Epidemiology of Osteoarthritis and associated comorbidities in the United Kingdom

By Subhashisa Swain, PT, MPH

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

October 2020

Declaration

I hereby declare that this thesis is the result of original research. It has largely been conducted by me and any assistance received is detailed below. It has not been accepted elsewhere for any degree, diploma, or other qualification. All the authors and works to which reference has been made are fully acknowledged.

Abstract

Background

Osteoarthritis (OA) is very common and is the main cause of chronic joint pain and disability in older people. According to this systematic review nearly 67% of people with OA had comorbidity. There is little information available on how the incidence and prevalence of OA has changed over the past 20 years in the UK, and what is the likelihood of having other chronic conditions, their progression, and associated outcomes.

Objectives

This research aimed to answer five questions: 1) how common is osteoarthritis in the UK and what are the trends over the past twenty years; 2) are people with osteoarthritis more likely to have other chronic conditions and multimorbidity (two or more conditions in an individual) than people without osteoarthritis; 3) in people with OA how do these long-term conditions coexist; 4) how does the group of long-term conditions progress with time; and 5) does the presence of long-term conditions in osteoarthritis add to the burden both to patients and to health services.

Methods

A large nationally representative UK primary care database known as the Clinical Practice Research Datalink (CPRD) GOLD was used for the study. Six different studies were performed in this thesis in people aged 20 years or more with OA and age, sex, and practice matched controls. These are: 1) epidemiology of osteoarthritis in the UK (chapter 3); 2) risk of comorbidities occurring before and after the diagnosis of osteoarthritis using both case-control and cohort design (chapter 4); 3) clusters of multimorbidity in people with OA and controls using latent class analysis (chapter 5); 4) illness pathways (transition and trajectories) of multimorbidity clusters in people with OA and controls using latent transition analysis and latent class growth analysis, respectively (chapter 6 and 7); 5) outcomes such

as all-cause mortality, outpatient visits, inpatient admission and disability adjusted life years (DALYs) associated with OA and their comorbidities (chapter 8).

Results

The prevalence of OA in the UK primary care in 2017 was 10.7% and the incidence was 6.8 per 1000 person-years in people aged 20 and over. OA was more common in women compared to men and increased with age, especially after age 40 years. The prevalence has increased at a rate of 1.4% per year since 1998, whereas the incidence is declining at a rate of -1.6% per year. The burden of joint pain defined as OA is quite high, constituting nearly one third of primary care adult patients.

People with OA are more likely