitis-a retrospective cohort study using routine data. *Seminars in Arthritis and Rheumatism.* 2012, **42**(2), pp.140-145.
401. Husain, M.J., Brophy, S., Macey, S., Pinder, L.M., Atkinson, M.D. et al. HERALD (health economics using routine anonymised linked data). *BMC Med Inform Decis Mak.* 2012, **12**, p.24.
402. Dean, L.E., Macfarlane, G.J. and Jones, G.T. Differences in the prevalence of ankylosing spondylitis in primary and secondary care: only one-third of patients are managed in rheumatology. *Rheumatology (Oxford).* 2016, **55**(10), pp.1820-1825.
403. Ramagopalan, S.V., Goldacre, R., Skingsley, A., Conlon, C. and Goldacre, M.J. Associations between selected immune-mediated diseases and tuberculosis: record-linkage studies. *BMC Med.* 2013, **11**, p.97.
404. Edson-Heredia, E., Zhu, B., Lefevre, C., Wang, M., Barrett, A. et al. Prevalence and incidence rates of cardiovascular, autoimmune, and other diseases in patients with psoriatic or psoriatic arthritis: a retrospective study using Clinical Practice Research Datalink. *Journal of the European Academy of Dermatology & Venereology.* 2015, **29**(5), pp.955-963.
405. Card, T.R., Langan, S.M. and Chu, T.P. Extra-Gastrointestinal Manifestations of Inflammatory Bowel Disease May Be Less Common Than Previously Reported. *Dig Dis Sci.* 2016, **61**(9), pp.2619-2626.
406. Ramagopalan, S.V., Goldacre, R., Disanto, G., Giovannoni, G. and Goldacre, M.J. Hospital admissions for vitamin D related conditions and subsequent immune-mediated disease: record-linkage studies. *BMC Med.* 2013, **11**, p.171.
407. Cooksey, R., Brophy, S., Dennis, M., Davies, H., Atkinson, M. et al. Severe flare as a predictor of poor outcome in ankylosing spondylitis: a cohort study using questionnaire and routine data linkage. *Rheumatology (Oxford).* 2015, **54**(9), pp.1563-1572.
408. Yeoh, D., Moffatt, T. and Karmani, S. Good outcomes of percutaneous fixation of spinal fractures in ankylosing spinal disorders. *Injury.* 2014, **45**(10), pp.1534-1538.
409. Ramagopalan, S.V., Pakpoor, J., Seminog, O., Goldacre, R., Graham, L. et al. Risk of subarachnoid haemorrhage in people admitted to hospital with selected immune-mediated diseases: record-linkage studies. *BMC Neurol.* 2013, **13**, p.176.
410. Essers, I., Stolwijk, C., Boonen, A., De Bruin, M.L., Bazelier, M.T. et al. Ankylosing spondylitis and risk of ischaemic heart disease: a population-based cohort study. *Annals of the Rheumatic Diseases.* 2016, **75**(1), pp.203-209.
411. Stolwijk, C., Essers, I., van Tubergen, A., Boonen, A., Bazelier, M.T. et al. The epidemiology of extra-articular manifestations in ankylosing spondylitis: a population-based matched cohort study. *Ann Rheum Dis.* 2015, **74**(7), pp.1373-1378.

412. Singh, J.A., Holmgren, A.R., Krug, H. and Noorbaloochi, S. Accuracy of the diagnoses of spondylarthritides in veterans affairs medical center databases. *Arthritis & Rheumatism.* 2007, **57**(4), pp.648-655.
413. Dostal, C., Pavelka, K., Zvarova, J., Hanzlicek, P. and Olejarova, M. Some principles of the development of a clinical database/national register of selected inflammatory rheumatic diseases in the Czech Republic. *Int J Med Inform.* 2006, **75**(3), pp.216-223.
414. Westerveld, L.A., van Bemmel, J.C., Dhert, W.J., Oner, F.C. and Verlaan, J.J. Clinical outcome after traumatic spinal fractures in patients with ankylosing spinal disorders compared with control patients. *The Spine Journal.* 2014, **14**(5), pp.729-740.
415. Backhaus, M., Citak, M., Kalicke, T., Sobottke, R., Russe, O. et al. Spine fractures in patients with ankylosing spondylitis: an analysis of 129 fractures after surgical treatment. *Orthopade.* 2011, **40**(10), pp.917-920, 922-914.
416. Jarvinen, P. Occurrence of ankylosing spondylitis in a nationwide series of twins. *Arthritis & Rheumatism.* 1995, **38**(3), pp.381-383.
417. Carter, E.T., McKenna, C.H., Brian, D.D. and Kurland, L.T. Epidemiology of Ankylosing spondylitis in Rochester, Minnesota, 1935-1973. *Arthritis & Rheumatism.* 1979, **22**(4), pp.365-370.
418. Buschiazzo, E., Maldonado-Cocco, J.A., Arturi, P., Citera, G., Berman, A. et al. Epidemiology of spondyloarthritis in Argentina. *Am J Med Sci.* 2011, **341**(4), pp.289-292.
419. Keller, J.J., Kang, J.H. and Lin, H.C. Association between ankylosing spondylitis and chronic periodontitis: a population-based study. *Arthritis & Rheumatism.* 2013, **65**(1), pp.167-173.
420. Yang, J.J., Tsai, M.S., Sun, H.Y., Hsieh, S.M., Chen, M.Y. et al. Autoimmune diseases-related arthritis in HIV-infected patients in the era of highly active antiretroviral therapy. *J Microbiol Immunol Infect.* 2015, **48**(2), pp.130-136.
421. Kathuria, P., Gordon, K.B. and Silverberg, J.I. Association of psoriasis and psoriatic arthritis with osteoporosis and pathological fractures. *J Am Acad Dermatol.* 2017, **76**(6), pp.1045-1053.e1043.
422. Vander Cruyssen, B., Munoz-Gomariz, E., Font, P., Mulero, J., de Vlam, K. et al. Hip involvement in ankylosing spondylitis: epidemiology and risk factors associated with hip replacement surgery. *Rheumatology (Oxford).* 2009, **49**(1), pp.73-81.
423. Adomaviciute, D., Pileckyte, M., Baranauskaite, A., Morvan, J., Dadoniene, J. et al. Prevalence survey of rheumatoid arthritis and spondyloarthropathy in Lithuania. *Scandinavian Journal of Rheumatology.* 2008, **37**(2), pp.113-119.
424. Shen, T.C., Lin, C.L., Wei, C.C., Chen, C.H., Tu, C.Y. et al. The risk of asthma in patients with ankylosing spondylitis: a population-based cohort study. *PLoS One.* 2015, **10**(2), p.e0116608.
425. Lee, C.C., Lee, S.H., Chang, I.J., Lu, T.C., Yuan, A. et al. Spontaneous pneumothorax associated with ankylosing spondylitis. *Rheumatology (Oxford).* 2005, **44**(12), pp.1538-1541.
426. Jakobsen, A.K., Jacobsson, L.T., Patschan, O., Askling, J. and Kristensen, L.E. Is nephrolithiasis an unrecognized extra-articular manifestation in ankylosing spondylitis? A prospective population-based Swedish national cohort study with matched general population comparator subjects. *PLoS One.* 2014, **9**(11), p.e113602.
427. Herrinton, L.J., Liu, L., Chen, L., Harrold, L.R., Raebel, M.A. et al. Association between anti-TNF-alpha therapy and all-cause mortality. *Pharmacoepidemiology and Drug Safety.* 2012, **21**(12), pp.1311-1320.
428. Hellgren, K., Dreyer, L., Arkema, E.V., Glintborg, B., Jacobsson, L.T. et al. Cancer risk in patients with spondyloarthritis treated with TNF inhibitors: a collaborative study from the ARTIS and DANBIO registers. *Annals of the Rheumatic Diseases.* 2017, **76**(1), pp.105-111.

429. Broms, G., Granath, F., Ekbom, A., Hellgren, K., Pedersen, L. et al. Low Risk of Birth Defects for Infants Whose Mothers Are Treated With Anti-Tumor Necrosis Factor Agents During Pregnancy. *Clin Gastroenterol Hepatol.* 2016, **14**(2), pp.234-241.e231-235.
430. Hellgren, K., Smedby, K.E., Backlin, C., Sundstrom, C., Feltelius, N. et al. Ankylosing spondylitis, psoriatic arthritis, and risk of malignant lymphoma: a cohort study based on nationwide prospectively recorded data from Sweden. *Arthritis Rheumatol.* 2013, **66**(5), pp.1282-1290.
431. Hemminki, K., Li, X., Sundquist, J. and Sundquist, K. Familial associations of rheumatoid arthritis with autoimmune diseases and related conditions. *Arthritis & Rheumatism.* 2009, **60**(3), pp.661-668.
432. Sun, L.M., Muo, C.H., Liang, J.A., Chang, S.N., Sung, F.C. et al. Increased risk of cancer for patients with ankylosing spondylitis: a nationwide population-based retrospective cohort study. *Scandinavian Journal of Rheumatology.* 2014, **43**(4), pp.301-306.
433. Brown, L.M., Gridley, G., Check, D. and Landgren, O. Risk of multiple myeloma and monoclonal gammopathy of undetermined significance among white and black male United States veterans with prior autoimmune, infectious, inflammatory, and allergic disorders. *Blood.* 2008, **111**(7), pp.3388-3394.
434. Haroon, N.N., Paterson, J.M., Li, P., Inman, R.D. and Haroon, N. Patients With Ankylosing Spondylitis Have Increased Cardiovascular and Cerebrovascular Mortality: A Population-Based Study. *Ann Intern Med.* 2015, **163**(6), pp.409-416.
435. Keller, J.J., Hsu, J.L., Lin, S.M., Chou, C.C., Wang, L.H. et al. Increased risk of stroke among patients with ankylosing spondylitis: a population-based matched-cohort study. *Rheumatol Int.* 2014, **34**(2), pp.255-263.
436. Zoller, B., Li, X., Sundquist, J. and Sundquist, K. Risk of subsequent ischemic and hemorrhagic stroke in patients hospitalized for immune-mediated diseases: a nationwide follow-up study from Sweden. *BMC Neurology.* 2012, **12**, p.41.
437. Zhu, T.Y., Tam, L.S., Lee, V.W., Hwang, W.W., Li, T.K. et al. Costs and quality of life of patients with ankylosing spondylitis in Hong Kong. *Rheumatology (Oxford).* 2008, **47**(9), pp.1422-1425.
438. Russo, S., Mariani, T.T., Migliorini, R., Marcellusi, A. and Mennini, F.S. The economic burden of musculoskeletal disorders on the Italian social security pension system estimated by a Monte Carlo simulation. *Reumatismo.* 2015, **67**(2), pp.45-56.
439. Strombeck, B., Jacobsson, L.T., Bremander, A., Englund, M., Heide, A. et al. Patients with ankylosing spondylitis have increased sick leave--a registry-based case-control study over 7 yrs. *Rheumatology (Oxford).* 2009, **48**(3), pp.289-292.
440. Ariza-Ariza, R., Hernandez-Cruz, B., Collantes, E., Batlle, E., Fernandez-Sueiro, J.L. et al. Work disability in patients with ankylosing spondylitis. *J Rheumatol.* 2009, **36**(11), pp.2512-2516.
441. Kristensen, L.E., Petersson, I.F., Geborek, P., Joud, A., Saxne, T. et al. Sick leave in patients with ankylosing spondylitis before and after anti-TNF therapy: a population-based cohort study. *Rheumatology (Oxford).* 2012, **51**(2), pp.243-249.
442. Wu, N., Lee, Y.C., Shah, N. and Harrison, D.J. Cost of biologics per treated patient across immune-mediated inflammatory disease indications in a pharmacy benefit management setting: a retrospective cohort study. *Clin Ther.* 2014, **36**(8), pp.1231-1241, 1241.e1231-1233.
443. Nell-Duxneuner, V., Schroeder, Y., Reichardt, B. and Bucsics, A. The use of TNF-inhibitors in ankylosing spondylitis in Austria from 2007 to 2009 - a retrospective analysis. *Int J Clin Pharmacol Ther.* 2012, **50**(12), pp.867-872.
444. Neilson, A.R., Sieper, J. and Deeg, M. Cost-effectiveness of etanercept in patients with severe ankylosing spondylitis in Germany. *Rheumatology (Oxford).* 2010, **49**(11), pp.2122-2134.

445. Cemeroglu, O., Yasar, Z.S., Saglam, M. and Cakirbay, H. Clinical and demographic findings of patients with rheumatoid arthritis and ankylosing spondylitis treated in a tertiary care center in Turkey. *Turk J Med Sci.* 2014, **44**(4), pp.595-600.
446. Relas, H., Kautiainen, H., Puolakka, K., Virta, L.J. and Leirisalo-Repo, M. Survival of disease-modifying antirheumatic drugs used as the first antirheumatic medication in the treatment of ankylosing spondylitis in Finland. A nationwide population-based register study. *Clinical Rheumatology.* 2013, **33**(8), pp.1135-1138.
447. Dabes, C.G., Almeida, A.M. and Acurcio Fde, A. Non-adherence to biological therapy in patients with rheumatic diseases in the Brazilian Unified National Health System in Minas Gerais State, Brazil. *Cad Saude Publica.* 2015, **31**(12), pp.2599-2609.
448. Lyu, R., Govoni, M., Ding, Q., Black, C.M., Kachroo, S. et al. Treatment persistence among patients with rheumatoid disease (RA, AS, PsA) treated with subcutaneous biologics in Germany. *Rheumatol Int.* 2016, **36**(1), pp.143-153.
449. Lie, E., Lindstrom, U., , Z.-S., strom, T., Olsen, I.C. et al. Tumour necrosis factor inhibitor treatment and occurrence of anterior uveitis in ankylosing spondylitis: results from the Swedish biologics register. *Annals of the Rheumatic Diseases.* 2017, **76**(9), pp.1515-1521.
450. Vastesaeger, N., Cruyssen, B.V., Mulero, J., Gratacos Masmitja, J., Zarco, P. et al. ASDAS high disease activity versus BASDAI elevation in patients with ankylosing spondylitis as selection criterion for anti-TNF therapy. *Reumatol Clin.* 2014, **10**(4), pp.204-209.
451. Munoz-Ortego, J., Vestergaard, P., Rubio, J.B., Wordsworth, P., Judge, A. et al. Ankylosing spondylitis is associated with an increased risk of vertebral and nonvertebral clinical fractures: a population-based cohort study. *J Bone Miner Res.* 2014, **29**(8), pp.1770-1776.
452. Lu, M.L., Tsai, T.T., Lai, P.L., Fu, T.S., Niu, C.C. et al. A retrospective study of treating thoracolumbar spine fractures in ankylosing spondylitis. *Eur J Orthop Surg Traumatol.* 2014, **24**, pp.S117-123.
453. Moussallem, C.D., McCutcheon, B.A., Clarke, M.J., Cui, Q., Currier, B.L. et al. Perioperative complications in open versus percutaneous treatment of spinal fractures in patients with an ankylosed spine. *J Clin Neurosci.* 2016, **30**, pp.88-92.
454. Muller, S., Hider, S.L., Raza, K., Stack, R.J., Hayward, R.A. et al. An algorithm to identify rheumatoid arthritis in primary care: a Clinical Practice Research Datalink study. *BMJ Open.* 2015, **5**(12), p.e009309.
455. Zhou, S.M., Fernandez-Gutierrez, F., Kennedy, J., Cooksey, R., Atkinson, M. et al. Defining Disease Phenotypes in Primary Care Electronic Health Records by a Machine Learning Approach: A Case Study in Identifying Rheumatoid Arthritis. *PLoS One.* 2016, **11**(5), p.e0154515.
456. Ford, E., Nicholson, A., Koeling, R., Tate, A., Carroll, J. et al. Optimising the use of electronic health records to estimate the incidence of rheumatoid arthritis in primary care: what information is hidden in free text? *BMC Med Res Methodol.* 2013, **13**, p.105.
457. Nicholson, A., Ford, E., Davies, K., Smith, H., Rait, G. et al. Optimising use of electronic health records to describe the presentation of rheumatoid arthritis in primary care: a strategy for developing code lists. *PLoS One.* 2013.
458. Thomas, S.L., Edwards, C.J., Smeeth, L., Cooper, C. and Hall, A.J. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? *Arthritis & Rheumatism.* 2008, **59**(9), pp.1314-1321.
459. Ford, E., Carroll, J., Smith, H., Davies, K., Koeling, R. et al. What evidence is there for a delay in diagnostic coding of RA in UK general practice records? An observational study of free text. *BMJ Open.* 2016, **6**(6), p.e010393.

460. Rodriguez, L.A.G., Tolosa, L.B., Ruigomez, A., Johansson, S. and Wallander, M.A. Rheumatoid arthritis in UK primary care: incidence and prior morbidity. *Scandinavian Journal of Rheumatology.* 2009, **38**(3), pp.173-177.
461. Jordan, K., Clarke, A.M., Symmons, D.P., Fleming, D., Porcheret, M. et al. Measuring disease prevalence: a comparison of musculoskeletal disease using four general practice consultation databases. *Br J Gen Pract.* 2007, **57**(534), pp.7-14.
462. Tascilar, K., Dell'Aniello, S., Hudson, M. and Suissa, S. Statins and Risk of Rheumatoid Arthritis: A Nested Case-Control Study. *Arthritis Rheumatol.* 2016, **68**(11), pp.2603-2611.
463. Jick, S.S., Choi, H., Li, L., McInnes, I.B. and Sattar, N. Hyperlipidaemia, statin use and the risk of developing rheumatoid arthritis. *Annals of the Rheumatic Diseases.* 2009, **68**(4), pp.546-551.
464. Simpson, C.R., Anderson, W.J., Helms, P.J., Taylor, M.W., Watson, L. et al. Coincidence of immune-mediated diseases driven by Th1 and Th2 subsets suggests a common aetiology. A population-based study using computerized general practice data. *Clin Exp Allergy.* 2002, **32**(1), pp.37-42.
465. Seminog, O.O., Seminog, A.B., Yeates, D. and Goldacre, M.J. Associations between Klinefelter's syndrome and autoimmune diseases: English national record linkage studies. *Autoimmunity.* 2015, **48**(2), pp.125-128.
466. Tremlett, H.L., Evans, J., Wiles, C.M. and Luscombe, D.K. Asthma and multiple sclerosis: an inverse association in a case-control general practice population. *QJM.* 2002, **95**(11), pp.753-756.
467. Thorarensen, S.M., Lu, N., Ogdie, A., Gelfand, J.M., Choi, H.K. et al. Physical trauma recorded in primary care is associated with the onset of psoriatic arthritis among patients with psoriasis. *Annals of the Rheumatic Diseases.* 2017, **76**(3), pp.521-525.
468. Norton, S., Koduri, G., Nikiphorou, E., Dixey, J., Williams, P. et al. A study of baseline prevalence and cumulative incidence of comorbidity and extra-articular manifestations in RA and their impact on outcome. *Rheumatology (Oxford).* 2013, **52**(1), pp.99-110.
469. Jafri, K., Bartels, C.M., Shin, D., Gelfand, J.M. and Ogdie, A. Incidence and Management of Cardiovascular Risk Factors in Psoriatic Arthritis and Rheumatoid Arthritis: A Population-Based Study. *Arthritis Care & Research.* 2017, **69**(1), pp.51-57.
470. Ogdie, A., Harter, L., Shin, D., Baker, J., Takeshita, J. et al. The risk of fracture among patients with psoriatic arthritis and psoriasis: a population-based study. *Annals of the Rheumatic Diseases.* 2017, **76**(5), pp.882-885.
471. Ogdie, A., Maliha, S., Shin, D., Love, T.J., Baker, J. et al. Cause-specific mortality in patients with psoriatic arthritis and rheumatoid arthritis. *Rheumatology (Oxford).* 2017, **56**(6), pp.907-911.
472. Thomas, E., Symmons, D.P., Brewster, D.H., Black, R.J. and Macfarlane, G.J. National study of cause-specific mortality in rheumatoid arthritis, juvenile chronic arthritis, and other rheumatic conditions: a 20 year followup study. *J Rheumatol.* 2003, **30**(5), pp.958-965.
473. Pujades-Rodriguez, M., Duyx, B., Thomas, S.L., Stogiannis, D., Rahman, A. et al. Rheumatoid Arthritis and Incidence of Twelve Initial Presentations of Cardiovascular Disease: A Population Record-Linkage Cohort Study in England. *PLoS One.* 2016, **11**(3), p.e0151245.
474. Geoghegan, J.M., Clark, D.I., Bainbridge, L.C., Smith, C. and Hubbard, R. Risk factors in carpal tunnel syndrome. *J Hand Surg Br.* 2004, **29**(4), pp.315-320.
475. Dubreuil, M., Rho, Y.H., Man, A., Zhu, Y., Zhang, Y. et al. Diabetes incidence in psoriatic arthritis, psoriasis and rheumatoid arthritis: a UK population-based cohort study. *Rheumatology (Oxford).* 2014, **53**(2), pp.346-352.
476. Cullen, D.J., Seager, J.M., Holmes, S., Doherty, M., Wilson, J.V. et al. Pharmacoepidemiology of non-steroidal anti-inflammatory drug use in

Nottingham general practices. *Aliment Pharmacol Ther.* 2000, **14**(2), pp.177-185.
477. Edwards, C.J., Campbell, J., van Staa, T. and Arden, N.K. Regional and temporal variation in the treatment of rheumatoid arthritis across the UK: a descriptive register-based cohort study. *BMJ Open.* 2012, **2**(6).
478. Edwards, C.J., Arden, N.K., Fisher, D., Saperia, J.C., Reading, I. et al. The changing use of disease-modifying anti-rheumatic drugs in individuals with rheumatoid arthritis from the United Kingdom General Practice Research Database. *Rheumatology (Oxford).* 2005, **44**(11), pp.1394-1398.
479. Judge, A., Wallace, G., Prieto-Alhambra, D., Arden, N.K. and Edwards, C.J. Can the publication of guidelines change the management of early rheumatoid arthritis? An interrupted time series analysis from the United Kingdom. *Rheumatology (Oxford).* 2015, **54**(12), pp.2244-2248.
480. Luqmani, R., Hennell, S., Estrach, C., Birrell, F., Bosworth, A. et al. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of rheumatoid arthritis (the first two years). *Rheumatology (Oxford).* 2006, **45**(9), pp.1167-1169.
481. Rachapalli, S.M., Williams, R., Walsh, D.A., Young, A., Kiely, P.D. et al. First-line DMARD choice in early rheumatoid arthritis--do prognostic factors play a role? *Rheumatology (Oxford).* 2010, **49**(7), pp.1267-1271.
482. Simon, T.A., Smitten, A.L., Franklin, J., Askling, J., Lacaille, D. et al. Malignancies in the rheumatoid arthritis abatacept clinical development programme: an epidemiological assessment. *Annals of the Rheumatic Diseases.* 2009, **68**(12), pp.1819-1826.
483. Hawley, S., Cordtz, R., Dreyer, L., Edwards, C.J., Arden, N.K. et al. Association between NICE guidance on biologic therapies with rates of hip and knee replacement among rheumatoid arthritis patients in England and Wales: An interrupted time-series analysis. *Seminars in Arthritis and Rheumatism.* 2017.
484. Edwards, C.J., Cooper, C., Fisher, D., Field, M., van Staa, T.P. et al. The importance of the disease process and disease-modifying antirheumatic drug treatment in the development of septic arthritis in patients with rheumatoid arthritis. *Arthritis & Rheumatism.* 2007, **57**(7), pp.1151-1157.
485. Crossfield, S.S., Lai, L.Y.H., Kingsbury, S.R., Baxter, P., Johnson, O. et al. Variation in methods, results and reporting in electronic health record-based studies evaluating routine care in gout: A systematic review. *PLoS One.* 2019, **14**(10).
486. Black, R.J., Joseph, R.M., Brown, B., Movahedi, M., Lunt, M. et al. Half of U.K. patients with rheumatoid arthritis are prescribed oral glucocorticoid therapy in primary care: a retrospective drug utilisation study. *Arthritis Research & Therapy.* 2015, **17**, p.375.
487. Souverein, P.C., Berard, A., Van Staa, T.P., Cooper, C., Egberts, A.C. et al. Use of oral glucocorticoids and risk of cardiovascular and cerebrovascular disease in a population based case-control study. *Heart.* 2004, **90**(8), pp.859-865.
488. Fischer, L.M., Schlienger, R.G., Matter, C.M., Jick, H. and Meier, C.R. Discontinuation of nonsteroidal anti-inflammatory drug therapy and risk of acute myocardial infarction. *Arch Intern Med.* 2004, **164**(22), pp.2472-2476.
489. Movahedi, M., Beauchamp, M.E., Abrahamowicz, M., Ray, D.W., Michaud, K. et al. Risk of Incident Diabetes Mellitus Associated With the Dosage and Duration of Oral Glucocorticoid Therapy in Patients With Rheumatoid Arthritis. *Arthritis Rheumatol.* 2016, **68**(5), pp.1089-1098.
490. Watson, D.J., Rhodes, T., Cai, B. and Guess, H.A. Lower risk of thromboembolic cardiovascular events with naproxen among patients with rheumatoid arthritis. *Arch Intern Med.* 2002, **162**(10), pp.1105-1110.
491. van Staa, T.P., Geusens, P., Zhang, B., Leufkens, H.G., Boonen, A. et al. Individual fracture risk and the cost-effectiveness of bisphosphonates in

patients using oral glucocorticoids. *Rheumatology (Oxford).* 2007, **46**(3), pp.460-466.
492. Sheng, X., Murphy, M.J., Macdonald, T.M. and Wei, L. Effectiveness of statins on total cholesterol and cardiovascular disease and all-cause mortality in osteoarthritis and rheumatoid arthritis. *J Rheumatol.* 2012, **39**(1), pp.32-40.
493. Forbes, H.J., Thomas, S.L., Smeeth, L. and Langan, S.M. Prescription of antiviral therapy after herpes zoster in general practice: who receives therapy? *Br J Gen Pract.* 2012, **62**(605), pp.e808-814.
494. Costello, R., Winthrop, K.L., Pye, S.R., Brown, B. and Dixon, W.G. Influenza and Pneumococcal Vaccination Uptake in Patients with Rheumatoid Arthritis Treated with Immunosuppressive Therapy in the UK: A Retrospective Cohort Study Using Data from the Clinical Practice Research Datalink. *PLoS One.* 2016, **11**(4), p.e0153848.
495. Pratt, A.G., Lendrem, D., Hargreaves, B., Aslam, O., Galloway, J.B. et al. Components of treatment delay in rheumatoid arthritis differ according to autoantibody status: validation of a single-centre observation using national audit data. *Rheumatology (Oxford).* 2016, **55**(10), pp.1843-1848.
496. Alemao, E., Cawston, H., Bourhis, F., Al, M., Rutten-van Molken, M. et al. Comparison of cardiovascular risk algorithms in patients with vs without rheumatoid arthritis and the role of C-reactive protein in predicting cardiovascular outcomes in rheumatoid arthritis. *Rheumatology (Oxford).* 2017, **56**(5), pp.777-786.
497. Joseph, R.M., Movahedi, M., Dixon, W.G. and Symmons, D.P. Smoking-Related Mortality in Patients With Early Rheumatoid Arthritis: A Retrospective Cohort Study Using the Clinical Practice Research Datalink. *Arthritis Care & Research.* 2016, **68**(11), pp.1598-1606.
498. Joseph, R.M., Movahedi, M., Dixon, W.G. and Symmons, D.P. Risks of smoking and benefits of smoking cessation on hospitalisations for cardiovascular events and respiratory infection in patients with rheumatoid arthritis: a retrospective cohort study using the Clinical Practice Research Datalink. *RMD Open.* 2017, **3**(2), p.e000506.
499. Rajakulendran, S., Gadsby, K. and Deighton, C. Rheumatoid arthritis, alcohol, leflunomide and methotrexate. Can changes to the BSR guidelines for leflunomide and methotrexate on alcohol consumption be justified? *Musculoskeletal Care.* 2008, **6**(4), pp.233-245.
500. Humphreys, J.H., Warner, A., Costello, R., Lunt, M., Verstappen, S.M.M. et al. Quantifying the hepatotoxic risk of alcohol consumption in patients with rheumatoid arthritis taking methotrexate. *Annals of the Rheumatic Diseases.* 2017, **76**(9), pp.1509-1514.
501. Pujades-Rodriguez, M., Morgan, A.W., Cubbon, R.M. and Wu, J. Dose-dependent oral glucocorticoid cardiovascular risk in people with immune-mediated inflammatory diseases. *medRxiv.* 2020.
502. Watson, D.J., Rhodes, T., Cai, B. and Guess, H.A. Lower risk of thromboembolic cardiovascular events with naproxen among patients with rheumatoid arthritis. *Archives of internal medicine.* 2002, **162**(10), pp.1105-1110.
503. Benjamin, O., Bansal, P., Goyal, A. and Lappin, S.L. Disease modifying anti-rheumatic drugs (DMARD). In: *StatPearls [Internet]*. StatPearls Publishing, 2019.
504. Partington, R.J., Muller, S., Helliwell, T., Mallen, C.D. and Sultan, A.A. Incidence, prevalence and treatment burden of polymyalgia rheumatica in the UK over two decades: a population-based study. *Annals of the Rheumatic Diseases.* 2018, **77**(12), pp.1750-1756.
505. Arthritis, V. *Versus Arthritis: About Arthritis.* [Online]. 2019. [Accessed 02 July 2019]. Available from: https://www.versusarthritis.org/
506. Department of Health. *NHS.* [Online]. 2019. [Accessed 02 July 2019]. Available from: www.nhs.uk

507. Wallace, B., Vummidi, D. and Khanna, D. Management of connective tissue diseases associated interstitial lung disease: a review of the published literature. *Curr Opin Rheumatol.* 2016, **28**(3), pp.236-245.
508. Dean, L.E., Jones, G.T., MacDonald, A.G., Downham, C., Sturrock, R.D. et al. Global prevalence of ankylosing spondylitis. *Rheumatology (Oxford).* 2014, **53**(4), pp.650-657.
509. Rudwaleit, M., Van Der Heijde, D., Landewé, R., Listing, J., Akkoc, N. et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Annals of the Rheumatic Diseases.* 2009, **68**(6), pp.777-783.
510. Tate, A.R., Dungey, S., Glew, S., Beloff, N., Williams, R. et al. Quality of recording of diabetes in the UK: how does the GP's method of coding clinical data affect incidence estimates? Cross-sectional study using the CPRD database. *BMJ Open.* 2017, **7**(1), p.e012905.
511. Keiding, N. Event history analysis and the cross-section. *Statistics in Medicine.* 2006, **25**(14), pp.2343-2364.
512. Spronk, I., Korevaar, J.C., Poos, R., Davids, R., Hilderink, H. et al. Calculating incidence rates and prevalence proportions: not as simple as it seems. *BMC Public Health.* 2019, **19**(1), p.512.
513. Sykes, M.P., Doll, H., Sengupta, R. and Gaffney, K. Delay to diagnosis in axial spondyloarthritis: are we improving in the UK? *Rheumatology (Oxford).* 2015, **54**(12), pp.2283-2284.
514. Maksymowych, W.P. MRI in ankylosing spondylitis. *Curr Opin Rheumatol.* 2009, **21**(4), pp.313-317.
515. Dincer, U., Cakar, E., Kiralp, M.Z. and Dursun, H. Diagnosis delay in patients with ankylosing spondylitis: possible reasons and proposals for new diagnostic criteria. *Clinical Rheumatology.* 2008, **27**(4), pp.457-462.
516. Jovaní, V., Blasco-Blasco, M., Ruiz-Cantero, M.T. and Pascual, E. Understanding how the diagnostic delay of spondyloarthritis differs between women and men: a systematic review and metaanalysis. *The Journal of Rheumatology.* 2017, **44**(2), pp.174-183.
517. Jois, R., Macgregor, A. and Gaffney, K. Recognition of inflammatory back pain and ankylosing spondylitis in primary care. *Rheumatology (Oxford).* 2008, **47**(9), pp.1364-1366.
518. Jones, A., Harrison, N., Jones, T., Rees, J.D. and Bennett, A.N. Time to diagnosis of axial spondylarthritis in clinical practice: signs of improving awareness? *Rheumatology (Oxford).* 2014, **53**(11), pp.2126-2127.
519. Hamilton, L., Gilbert, A., Skerrett, J., Dickinson, S. and Gaffney, K. Services for people with ankylosing spondylitis in the UK—a survey of rheumatologists and patients. *Rheumatology (Oxford).* 2011, **50**(11), pp.1991-1998.
520. Magrey, M., Yi, E., Wolin, D., Price, M., Chirila, C. et al. Understanding Barriers in the Pathway to Diagnosis of Ankylosing Spondylitis: Results From a US Survey of 1690 Physicians From 10 Specialties. *ACR Open Rheumatology.* 2020, **2**(10), pp.616-626.
521. Bashir, M.T., Iversen, L. and Burton, C. Clinical features in primary care electronic records before diagnosis of ankylosing spondylitis: a nested case-control study. *BMC family practice.* 2020, **21**(1), pp.1-9.
522. Irvine, S., Munro, R. and Porter, D. Early referral, diagnosis, and treatment of rheumatoid arthritis: evidence for changing medical practice. *Annals of the Rheumatic Diseases.* 1999, **58**(8), pp.510-513.
523. Buttgereit, F. and Bijlsma, J.W. Glucocorticoids in rheumatoid arthritis: the picture is shaping up. *Annals of the Rheumatic Diseases.* 2017.
524. Black, R.J., Joseph, R.M., Brown, B., Movahedi, M., Lunt, M. et al. Half of UK patients with rheumatoid arthritis are prescribed oral glucocorticoid therapy in primary care: a retrospective drug utilisation study. *Arthritis Research & Therapy.* 2015, **17**(1), p.375.

525. Quint, J.K., Müllerova, H., DiSantostefano, R.L., Forbes, H., Eaton, S. et al. Validation of chronic obstructive pulmonary disease recording in the Clinical Practice Research Datalink (CPRD-GOLD). *BMJ Open.* 2014, **4**(7), p.e005540.
526. EBM DataLab. *Open Prescribing.* [Online]. 2019. [Accessed 06 November 2019]. Available from: https://openprescribing.net/
527. Fardet, L., Petersen, I. and Nazareth, I. Prevalence of long-term oral glucocorticoid prescriptions in the UK over the past 20 years. *Rheumatology (Oxford).* 2011, **50**(11), pp.1982-1990.
528. Sharma, M., Nazareth, I. and Petersen, I. Trends in incidence, prevalence and prescribing in type 2 diabetes mellitus between 2000 and 2013 in primary care: a retrospective cohort study. *BMJ Open.* 2016, **6**(1).
529. Dunteman, G.H. and Ho, M.-H.R. *An introduction to generalized linear models.* Sage Publications, 2005.
530. Hayat, M.J. and Higgins, M. Understanding poisson regression. *Journal of Nursing Education.* 2014, **53**(4), pp.207-215.
531. Jansakul, N. and Hinde, J. Score tests for zero-inflated Poisson models. *Computational Statistics and Data Analysis.* 2002, **40**(1), pp.75-96.
532. Warton, D.I. Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. *Environmetrics.* 2005, **16**(3), pp.275-289.
533. Joe, H. Accuracy of Laplace approximation for discrete response mixed models. *Computational Statistics and Data Analysis.* 2008, **52**(12), pp.5066-5074.
534. Hausman, J.A., Hall, B.H. and Griliches, Z. Econometric models for count data with an application to the patents-R&D relationship. *Econometrica.* 1984, **52**(4).
535. Wu, J., Mackie, S.L. and Pujades-Rodriguez, M. Glucocorticoid Dose-Dependent Risk of Type 2 Diabetes in Six Immune-Mediated Inflammatory Diseases: A Population-Based Cohort Analysis. *Preprints with the Lancet.* 2019.
536. Mebrahtu, T.F., Morgan, A.W., West, R.M., Stewart, P.M. and Pujades-Rodriguez, M. Oral glucocorticoids and incidence of hypertension in people with chronic inflammatory diseases: a population-based cohort study. *Canadian Medical Association Journal.* 2020, **192**(12), pp.E295-E301.
537. Stouten, V., Westhovens, R., Pazmino, S., De Cock, D., Van der Elst, K. et al. Effectiveness of different combinations of DMARDs and glucocorticoid bridging in early rheumatoid arthritis: two-year results of CareRA. *Rheumatology (Oxford).* 2019, **58**(12), pp.2284-2294.
538. Hui, S.L., Slemenda, C.W. and Johnston, C.C. Age and bone mass as predictors of fracture in a prospective study. *The Journal of Clinical Investigation.* 1988, **81**(6), pp.1804-1809.
539. Studenic, P., Radner, H., Smolen, J.S. and Aletaha, D. Discrepancies between patients and physicians in their perceptions of rheumatoid arthritis disease activity. *Arthritis & Rheumatism.* 2012, **64**(9), pp.2814-2823.
540. Mathieu, S., Couderc, M., Pereira, B., Dubost, J.-J., Malochet-Guinamand, S. et al. Prevalence of Migraine and Neuropathic Pain in Rheumatic Diseases. *Journal of Clinical Medicine.* 2020, **9**(6), p.1890.
541. Studenic, P., Smolen, J.S. and Aletaha, D. Near misses of ACR/EULAR criteria for remission: effects of patient global assessment in Boolean and index-based definitions. *Annals of the Rheumatic Diseases.* 2012, **71**(10), pp.1702-1705.
542. Bresalier, R., Sandler, R., Quan, H., Bolognese, J., Oxenius, B. et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med.* 2005, **352**, pp.1092-1102.
543. Hayes, K.N., Ban, J., Athanasiadis, G., Burden, A.M. and Cadarette, S.M. Time trends in oral bisphosphonate initiation in Ontario, Canada over 20 years reflect drug policy and healthcare delivery changes. *Osteoporosis International.* 2019, **30**(11), pp.2311-2319.

544. Klop, C., Gibson-Smith, D., Elders, P., Welsing, P., Leufkens, H. et al. Anti-osteoporosis drug prescribing after hip fracture in the UK: 2000–2010. *Osteoporosis International.* 2015, **26**(7), pp.1919-1928.
545. Wu, J., Keeley, A., Mallen, C., Morgan, A. and Pujades Rodriguez, M. Incidence of infections associated with oral glucocorticoid dose in people diagnosed with polymyalgia rheumatica and giant cell arteritis: a cohort study in England. *Canadian Medical Association Journal.* 2019.
546. Merck and Co Inc. *Merck announces voluntary worldwide withdrawal of Vioxx [news release; 30 September 2004].* [Online]. 2004. [Accessed 10 February 2020]. Available from: http://www.merck.com/newsroom/vioxx/pdf/vioxx_press_release_final.pdf
547. Suh, D.-C., Hunsche, E., Shin, H.-C. and Mavros, P. Co-prescribing of proton pump inhibitors among chronic users of NSAIDs in the UK. *Rheumatology (Oxford).* 2008, **47**(4), pp.458-463.
548. Dagli, R.J. and Sharma, A. Polypharmacy: a global risk factor for elderly people. *Journal of International Oral Health.* 2014, **6**(6), p.i.
549. Vinogradova, Y., Coupland, C. and Hippisley-Cox, J. Use of combined oral contraceptives and risk of venous thromboembolism: nested case-control studies using the QResearch and CPRD databases. *BMJ.* 2015, **350**, p.h2135.
550. Khan, T., Alvand, A., Prieto-Alhambra, D., Culliford, D.J., Judge, A. et al. ACL and meniscal injuries increase the risk of primary total knee replacement for osteoarthritis: a matched case–control study using the Clinical Practice Research Datalink (CPRD). *British Journal of Sports Medicine.* 2019, **53**(15), pp.965-968.
551. Zhang, J., Haines, C., Watson, A.J., Hart, A.R., Platt, M.J. et al. Oral antibiotic use and risk of colorectal cancer in the United Kingdom, 1989–2012: a matched case–control study. *Gut.* 2019, **68**(11), pp.1971-1978.
552. Rockenschaub, P., Nguyen, V., Aldridge, R.W., Acosta, D., García-Gómez, J.M. et al. Data-driven discovery of changes in clinical code usage over time: a case-study on changes in cardiovascular disease recording in two English electronic health records databases (2001–2015). *BMJ Open.* 2020, **10**(2).
553. Mebrahtu, T.F., Morgan, A.W., Keeley, A., Baxter, P.D., Stewart, P.M. et al. Dose dependency of iatrogenic glucocorticoid excess and adrenal insufficiency and mortality: a cohort study in England. *The Journal of Clinical Endocrinology and Metabolism.* 2019, **104**(9), pp.3757-3767.
554. Scherer, H.U., Häupl, T. and Burmester, G.R. The etiology of rheumatoid arthritis. *Journal of Autoimmunity.* 2020, p.102400.
555. Vinogradova, Y., Coupland, C., Brindle, P. and Hippisley-Cox, J. Discontinuation and restarting in patients on statin treatment: prospective open cohort study using a primary care database. *BMJ.* 2016, **353**, p.i3305.
556. Diab, D.L. and Watts, N.B. *Bisphosphonate drug holiday: who, when and how long*. SAGE Publications Sage UK: London, England. 2013.
557. Fauci, A.S. Mechanisms of the Immunosuppressive and Anti-Inflammatory Effects of Glucorcorticosteroids. *Journal of Immunopharmacology.* 1978, **1**(1), pp.1-25.
558. van Paassen, J., Vos, J.S., Hoekstra, E.M., Neumann, K.M., Boot, P.C. et al. Corticosteroid use in COVID-19 patients: a systematic review and meta-analysis on clinical outcomes. *Critical Care.* 2020, **24**(1), pp.1-22.
559. Jick, S.S., Lieberman, E.S., Rahman, M.U. and Choi, H.K. Glucocorticoid use, other associated factors, and the risk of tuberculosis. *Arthritis Care & Research.* 2006, **55**(1), pp.19-26.
560. Gianfrancesco, M., Hyrich, K.L., Al-Adely, S., Carmona, L., Danila, M.I. et al. Characteristics associated with hospitalisation for COVID-19 in people with rheumatic disease: data from the COVID-19 Global Rheumatology Alliance physician-reported registry. *Annals of the Rheumatic Diseases.* 2020, **79**(7), pp.859-866.

561. Aggarwal, R. and Malaviya, A.N. Diagnosis delay in patients with ankylosing spondylitis: factors and outcomes—an Indian perspective. *Clinical Rheumatology.* 2009, **28**(3), pp.327-331.
562. Kalweit, M., Walker, U.A., Finckh, A., Müller, R., Kalweit, G. et al. Personalized prediction of disease activity in patients with rheumatoid arthritis using an adaptive deep neural network. 2021, **16**(6), p.e0252289.
563. McCloskey, E.V., Johansson, H., Oden, A. and Kanis, J.A.J.C.o.r. From relative risk to absolute fracture risk calculation: the FRAX algorithm. 2009, **7**(3), pp.77-83.
564. Speyer, C.B., Li, D., Guan, H., Yoshida, K., Stevens, E. et al. Comparison of an administrative algorithm for SLE disease severity to clinical SLE disease activity index scores. 2020, **40**(2), pp.257-261.
565. Clegg, A., Bates, C., Young, J., Ryan, R., Nichols, L. et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. 2016, **45**(3), pp.353-360.
566. Gulliford, M., Sun, X., Anjuman, T., Yelland, E. and Murray-Thomas, T. Antibiotic prescribing records in two UK primary care electronic health record systems. Comparison of the CPRD GOLD and CPRD Aurum databases. 2020.
567. Presser, L., Hruskova, M., Rowbottom, H. and Kancir, J. Care. data and access to UK health records: patient privacy and public trust. *Technology Science.* 2015, pp.1-35.
568. Smolen, J., Landewé, R., Breedveld, F., Buch, T., Burmester, G. et al. EULAR recommendations for the managements of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 Update. *Rheumatologia.* 2014, **28**(1), pp.1-25.
569. Oza, A., Lu, N., Schoenfeld, S.R., Fisher, M.C., Dubreuil, M. et al. Survival benefit of statin use in ankylosing spondylitis: a general population-based cohort study. *Annals of the Rheumatic Diseases.* 2017, **76**(10), pp.1737-1742.
570. Atzeni, F., Nucera, V., Galloway, J., Zoltán, S. and Nurmohamed, M.J.E.O.o.B.T. Cardiovascular risk in ankylosing spondylitis and the effect of anti-TNF drugs: a narrative review. *Expert Opinion on Biological Therapy.* 2020, **20**(5), pp.517-524.
571. Van Der Aalst, W. Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems.* 2012, **3**(2), pp.1-17.
572. Denaxas, S., Gonzalez-Izquierdo, A., Direk, K., Fitzpatrick, N., Banerjee, A. et al. UK phenomics platform for developing and validating EHR phenotypes: CALIBER. *bioRxiv.* 2019, p.539403.
573. Springate, D.A., Kontopantelis, E., Ashcroft, D.M., Olier, I., Parisi, R. et al. ClinicalCodes: an online clinical codes repository to improve the validity and reproducibility of research using electronic medical records. *PLoS One.* 2014, **9**(6), p.e99825.
574. Schuurmans, A.A., de Looff, P., Nijhof, K.S., Rosada, C., Scholte, R.H. et al. Validity of the Empatica E4 Wristband to Measure Heart Rate Variability (HRV) Parameters: a Comparison to Electrocardiography (ECG). *J Med Syst.* 2020, **44**(11), pp.1-11.
575. Hafezi, H., Robertson, T.L., Moon, G.D., Au-Yeung, K.-Y., Zdeblick, M.J. et al. An ingestible sensor for measuring medication adherence. *IEEE Transactions on Biomedical Engineering.* 2014, **62**(1), pp.99-109.
576. RIVIAM Digital Care. *RIVIAM Integrates with TPP SystmOne.* [Online]. 2019. [Accessed 11 January 2021]. Available from: https://www.riviam.com/what-we-do/gp-system-integration/integrating-with-tpp-systm1
577. Genes, N., Violante, S., Cetrangol, C., Rogers, L., Schadt, E.E. et al. From smartphone to EHR: a case report on integrating patient-generated health data. *NPJ Digital Medicine.* 2018, **1**(1), pp.1-6.
578. Ashton, K. That 'internet of things' thing. *RFID journal.* 2009, **22**(7), pp.97-114.

579. Noorbakhsh-Sabet, N., Zand, R., Zhang, Y. and Abedi, V. Artificial intelligence transforms the future of health care. *The American Journal of Medicine.* 2019, **132**(7), pp.795-801.
580. Chen, Y., Argentinis, J.E. and Weber, G. IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research. *Clinical Therapeutics.* 2016, **38**(4), pp.688-701.
581. Gupta, S., Kar, A.K., Baabdullah, A. and Al-Khowaiter, W.A. Big data with cognitive computing: A review for the future. *International Journal of Information Management.* 2018, **42**, pp.78-89.
582. Huang, S.-C., Pareek, A., Seyyedi, S., Banerjee, I. and Lungren, M.P. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. *NPJ Digital Medicine.* 2020, **3**(1), pp.1-9.
583. Arbabshirani, M.R., Fornwalt, B.K., Mongelluzzo, G.J., Suever, J.D., Geise, B.D. et al. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. *NPJ Digital Medicine.* 2018, **1**(1), pp.1-7.
584. Sopic, D., Aminifar, A., Aminifar, A. and Atienza, D. Real-time event-driven classification technique for early detection and prevention of myocardial infarction on wearable systems. *IEEE Transactions on Biomedical Circuits and Systems.* 2018, **12**(5), pp.982-992.
585. He, L. and Cao, C. Automated depression analysis using convolutional neural networks from speech. *Journal of Biomedical Informatics.* 2018, **83**, pp.103-111.
586. Barrett, M.A., Humblet, O., Marcus, J.E., Henderson, K., Smith, T. et al. Effect of a mobile health, sensor-driven asthma management platform on asthma control. *Annals of Allergy, Asthma & Immunology.* 2017, **119**(5), pp.415-421. e411.
587. Luik, A.I., Kyle, S.D. and Espie, C.A. Digital cognitive behavioral therapy (dCBT) for insomnia: a state-of-the-science review. *Current Sleep Medicine Reports.* 2017, **3**(2), pp.48-56.
588. Sim, I. Mobile devices and health. *New England Journal of Medicine.* 2019, **381**(10), pp.956-968.
589. Jensen, P.B., Jensen, L.J. and Brunak, S. Mining electronic health records: towards better research applications and clinical care. *Nature Reviews Genetics.* 2012, **13**(6), pp.395-405.
590. Google. *DeepMind: What if solving one problem coul unlock solutions to thousands more?* [Online]. 2021. [Accessed 25 February 2021]. Available from: https://deepmind.com/
591. Observational Health Data Sciences and Informatics. *OMOP Common Data Model.* [Online]. 2021. [Accessed 12 February 2021]. Available from: https://www.ohdsi.org/data-standardization/the-common-data-model/
592. Estiri, H., Strasser, Z.H., Klann, J.G., Naseri, P., Wagholikar, K.B. et al. Predicting COVID-19 mortality with electronic medical records. *NPJ Digital Medicine.* 2021, **4**(1), pp.1-10.
593. Clough, H.E., McIntyre, K.M., Patterson, G.E., Harris, J.P. and Rushton, J. Use of routine death and illness surveillance data to provide insight for UK pandemic planning: lessons from COVID-19. *BMJ Open.* 2021, **11**(2), p.e044707.
594. UK Biobank. *UK Biobank: Hospital inpatient data (version 3.0).* [Online]. 2020. [Accessed 20 January 2020]. Available from: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/HospitalEpisodeStatistics.pdf
595. Dagliati, A., Malovini, A., Tibollo, V. and Bellazzi, R. Health informatics and EHR to support clinical research in the COVID-19 pandemic: an overview. *Briefings in Bioinformatics.* 2021.
596. Hyrich, K.L. and Zink, A. What can rheumatology expect from real-world data? *Rheumatology (Oxford).* 2020, **59**(1), pp.12-13.

597. Li, R., Chen, Y. and Moore, J.H. Integration of genetic and clinical information to improve imputation of data missing from electronic health records. *Journal of the American Medical Informatics Association.* 2019, **26**(10), pp.1056-1063.
598. Rozinat, A. *Data Quality Problems In Process Mining And What To Do About Them — Part 1: Formatting Errors.* [Online]. 2016. [Accessed 02 March 2018].
599. Weber, G.M., Mandl, K.D. and Kohane, I.S. Finding the missing link for big biomedical data. *JAMA.* 2014, **311**(24), pp.2479-2480.
600. Pisaniello, H.L. and Dixon, W.G. What does digitalization hold for the creation of real-world evidence? *Rheumatology (Oxford).* 2020, **59**(1), pp.39-45.
601. De Vera, M., Rahman, M.M., Rankin, J., Kopec, J., Gao, X. et al. Gout and the risk of Parkinson's disease: a cohort study. *Arthritis & Rheumatism.* 2008, **59**(11), pp.1549-1554.
602. Soriano, L.C., Rothenbacher, D., Choi, H.K. and Rodriguez, L.A.G. Contemporary epidemiology of gout in the UK general population. *Arthritis Research & Therapy.* 2011, **13**(2), p.3.
603. Cheyoe, N., Kuning, M. and Lim, A. The prevalence of chronic kidney disease among gout patients in Nongjik hospital, Pattani province. *Thai Journal of Pharmaceutical Sciences.* 2012, **36**(4), pp.144-149.
604. Harrold, L.R., Mazor, K.M., Peterson, D., Naz, N., Firneno, C. et al. Patients' knowledge and beliefs concerning gout and its treatment: a population based study. *BMC Musculoskelet Disord.* 2012, **13**(180).
605. Zarowitz, B.J. and O'Shea, T.E. Demographic and clinical profile of nursing facility residents with gout. *Consult Pharm.* 2013, **28**(6), pp.370-382.
606. George, M., Pullman-Mooar, S., Hussain, F. and Schumacher, H.R. Evaluating Appropriate Use of Prophylactic Colchicine for Gout Flare Prevention. *Arthritis Care & Research.* 2014, **66**(8), pp.1258-1262.
607. Jackson, G., Dalbeth, N., Te Karu, L., Winnard, D., Gow, P. et al. Variation in gout care in Aotearoa New Zealand: a national analysis of quality markers. *N Z Med J.* 2014, **127**(1404), pp.37-47.
608. MacFarlane, L.A., Liu, C.C. and Solomon, D.H. The effect of initiating pharmacologic insulin on serum uric acid levels in patients with diabetes: A matched cohort analysis. *Seminars in Arthritis and Rheumatism.* 2014, **44**(5), pp.592-596.
609. Meek, I.L., Vonkeman, H.E. and Van De Laar, M.A.F.J. Hyperuricaemia: A marker of increased cardiovascular risk in rheumatic patients: Analysis of the ACT-CVD cohort. *BMC Musculoskelet Disord.* 2014, **15**(1).
610. Park, J.W., Ko, D.J., Yoo, J.J., Chang, S.H., Cho, H.J. et al. Clinical factors and treatment outcomes associated with failure in the detection of urate crystal in patients with acute gouty arthritis. *Korean J Intern Med.* 2014, **29**(3), pp.361-369.
611. Hmar, R.C., Kannangara, D.R.W., Ramasamy, S.N., Baysari, M.T., Williams, K.M. et al. Understanding and improving the use of allopurinol in a teaching hospital. *Intern Med J.* 2015, **45**(4), pp.383-390.
612. Kerr, G.S., Richards, J.S., Nunziato, C.A., Patterson, O.V., DuVall, S.L. et al. Measuring physician adherence with gout quality indicators: a role for natural language processing. *Arthritis Care & Research.* 2015, **67**(2), pp.273-279.
613. Lu, N., Dubreuil, M., Zhang, Y., Neogi, T., Rai, S.K. et al. Gout and the risk of Alzheimer's disease: a population-based, BMI-matched cohort study. *Annals of the Rheumatic Diseases.* 2015, **75**(3), pp.547-551.
614. Robinson, P.C., Taylor, W.J. and Dalbeth, N. An Observational Study of Gout Prevalence and Quality of Care in a National Australian General Practice Population. *J Rheumatol.* 2015, **42**(9), pp.1702-1707.
615. Dehlin, M., Drivelegka, P., Sigurdardottir, V., Svärd, A. and Jacobsson, L.T.H. Incidence and prevalence of gout in Western Sweden. *Arthritis Research & Therapy.* 2016, **18**(164).
616. Nyberg, F., Horne, L., Morlock, R., Nuevo, J., Storgard, C. et al. Comorbidity Burden in Trial-Aligned Patients with Established Gout in Germany, UK, US,

and France: A Retrospective Analysis. *Advances in Therapy.* 2016, **33**(7), pp.1180-1198.
617. Chang, K., Yokose, C., Tenner, C., Oh, C., Donnino, R. et al. Association Between Gout and Aortic Stenosis. *The American Journal of Medicine.* 2017, **130**(2).
618. Lee, J.H., Yang, J., Shin, K., Lee, G.H., Lee, W.W. et al. Elderly Patients Exhibit Stronger Inflammatory Responses during Gout Attacks. *J Korean Med Sci.* 2017, **32**(12), pp.1967-1973.
619. Olaru, L., Soong, L., Dhillon, S. and Yacyshyn, E. Coexistent rheumatoid arthritis and gout: a case series and review of the literature. *Clinical Rheumatology.* 2017, **36**(12), pp.2835-2838.
620. Bevis, M., Blagojevic-Bucknall, M., Mallen, C., Hider, S. and Roddy, E. Comorbidity clusters in people with gout: an observational cohort study with linked medical record review. *Rheumatology (Oxford).* 2018, **57**(8), pp.1358-1363.
621. Dehlin, M. and Jacobsson, L.T.H. Trends in Gout Hospitalization in Sweden. *J Rheumatol.* 2018, **45**(1), pp.145-146.
622. Hassan, E. and Choudry, B. The compliance of guidelines set by the British Society for Rheumatology for managing Gout. *British Journal of Healthcare Management.* 2018, **24**(1), pp.26-30.
623. Jung, J.Y., Choi, Y., Suh, C.H., Yoon, D. and Kim, H.A. Effect of fenofibrate on uric acid level in patients with gout. *Scientific Reports.* 2018, **8**(1), pp.1-9.
624. Keller, S.F., Rai, S.K., Lu, N., Oza, A., Jorge, A.M. et al. Statin use and mortality in gout: A general population-based cohort study. *Semin Arthritis Rheum.* 2018, **48**(3), pp.449-455.
625. Lee, J.S., Won, J., Kwon, O.C., Lee, S.S. and Oh, J.S. Hepatic Safety of Febuxostat Compared with Allopurinol in Gout Patients with Fatty Liver Disease. *J Rheumatol.* 2018, **46**(5), pp.527-531.
626. Lin, L.W., Teng, G.G., Lim, A.Y.N. and Yoong, J.S.Y. Cost-effectiveness of an adherence-enhancing intervention for gout based on real-world data. *International Journal of Rheumatic Diseases.* 2018, **22**(4), pp.545-554.
627. Roughley, M., Sultan, A.A., Clarson, L., Muller, S., Whittle, R. et al. Risk of chronic kidney disease in patients with gout and the impact of urate lowering therapy: A population-based cohort study. *Arthritis Res Ther.* 2018, **20**(1), pp.1-10.
628. Vargas-Santos, A.B., Peloquin, C.E., Zhang, Y.Q. and Neogi, T. Association of Chronic Kidney Disease With Allopurinol Use in Gout Treatment. *JAMA Intern Med.* 2018, **178**(11), pp.1526-1533.

Appendices

Appendix A

Table A 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (336)

Section/topic	#	Checklist item	Reported
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Yes
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Yes
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	Yes
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	Yes
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	Yes
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	Yes
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	Yes
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Yes
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Yes
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	Yes
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Yes

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	Yes
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	Yes
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	Yes
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	Not applicable
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Yes
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Yes
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Yes
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Yes
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Not applicable
Synthesis of results	21	Present the main results of the review. If meta-analyses are done, include for each, confidence intervals and measures of consistency	Yes
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Not applicable
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Yes
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	Yes
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	Yes
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	Yes
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	Yes

Table A 2. Aim and study characteristics of gout EHR-based studies (N = 75)

Author (Reference)	Title	Aim	Year Published	Country	Setting
Arromdee et al. (119)	Epidemiology of gout: is the incidence rising?	Epidemiology of gout	2002	USA	Primary care, hospital
Mikuls et al. (126)	Gout epidemiology: results from the UK General Practice Research Database, 1990-1999	Epidemiology of gout; Patient management	2005	UK	Primary care
Mikuls et al. (150)	Suboptimal physician adherence to quality indicators for the management of gout and asymptomatic hyperuricaemia: results from the UK General Practice Research Database (GPRD)	Adherence to clinical guidelines	2005	UK	Primary care
Alonso et al. (131)	Gout and risk of Parkinson disease: a prospective study	Epidemiology of gout	2007	UK	Primary care
Singh et al. (381)	Quality of care for gout in the US needs improvement	Adherence to clinical guidelines	2007	USA	Primary care, hospital
DeVera et al. (601)	Gout and the risk of Parkinson's disease: a cohort study	Epidemiology of gout	2008	Canada	Primary care, hospital
Singh et al. (368)	Opportunities for improving medication use and monitoring in gout	Adherence to clinical guidelines	2008	USA	Primary care, hospital
Harrold et al. (367)	The dynamics of chronic gout treatment: Medication gaps and return to therapy	Adherence and gaps in therapy	2010	USA	Primary care, hospital
Roddy et al. (357)	Prescription and comorbidity screening following consultation for acute gout in primary care	Patient management	2010	UK	Primary care
Wahedduddin et al. (365)	Gout in the Hmong in the United States	Epidemiology of gout	2010	USA	Primary care, hospital
Keenan et al. (353)	Prevalence of contraindications and prescription of pharmacologic therapies for gout	Patient management	2011	USA	Primary care, hospital
Rothenbacher et al. (117)	Frequency and risk factors of gout flares in a large population-based cohort of incident gout	Epidemiology of gout	2011	UK	Primary care

Soriano et al. (602)	Contemporary epidemiology of gout in the UK general population	Epidemiology of gout; Patient management	2011	UK	Primary care
Cheyoe et al. (603)	The prevalence of chronic kidney disease among gout patients in Nongjik hospital, Pattani province	Epidemiology of gout	2012	Thailand	Hospital
Crittenden et al. (384)	Colchicine use is associated with decreased prevalence of myocardial infarction in patients with gout	Treatment safety	2012	USA	Hospital
Harrold et al. (604)	Patients' knowledge and beliefs concerning gout and its treatment: a population based study	Patient knowledge, beliefs and education	2012	USA	Primary care, hospital
Cottrell et al. (60)	Improvement in the management of gout is vital and overdue: an audit from a UK primary care medical practice	Adherence to clinical guidelines	2013	UK	Primary care
Pui et al. (377)	Efficacy and tolerability of probenecid as urate-lowering therapy in gout; clinical experience in high-prevalence population	Treatment effectiveness; Treatment safety	2013	New Zealand	Outpatient
Ryu et al. (379)	Clinical risk factors for adverse events in allopurinol users	Treatment safety	2013	Korea	Hospital
Zandman-Goddard et al. (369)	Rates of adherence and persistence with allopurinol therapy among gout patients in Israel	Adherence and gaps in therapy	2013	Israel	Primary care, hospital
Zarowitz and O'Shea (605)	Demographic and clinical profile of nursing facility residents with gout	Epidemiology of gout; Patient management	2013		Community
George et al. (606)	Evaluating appropriate use of prophylactic colchicine for gout flare prevention	Adherence to clinical guidelines	2014	USA	Primary care
Hatoum et al. (380)	Achieving Serum Urate Goal: A comparative effectiveness study between allopurinol and febuxostat	Treatment effectiveness	2014	USA	Primary care
Jackson et al. (607)	Variation in gout care in Aotearoa New Zealand: a national analysis of quality markers	Adherence to clinical guidelines	2014	Aotearoa New Zealand	Primary care, hospital
Kuo et al. (354)	Eligibility for and prescription of urate-lowering treatment in patients with incident gout in England	Adherence to clinical guidelines	2014	UK	Primary care
Kuo et al. (359)	Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study	Epidemiology of gout; Patient management	2014	UK	Primary care

MacFarlane et al. (608)	The effect of initiating pharmacologic insulin on serum uric acid levels in patients with diabetes: A matched cohort analysis	Epidemiology of gout	2014	USA	Hospital
Meek et al. (609)	Hyperuricaemia: A marker of increased cardiovascular risk in rheumatic patients: Analysis of the ACT-CVD cohort	Epidemiology of gout	2014	Netherlands	Hospital
Park et al. (610)	Clinical factors and treatment outcomes associated with failure in the detection of urate crystal in patients with acute gouty arthritis	Patient management	2014	Korea	Hospital
Rashid et al. (366)	Modifiable factors associated with allopurinol adherence and outcomes among patients with gout in an integrated healthcare system	Adherence and gaps in therapy; treatment effectiveness	2014	USA	Primary care, hospital
Rho et al. (127)	Independent impact of gout on the risk of diabetes mellitus among women and men: a population-based, BMI-matched cohort study	Epidemiology of gout	2014	UK	Primary care
Dennison et al. (389)	Is allopurinol use associated with an excess risk of osteoporotic fracture? A national prescription registry study	Treatment safety	2015	Denmark	Primary care, hospital
Hmar et al. (611)	Understanding and improving the use of allopurinol in a teaching hospital	Adherence to clinical guidelines	2015	Australia	Hospital
Kerr et al. (612)	Measuring physician adherence with gout quality indicators: a role for natural language processing	Adherence to clinical guidelines	2015	USA	Primary care, hospital
Kuo et al. (391)	Effect of allopurinol on all-cause mortality in adults with incident gout: propensity score-matched landmark analysis	Treatment safety	2015	UK	Primary care
Lu et al. (613)	Gout and the risk of Alzheimer's disease: a population-based, BMI-matched cohort study	Epidemiology of gout	2015	UK	Primary care
Mantarro et al. (370)	Allopurinol adherence among patients with gout: an Italian general practice database study	Adherence and gaps in therapy; treatment effectiveness	2015	Italy	Primary care
Rashid et al. (156)	Patient and clinical characteristics associated with gout flares in an integrated healthcare system	Epidemiology of gout	2015	USA	Primary care, hospital
Robinson et al. (614)	An observational study of gout prevalence and quality of care in a national Australian general practice population	Epidemiology of gout; Adherence to clinical guidelines	2015	Australia	Primary care

Thueringer et al. (383)	Anakinra for the treatment of acute severe gout in critically ill patients	Treatment effectiveness; Treatment safety	2015	USA	Hospital
Dehlin et al. (615)	Incidence and prevalence of gout in Western Sweden	Epidemiology of gout; Patient management	2016	Sweden	Primary care, hospital
Kapetanovic et al. (362)	Prevalence and incidence of gout in southern Sweden from the socioeconomic perspective	Epidemiology of gout	2016	Sweden	Primary care, hospital
Kwon et al. (386)	Risk of colchicine-associated myopathy in gout: influence of concomitant use of statin	Treatment safety	2016	Korea	Tertiary referral hospital
Morlock et al. (378)	Disease control, health resource use, healthcare costs, and predictors in gout patients in the United States, the United Kingdom, Germany, and France: A retrospective analysis	Patient management	2016	UK, Germany, France, USA	Primary care, hospital subset
Nyberg et al. (616)	Comorbidity burden in trial-aligned patients with established gout in Germany, UK, US, and France: A retrospective analysis	Epidemiology of gout	2016	UK, Germany, France, USA	Primary care, hospital subset
Solomon et al. (385)	Effects of colchicine on risk of cardiovascular events and mortality among patients with gout: a cohort study using electronic medical records linked with Medicare claims	Treatment safety	2016	USA	Hospital
Chang et al. (617)	Association between gout and aortic stenosis	Epidemiology of gout	2017	USA	Primary care, hospital
Clarson et al. (356)	Factors influencing allopurinol initiation in primary care	Adherence to clinical guidelines	2017	UK	Primary care
Coburn et al. (372)	Allopurinol Medication Adherence as a Mediator of Optimal Outcomes in Gout Management	Adherence and gaps in therapy; treatment effectiveness	2017	USA	Primary care, hospital
Dehlin et al. (358)	Factors associated with initiation and persistence of urate-lowering therapy	Adherence and gaps in therapy	2017	Sweden	Primary care, hospital
Fisher et al. (364)	The unclosing premature mortality gap in gout: a general population-based study	Epidemiology of gout	2017	UK	Primary care

Hughes et al. (373)	Monitoring of urate-lowering therapy among us veterans following the 2012 American College of Rheumatology Guidelines for Management of Gout	Adherence to clinical guidelines	2017	USA	Primary care, hospital
Janssen et al. (382)	Quality of care in gout: a clinical audit on treating to the target with urate lowering therapy in real-world gout patients	Treatment effectiveness	2017	Netherlands	Outpatient
Landgren et al. (376)	Incidence of and risk factors for nephrolithiasis in patients with gout and the general population, a cohort study	Epidemiology of gout	2017	Sweden	Primary care, hospital
Lee et al. (618)	Elderly patients exhibit stronger inflammatory responses during gout attacks	Epidemiology of gout	2017	Korea	Primary care, hospital
Maravic et al. (375)	Persistent clinical inertia in gout in 2014: An observational French longitudinal patient database study	Epidemiology of gout; Patient management	2017	France	Primary care
Olaru et al. (619)	Coexistent rheumatoid arthritis and gout: a case series and review of the literature	Epidemiology of gout	2017	USA	Outpatient
Rai et al. (363)	The rising prevalence and incidence of gout in British Columbia, Canada: Population-based trends from 2000 to 2012	Epidemiology of gout; Patient management	2017	Canada	Primary care, hospital
Sigurdardottir et al. (360)	Work disability in gout: a population-based case-control study	Epidemiology of gout	2017	Sweden	Primary care, hospital
Spaetgens et al. (387)	Risk of infections in patients with gout: a population-based cohort study	Epidemiology of gout	2017	UK	Primary care
Sultan et al. (374)	Gout and subsequent erectile dysfunction: a population- based cohort study from England	Epidemiology of gout	2017	UK	Primary care
Bevis et al. (620)	Comorbidity clusters in people with gout: an observational cohort study with linked medical record review	Epidemiology of gout	2018	UK	Primary care
Coburn et al. (371)	Allopurinol dose escalation and mortality among patients with gout: a national propensity-matched cohort study	Treatment safety	2018	USA	Primary care, hospital
Dehlin and Jacobsson (621)	Trends in gout hospitalization in Sweden	Patient management	2018	Sweden	Hospital

Hassan and Choudry (622)	The compliance of guidelines set by the British Society for Rheumatology for managing Gout	Adherence to clinical guidelines	2018		Primary care
Jung et al. (623)	Effect of fenofibrate on uric acid level in patients with gout	Treatment effectiveness	2018	Korea	Hospital
Keller et al. (624)	Statin use and mortality in gout: A general population-based cohort study	Treatment safety	2018	UK	Primary care
Kuo et al. (390)	Urate-lowering treatment and risk of total joint replacement in patients with gout	Epidemiology of gout; Treatment safety	2018	UK (also Taiwan but not EHR)	Primary care
Lee et al. (625)	Hepatic Safety of Febuxostat Compared with Allopurinol in Gout Patients with Fatty Liver Disease	Treatment safety	2018	Korea	Hospital
Lin et al. (626)	Cost-effectiveness of an adherence-enhancing intervention for gout based on real-world data	Patient management	2018	Singapore	Outpatient
Roughley et al. (627)	Risk of chronic kidney disease in patients with gout and the impact of urate lowering therapy: A population-based cohort study	Epidemiology of gout; Treatment safety	2018	UK	Primary care
Scheepers et al. (355)	Medication adherence among gout patients initiated allopurinol: a retrospective cohort study in the Clinical Practice Research Datalink (CPRD)	Adherence and gaps in therapy	2018	UK	Primary care
Sultan et al. (388)	Risk of fragility fracture among patients with gout and the effect of urate-lowering therapy	Epidemiology of gout; Treatment safety	2018	UK	Primary care
Vargas-Santos et al. (628)	Association of chronic kidney disease with allopurinol use in gout treatment	Treatment safety	2018	UK	Primary care
Mikuls et al. (352)	Adherence and outcomes with urate-lowering therapy: a site-randomized trial	Adherence and gaps in therapy	2019	USA	Primary care, hospital

Note: UK = United Kingdom; US = United States, USA = United States of America; ACT-CVD = Arthritis Center Twente CardioVascular Disease; BMI = Body mass index

Table A 3. Detail on the gout definition, gout cohort, reporting and study quality of gout EHR-based studies (N = 75)

Author (Reference)	Gout Definition Detail*	Gout Definition	Size	Male, %	Mean Age (SD)	Gout Code Provided	CoR Score	RoB Score
Arromdee et al. (119)	Meet ACR criteria, adjudicated by rheumatologist	Stringent	120	76.7			4.1	10.0
Mikuls et al. (126)	≥1 diagnosis	Liberal	63105				5.0	11.4
Mikuls et al. (150)	≥1 diagnosis	Liberal	63105	78	61 (15)		4.1	11.7
Alonso et al. (131)	≥1 diagnosis	Liberal					5.9	14.3
Singh et al. (381)	≥1 diagnosis	Liberal	663	99.4	67.9 (9.7)	Diagnosis	4.2	14.2
DeVera et al. (601)	≥1 diagnosis	Liberal	11258	66.5	74.1 (6.5)	Diagnosis	5.5	12.1
Singh et al. (368)	≥1 diagnosis	Liberal	643	99.4	67.9 (9.7)		3.5	13.3
Harrold et al. (367)	≥1 diagnosis	Liberal	4166	76	63 (15)		3.5	12.5
Roddy et al. (357)	≥1 diagnosis	Liberal	673	77	63.2 (14.5)	Diagnosis	5.0	7.5
Wahedduddin et al. (365)	≥2 diagnoses , ≥2 diagnoses and meet ACR criteria	Stringent	173	80.92			5.8	11.4
Keenan et al. (353)	≥1 diagnosis; ≥1 diagnosis and current prescription, crystal identification, evidence of tophus or a score >5/12 clinical criteria for diagnosis	Stringent	575	99.3	71.75 (11.64)	Diagnosis	5.9	13.3
Rothenbacher et al. (117)	≥1 diagnosis	Liberal	23857	72.8	61.9 (14.5)		6.4	10.8
Soriano et al. (602)	≥1 diagnosis or drug; ≥1 diagnosis and drug	Stringent	24768	72.46			5.5	12.5
Cheyoe et al. (603)	≥1 diagnosis	Liberal	154	75.3	63.1 (12.9)	Diagnosis	3.6	11.3

Crittenden et al. (384)	≥1 diagnosis	Liberal	1288	99	71.3	Diagnosis	6.4	10.0
Harrold et al. (604)	≥1 diagnosis and survey response	Stringent	479			Diagnosis	3.6	8.3
Cottrell et al. (60)	≥1 diagnosis	Liberal	305	74	65.5	Diagnosis	6.8	12.5
Pui et al. (377)	Meet ACR criteria	Stringent	57	77	57 (16)		2.3	12.1
Ryu et al. (379)	Not specified	NS	66				1.8	11.7
Zandman-Goddard et al. (369)	≥1 specialist diagnosis	Stringent	7644	72		Diagnosis	5.0	10.8
Zarowitz and O'Shea (605)	≥1 diagnosis	Liberal	2487	52.8		Diagnosis	6.9	12.5
George et al. (606)	Not specified	NS	126	99.2			3.5	9.0
Hatoum et al. (380)	≥1 diagnosis	Liberal	18389	69	63.7 (13.37)		5.0	12.1
Jackson et al. (607)	≥1 diagnosis or drug	Liberal	114703			Diagnosis	4.6	11.7
Kuo et al. (354)	Not specified	NS	52164	73.37	62.5		3.2	10.8
Kuo et al. (359)	≥1 diagnosis	Liberal	115608			Diagnosis	6.4	12.5
MacFarlane et al. (608)	Not specified	NS	46	47.8	57		1.9	13.3
Meek et al. (609)	Not specified	NS	172	89	59.6 (10.8)		2.3	15.0
Park et al. (610)	≥1 diagnosis and test	Stringent	179	94.4	62.6 (16.4)		4.2	8.8
Rashid et al. (366)	≥2 diagnoses	Stringent	13341	78	60.2 (13.9)		5.9	13.3
Rho et al. (127)	≥1 diagnosis; ≥1 diagnosis and drug	Stringent	35339	72.57	62.7		5.0	11.4

Dennison et al. (389)	Not specified	NS	86039	67	63 (15.1)		4.6	12.9
Hmar et al. (611)	≥1 drug	Liberal	1304	75	74		5.0	12.5
Kerr et al. (612)	≥2 diagnoses	Stringent	2280	99.1	66.8 (12.2)		5.0	8.3
Kuo et al. (391)	≥1 diagnosis	Liberal	23332	73.91			6.4	13.6
Lu et al. (613)	≥1 diagnosis; ≥1 diagnosis and drug	Stringent	59224	70.8	65.3 (12.2)		6.4	11.4
Mantarro et al. (370)	≥1 diagnosis or free-text keyword in notes	Liberal	3570	80	65	Diagnosis and Medication	5.4	12.1
Rashid et al. (156)	≥2 outpatient or ≥1 inpatient diagnosis	Stringent	8828	81	55-64	Diagnosis	7.7	12.5
Robinson et al. (614)	≥1 diagnosis or drug	Liberal	22776	81.59			6.2	10.0
Thueringer et al. (383)	Not specified	NS	13	92.3	58		2.7	13.0
Dehlin et al. (615)	≥1 diagnosis; 1 primary diagnosis; 1 rheumatology diagnosis or ≥2 diagnoses	Stringent	30430	70	69 (14)	Diagnosis and Medication	8.1	15.0
Kapetanovic et al. (362)	≥1 diagnosis; ≥2 diagnoses or ≥1 hospital physician diagnosis	Stringent			69	Diagnosis and Medication	5.8	15.0
Kwon et al. (386)	Meet Preliminary criteria	Stringent	674	97			2.7	10.0
Morlock et al. (378)	≥1 diagnosis or drug and CKD, urolithiasis, tophus or 2+ flares	Stringent	197805				4.6	11.4
Nyberg et al. (616)	≥1 diagnosis or drug and CKD, urolithiasis, tophus or 2+ flares	Stringent	198654				6.2	11.3
Solomon et al. (385)	≥1 diagnosis	Liberal	1002	63.7	72.5 (11.5)		6.5	12.9
Chang et al. (617)	≥1 diagnosis or free-text keyword in notes	Liberal	52				2.3	12.1
Clarson et al. (356)	≥1 diagnosis	Liberal	8142	70.9	65.4 (10.2)		5.0	10.8

Coburn et al. (372)	≥1 diagnosis and survey response	Stringent	612	98	72.1 (10.7)		6.5	10.7
Dehlin et al. (358)	≥1 diagnosis	Liberal	7709	68	66	Diagnosis and Medication	7.3	11.7
Fisher et al. (364)	≥1 diagnosis; ≥1 diagnosis and drug	Stringent	103261	74	61.9 (15)		5.0	13.3
Hughes et al. (373)	≥1 diagnosis	Liberal	505	97	62.3 (12.4)		5.0	12.5
Janssen et al. (382)	Not specified	NS	177		67.5 (12)		4.1	15.0
Landgren et al. (376)	≥1 diagnosis	Liberal	29171			Diagnosis and Medication	7.3	12.1
Lee et al. (618)	≥1 hospital diagnosis with or without MSU crystals and adjudicated by Rheumatologist	Stringent	254	87.8		Diagnosis	4.1	10.0
Maravic et al. (375)	≥1 free-text keyword in notes	Liberal	14400	84.4	67.5		5.9	10.0
Olaru et al. (619)	≥1 test	Stringent	13	69.23	68.6		5.9	12.5
Rai et al. (363)	≥1 diagnosis; ≥2 diagnoses or ≥1 primary diagnosis	Stringent	171165	68	63 (15.4)	Diagnosis	5.0	15.0
Sigurdardottir et al. (360)	≥1 diagnosis	Liberal	4571	77	51.3 (8.3)		8.5	12.5
Spaetgens et al. (387)	≥1 diagnosis	Liberal	131565	73.9	64 (13.5)		6.5	12.1
Sultan et al. (374)	≥1 diagnosis	Liberal	9653	100			5.0	13.6
Bevis et al. (620)	≥1 diagnosis or drug and survey response	Stringent	1079	84	65.5 (12.5)	Diagnosis	5.8	13.3
Coburn et al. (371)	≥2 diagnoses	Stringent	25378	99.7		Diagnosis	5.4	12.1

Dehlin and Jacobsson (621)	≥1 diagnosis	Liberal				Diagnosis and Medication	6.2	15.0
Hassan and Choudry (622)	≥1 drug	Liberal	112				3.6	11.7
Jung et al. (623)	≥1 diagnosis and chart review	Stringent	863	91.4	50.6 (14.9)	Diagnosis	6.8	12.5
Keller et al. (624)	≥1 diagnosis; ≥1 diagnosis and drug	Stringent	36014	81	63		4.1	12.1
Kuo et al. (390)	≥1 diagnosis	Liberal	34505		61.4 (14.8)	Diagnosis	6.3	11.4
Lee et al. (625)	Meet ACR criteria	Stringent	134				3.2	10.0
Lin et al. (626)	≥1 test	Stringent	53	73.6	61.3 (15.12)		7.3	13.6
Roughley et al. (627)	≥1 diagnosis	Liberal	41446	81	57.2 (13.6)		5.8	12.1
Scheepers et al. (355)	≥1 diagnosis	Liberal	48280	76	64.6 (13.2)		5.9	11.7
Sultan et al. (388)	≥1 diagnosis	Liberal	31781	72.9	63.5 (12.5)		5.9	11.4
Vargas-Santos et al. (628)	≥1 diagnosis	Liberal	9520	83.5	57.4 (13.6)		7.3	12.9
Mikuls et al. (352)	≥1 diagnosis	Liberal	782	82.7	58.0 (14.4)	Diagnosis	5.0	11.4

Note: *All elements were coded except free-text keywords; where studies use multiple definitions, these are separated by ";". ACR, American College of Rheumatology; CKD, chronic kidney disease; CoR, comprehensiveness of reporting; MSU, monosodium urate; NS, not stated; RoB, risk of bias; SD, standard deviation

Appendix B

Table B 1. Read Codes used to determine diagnosis of rheumatoid arthritis, in comparison to the code-list used by Abhishek et al (2017) (265)

Code	Term Description	Study code	Abhishek et al.
F3712	Polyneuropathy+rheumatoid arth	Y	Y
F3964	Myopathy+rheumatoid arthritis	Y	Y
N040.	Rheumatoid arthritis	Y	Y
N0400	Rheumatoid arthritis-Cx spine	Y	Y
N0401	Oth rheumatoid arthritis-spine	Y	Y
N0402	Rheumatoid arthritis-shoulder	Y	Y
N0405	Rheumatoid arthritis of elbow	Y	Y
N0406	Rheumatoid arthritis-dist RUJ	Y	Y
N0407	Rheumatoid arthritis of wrist	Y	Y
N0408	Rheumatoid arthritis-MCP joint	Y	Y
N0409	Rheumatoid arthritis-PIPJ-fing	Y	Y
N040A	Rheumatoid arthritis-DIPJ-fing	Y	Y
N040B	Rheumatoid arthritis of hip	Y	Y
N040D	Rheumatoid arthritis of knee	Y	Y
N040F	Rheumatoid arthritis of ankle	Y	Y
N040G	Rheumatoid arthr-subtalar jnt	Y	Y
N040H	Rheumatoid arthr-talonav joint	Y	Y
N040J	Rheumatoid arthr-oth tarsal jt	Y	Y
N040K	Rheumatoid arthr-1st MTP joint	Y	Y
N040P	Seronegative rheumat arthritis	Y	Y
N040S	Rheumat arthr - multiple joint	Y	Y
N040T	Flare of rheumatoid arthritis	Y	Y
N041.	Felty's syndrome	Y	Y
N042.	Other rh.arthr.+visc/syst.dis.	Y	Y
N042z	Rh.arthr.+visc/syst.dis.NOS	Y	Y
N047.	Seropositive errosive RA	Y	Y
N04X.	Seroposit rheum arthr, unsp	Y	Y
Nyu11	[X]O sero+ve rheumat arthritis	Y	Y
Nyu12	[X]Oth spcf rheumatd arthritis	Y	Y
Nyu1G	[X]Seroposit rheum arthr, unsp	Y	Y
14G1.	H/O: rheumatoid arthritis	Y	

38DZ.	Diseas activ scor rheu arthrit	Y	
38DZ0	DAS 28 joint rheumatoid arthri	Y	
38Vs.	RAID questionnaire	Y	
66HB0	Rheumatoid arthritis annul rev	Y	
7P203	Del rehab rheumatoid arthritis	Y	
9hR..	Exp rep: rheumtd arth qual ind	Y	
9hR0.	Ex rhm art qual ind: pt unsuit	Y	
9hR1.	Ex rheum arth qua ind: inf dis	Y	
9mM..	RA monitoring invitation	Y	
9mM0.	RA monitor invitation 1st lett	Y	
9mM1.	RA monitor invitation 2nd lett	Y	
9mM2.	RA monitor invitation 3rd lett	Y	
9mM3.	RA monitoring verbal invitatin	Y	
9mM4.	RA monitor telephone invitatin	Y	
N0403	Rheumatoid arthr-sternoclav jt	Y	
N0404	Rheumatoid arthr-acromioclav j	Y	
N040C	Rheumatoid arthritis of SIJ	Y	
N040E	Rheumatoid arthr of tib-fib jt	Y	
N040L	Rheumatoid arthr-lesser MTP jt	Y	
N040M	Rheumatoid arthr-IP joint-toe	Y	
Nyu10	[X]Rheum arthrit+inv/o org/sys	Y	
N043.	Juvenile rheumatoid arthritis - Still's disease		Y
N0422	Rheumatoid nodule		Y
N045.	Other juvenile arthritis		Y
H570.	Rheumatoid lung		Y
66H..	Rheumatoid arthrit. monitoring		Y
N040Q	Rheumatoid bursitis		Y
N0432	Pauciarticular juvenile rheumatoid arthritis		Y
N3622	Swan-neck finger deformity		Y
N043z	Juvenile rheumatoid arthritis NOS		Y
N04..	Rheumatoid arthritis and other inflammatory polyarthropathy		Y
N04y0	Fibrosing alveolitis associated with rheumatoid arthritis		Y
N040N	Rheumatoid vasculitis		Y
N0451	Juvenile seronegative polyarthritis		Y
N0455	Juvenile rheumatoid arthritis		Y

N04y0	Rheumatoid lung	Y
2G27.	O/E-hands-rheumatoid spindling	Y
N0433	Monarticular juvenile rheumatoid arthritis	Y
2G25.	O/E - ulnar deviation	Y
G5yA.	Rheumatoid carditis	Y
N0421	Rheumatoid lung disease	Y
N0456	Pauciarticular onset juvenile chronic arthritis	Y
N0431	Acute polyarticular juvenile rheumatoid arthritis	Y
G5y8.	Rheumatoid myocarditis	Y
N0430	Juvenile rheumatoid arthropathy unspecified	Y
N040R	Rheumatoid nodule	Y
2G25.	O/E - hands - ulnar deviation	Y
N04y0	Caplan's syndrome	Y
Nyu15	[X]Other juvenile arthritis	Y

Table B 2. Read Codes used to define osteoporosis

Code	Term Description
N331L00	collapse of vertebra due to osteoporosis nos
N331J00	collapse of lumbar vertebra due to osteoporosis
N331M00	fragility fracture due to unspecified osteoporosis
N331900	osteoporosis + pathological fracture thoracic vertebrae
N331800	osteoporosis + pathological fracture lumbar vertebrae
NyuB800	[x]unspecified osteoporosis with pathological fracture
N331K00	collapse of thoracic vertebra due to osteoporosis
N331600	idiopathic osteoporosis with pathological fracture
N331300	osteoporosis of disuse with pathological fracture
N331.14	osteoporotic vertebral collapse
N331H00	collapse of cervical vertebra due to osteoporosis
N331A00	osteoporosis + pathological fracture cervical vertebrae
N331M11	minimal trauma fracture due to unspecified osteoporosis
66aB.00	osteoporosis - no treatment response
66aA.00	osteoporosis - treatment response
N330.00	Osteoporosis
N330B00	vertebral osteoporosis
N330000	osteoporosis, unspecified

N330C00	osteoporosis localized to spine
66a9.00	osteoporosis - falls prevention
N330z00	osteoporosis nos
N374600	osteoporotic kyphosis
N330300	idiopathic osteoporosis
N330800	localized osteoporosis – lequesne
N330400	dissuse osteoporosis
9hP..00	exception reporting: osteoporosis quality indicators
N330100	senile osteoporosis
N331B00	postmenopausal osteoporosis with pathological fracture
N330200	postmenopausal osteoporosis
N331500	drug-induced osteoporosis with pathological fracture
N330500	drug-induced osteoporosis
N330D00	osteoporosis due to corticosteroids
N330A00	osteoporosis in endocrine disorders
N331200	postoophorectomy osteoporosis with pathological fracture
N330600	postoophorectomy osteoporosis
N331400	postsurgical malabsorption osteoporosis with path fracture
N330700	postsurgical malabsorption osteoporosis
NyuB000	[x]other osteoporosis with pathological fracture
NyuB100	[x]other osteoporosis
N330900	osteoporosis in multiple myelomatosis
NyuB200	[x]osteoporosis in other disorders classified elsewhere

Appendix C

Regression Modelling Code

The following Poisson regression modelling analysis was run using R version 3.6.1, for all RA patients and for incident RA patients in the year following diagnosis, and the subset of each with IMD recorded.

Fixed effects Poisson regression

(chosen for: long-term corticosteroid prescribing in the year following incident diagnosis in the full cohort (Table 39); long-term corticosteroid prescribing for all RA patients in the IMD subset; long-term corticosteroid prescribing in the year following incident diagnosis in the IMD subset; long-term NSAID prescribing in the year following incident diagnosis in the IMD subset)

```
summary(model1 <- glm(longTermPrescribing ~ Year + Sex + AgeGroup + imd2015_5
+ offset(logPopulation), data = allData, family = poisson(link = "log")))
```

(chosen for: long-term NSAID prescribing for all RA patients in the IMD subset)

```
summary(model2 <- glm(longTermPrescribing ~ Year + Sex + AgeGroup + imd2015_5
+ offset(logPopulation), data = subsetIMD, family = poisson(link = "log")))
```

Fixed effects quasi-Poisson regression

(chosen for long-term corticosteroid and NSAID prescribing in all RA patients (Table 39))

```
summary(model3 <- glm(longTermPrescribing ~ Year + Sex + AgeGroup + imd2015_5
+ offset(logPopulation), data = allData, family = quasipoisson(link = "log")))
```

```
summary(model4 <- glm(longTermPrescribing ~ Year + Sex + AgeGroup + imd2015_5
+ offset(logPopulation), data = subsetIMD, family = quasipoisson(link = "log")))
```

Random effects Poisson regression

(chosen for long-term NSAID prescribing in the year following incident diagnosis (Table 39))

summary(model5 <- glmer(longTermPrescribing ~ Year + Sex + AgeGroup + offset(logpopn) + (1 | ProjectPracID), data = allData, family = poisson(link = "log"), nAGQ=0))

summary(model6 <- glmer(longTermPrescribing ~ Year + Sex + AgeGroup + + imd2015 + offset(logpopn) + (1 | ProjectPracID), data = subsetIMD, family = poisson(link = "log"), nAGQ=0))

Zero-inflation Poisson regression

summary(model7 <- mixed_model(longTermPrescribing ~ Year + Sex + AgeGroup + offset(logpopn), random= ~1 | ProjectPracID, data = allData, family = zi.poisson(), zi_fixed=~Year, zi_random= ~1 | ProjectPracID))

summary(model8 <- mixed_model(longTermPrescribing ~ Year + Sex + AgeGroup + imd2015 + offset(logpopn), random= ~1 | ProjectPracID, data = allData, family = zi.poisson(), zi_fixed=~Year, zi_random= ~1 | ProjectPracID))

For each model, robust standard errors and P values with 95% confidence intervals were obtained using the parameter estimates and their standard errors (presented for model1 as an example)

cov.m1<-vcovHC(model1, type="HC0")

std.err<-sqrt(diag(cov.m1))

r.est<-cbind(Estimate=coef(model1), "Robust SE" = std.err, "P(>|z|)" = 2*pnorm(abs(coef(model1)/std.err), lower.tail=FALSE), LL=coef(model1)-1.96*std.err, UL = coef(model1) + 1.96*std.err)

r.est

Incidence rate ratios were calculated, with the Delta method to define standard error (presented for model1 as an example)

g<-deltamethod(list(~exp(x1),~exp(x2),~exp(x3),~exp(x4)), coef(model1), cov.model1)

rexp.est<-exp(r.est[,-3])

rexp.est[, "Robust SE"] <- g

rexp.est

Fixed and random effects models were compared using the Hausman test:

```
library(plm)
phtest(model1, model5)
```

Appendix D

Table D 1. Poisson regression and quasi-Poisson regression modelling of corticosteroid prescribing and Poisson regression with GP practice as a random intercept: coefficient estimates, their standard errors (SEs) and *z*- or *t*- value

Variable	Poisson			Quasi-Poisson			Poisson with random intercept		
	Estimate	SE	*z*-value	Estimate	SE	*t*-value	Estimate	SE	*z*-value
Intercept	-1.981	0.053	-37.547	-1.981	0.067	-29.519	-2.023	0.055	-36.695
Year									
1999	0.021	0.028	0.721	0.021	0.036	0.567	0.019	0.029	0.651
2000	0.051	0.027	1.891	0.054	0.035	1.566	0.033	0.027	1.205
2001	0.042	0.026	1.585	0.042	0.033	1.246	0.026	0.026	1.002
2002	0.022	0.026	0.847	0.022	0.033	0.666	0.013	0.026	0.497
2003	0.011	0.025	0.435	0.011	0.032	0.342	0.012	0.026	0.458
2004	-0.007	0.025	-0.270	-0.004	0.032	-0.126	-0.004	0.025	-0.169
2005	-0.021	0.025	-0.832	-0.021	0.032	-0.654	-0.016	0.025	-0.656
2006	-0.038	0.025	-1.548	-0.038	0.031	-1.217	-0.033	0.025	-1.318
2007	-0.037	0.025	-1.503	-0.037	0.031	-1.181	-0.034	0.025	-1.345
2008	-0.048	0.025	-1.949	-0.045	0.031	-1.445	-0.047	0.025	-1.855
2009	-0.057	0.025	-2.307	-0.057	0.031	-1.813	-0.052	0.025	-2.050
2010	-0.071	0.025	-2.850	-0.071	0.032	-2.240	-0.066	0.025	-2.604
2011	-0.085	0.025	-3.386	-0.085	0.032	-2.662	-0.083	0.025	-3.280
2012	-0.094	0.025	-3.743	-0.091	0.032	-2.856	-0.096	0.026	-3.768
2013	-0.135	0.025	-5.308	-0.135	0.032	-4.173	-0.136	0.026	-5.243
2014	-0.188	0.026	-7.323	-0.188	0.033	-5.757	-0.188	0.026	-7.143
2015	-0.215	0.026	-8.186	-0.215	0.033	-6.434	-0.216	0.027	-8.038
2016	-0.231	0.027	-8.470	-0.229	0.035	-6.579	-0.247	0.028	-8.781
2017	-0.288	0.029	-10.079	-0.288	0.036	-7.923	-0.303	0.029	-10.293
Sex									
Female	-0.044	0.007	-6.251	-0.044	0.009	-4.914	-0.041	0.007	-5.763
Age-group									
30-39	0.072	0.053	1.366	0.072	0.067	1.074	0.052	0.053	0.990
40-49	0.048	0.050	0.957	0.048	0.064	0.752	0.027	0.050	0.532
50-59	0.240	0.049	4.903	0.240	0.062	3.854	0.226	0.049	4.614
60-69	0.514	0.049	10.583	0.514	0.062	8.317	0.500	0.049	10.264
70-79	0.731	0.049	15.061	0.731	0.062	11.836	0.712	0.049	14.642

80-89	0.781	0.049	16.026	0.781	0.062	12.595	0.764	0.049	15.627
90-99	0.469	0.053	8.818	0.469	0.068	6.930	0.459	0.053	8.605

Table D 2. Poisson regression modelling of corticosteroid prescribing in the year post-diagnosis and Poisson regression with GP practice as a random intercept: Coefficient estimates, their standard errors (SEs) and *z*-value

Variable	Poisson			Poisson with random intercept		
	Estimate	SE	*z*-value	Estimate	SE	*z*-value
Intercept	-1.917	0.130	-14.797	-1.951	0.131	-14.947
Year						
1999	0.027	0.107	0.247	0.005	0.108	0.043
2000	0.026	0.102	0.257	0.004	0.102	0.041
2001	-0.030	0.099	-0.300	-0.040	0.100	-0.399
2002	-0.002	0.098	-0.022	-0.016	0.098	-0.163
2003	0.053	0.095	0.561	0.032	0.096	0.336
2004	-0.002	0.094	-0.022	-0.012	0.095	-0.129
2005	0.029	0.094	0.306	0.011	0.095	0.111
2006	0.098	0.094	1.036	0.092	0.095	0.969
2007	0.215	0.093	2.305	0.205	0.094	2.176
2008	0.167	0.094	1.776	0.154	0.095	1.619
2009	0.171	0.093	1.829	0.158	0.094	1.678
2010	0.168	0.094	1.777	0.155	0.096	1.621
2011	0.126	0.095	1.334	0.141	0.096	1.470
2012	0.148	0.094	1.578	0.170	0.095	1.782
2013	-0.261	0.093	-2.810	-0.264	0.094	-2.802
2014	-0.292	0.096	-3.051	-0.288	0.097	-2.975
2015	-0.244	0.101	-2.420	-0.200	0.103	-1.945
2016	-0.086	0.103	-0.830	-0.077	0.105	-0.734
2017	-0.321	0.129	-2.484	-0.308	0.131	-2.361
Sex						
Female	-0.123	0.025	-4.838	-0.116	0.025	-4.540
Age-group						
30-39	0.045	0.116	0.385	0.047	0.116	0.407
40-49	0.031	0.109	0.282	0.038	0.108	0.354
50-59	0.285	0.105	2.723	0.291	0.104	2.783
60-69	0.550	0.103	5.313	0.547	0.103	5.296

70-79	0.853	0.103	8.268	0.851	0.103	8.267
80-89	0.956	0.106	9.033	0.952	0.106	9.007
90-99	0.757	0.156	4.844	0.744	0.156	4.760

Table D 3. Poisson regression and quasi-Poisson regression modelling of NSAID prescribing and Poisson regression with GP practice as a random intercept: Coefficient estimates, their standard errors (SEs) and *z*- or *t*- value

Variable	Poisson			Quasi-Poisson			Poisson with random intercept		
	Estimate	SE	*z*-value	Estimate	SE	*t*-value	Estimate	SE	*z*-value
Intercept	-0.849	0.033	-26.123	-0.850	0.038	-22.146	-0.836	0.034	-24.792
Year									
1999	0.003	0.019	0.167	0.003	0.023	0.142	0.005	0.019	0.242
2000	-0.038	0.019	-2.031	-0.035	0.022	-1.597	-0.041	0.019	-2.171
2001	-0.054	0.018	-3.010	-0.054	0.021	-2.550	-0.054	0.018	-2.946
2002	-0.066	0.018	-3.744	-0.066	0.021	-3.171	-0.068	0.018	-3.787
2003	-0.086	0.017	-4.946	-0.086	0.021	-4.189	-0.085	0.018	-4.837
2004	-0.087	0.017	-5.038	-0.084	0.020	-4.133	-0.089	0.017	-5.129
2005	-0.053	0.017	-3.156	-0.053	0.020	-2.673	-0.061	0.017	-3.539
2006	-0.120	0.017	-7.081	-0.120	0.020	-5.999	-0.129	0.017	-7.485
2007	-0.173	0.017	-10.105	-0.173	0.020	-8.560	-0.182	0.017	-10.489
2008	-0.214	0.017	-12.463	-0.211	0.020	-10.423	-0.226	0.017	-12.932
2009	-0.260	0.017	-15.038	-0.260	0.020	-12.739	-0.273	0.018	-15.496
2010	-0.300	0.017	-17.178	-0.300	0.021	-14.552	-0.313	0.018	-17.593
2011	-0.319	0.018	-18.110	-0.319	0.021	-15.341	-0.334	0.018	-18.590
2012	-0.366	0.018	-20.588	-0.363	0.021	-17.310	-0.381	0.018	-21.019
2013	-0.388	0.018	-21.553	-0.388	0.021	-18.258	-0.406	0.018	-22.087
2014	-0.419	0.018	-22.981	-0.419	0.022	-19.467	-0.444	0.019	-23.826
2015	-0.437	0.019	-23.406	-0.437	0.022	-19.827	-0.464	0.019	-24.260
2016	-0.463	0.020	-23.650	-0.460	0.023	-19.915	-0.501	0.020	-24.902
2017	-0.558	0.021	-26.793	-0.558	0.025	-22.696	-0.608	0.021	-28.398
Sex									
Female	-0.002	0.005	-0.291	-0.002	0.006	-0.247	-0.001	0.005	-0.114
Age-group									
30-39	0.136	0.032	4.314	-0.002	0.006	-0.247	0.122	0.032	3.846
40-49	0.274	0.030	9.160	-0.002	0.006	-0.247	0.259	0.030	8.627

50-59	0.289	0.029	9.816	-0.002	0.006	-0.247	0.272	0.030	9.177
60-69	0.219	0.029	7.452	-0.002	0.006	-0.247	0.199	0.029	6.753
70-79	-0.005	0.029	-0.186	-0.002	0.006	-0.247	-0.022	0.030	-0.733
80-89	-0.303	0.030	-10.061	-0.002	0.006	-0.247	-0.314	0.030	-10.382
90-99	-0.578	0.037	-15.620	-0.002	0.006	-0.247	-0.584	0.037	-15.751

Table D 4. Poisson regression modelling of NSAID prescribing in the year post-diagnosis and Poisson regression with GP practice as a random intercept: Coefficient estimates, their standard errors (SEs) and *z*-value

Variable	Poisson			Poisson with random intercept		
	Estimate	SE	*z*-value	Estimate	SE	*z*-value
Intercept	-0.601	0.078	-7.747	-0.595	0.078	-7.624
Year						
1999	-0.005	0.068	-0.080	-0.004	0.068	-0.063
2000	-0.038	0.065	-0.583	-0.038	0.065	-0.589
2001	-0.092	0.063	-1.472	-0.092	0.063	-1.453
2002	-0.067	0.061	-1.092	-0.067	0.062	-1.085
2003	-0.115	0.061	-1.892	-0.114	0.061	-1.871
2004	-0.049	0.059	-0.826	-0.045	0.060	-0.757
2005	-0.090	0.060	-1.509	-0.088	0.060	-1.460
2006	-0.197	0.061	-3.219	-0.198	0.062	-3.209
2007	-0.189	0.061	-3.079	-0.193	0.062	-3.108
2008	-0.279	0.063	-4.443	-0.280	0.063	-4.427
2009	-0.389	0.063	-6.136	-0.391	0.064	-6.121
2010	-0.361	0.064	-5.639	-0.360	0.065	-5.581
2011	-0.451	0.065	-6.973	-0.442	0.065	-6.773
2012	-0.518	0.066	-7.908	-0.502	0.066	-7.583
2013	-0.583	0.061	-9.575	-0.580	0.061	-9.444
2014	-0.730	0.064	-11.354	-0.730	0.065	-11.256
2015	-0.635	0.068	-9.341	-0.607	0.069	-8.794
2016	-0.699	0.074	-9.484	-0.683	0.075	-9.143
2017	-0.904	0.097	-9.279	-0.908	0.098	-9.259
Sex						
Female	0.021	0.019	1.089	0.021	0.019	1.070
Age-group						
30-39	0.126	0.067	1.894	0.119	0.067	1.781

40-49	0.158	0.062	2.537	0.155	0.062	2.473
50-59	0.147	0.061	2.418	0.142	0.061	2.322
60-69	0.088	0.061	1.444	0.081	0.061	1.335
70-79	-0.096	0.062	-1.544	-0.101	0.062	-1.626
80-89	-0.322	0.068	-4.711	-0.327	0.069	-4.772
90-99	-0.294	0.127	-2.305	-0.294	0.128	-2.306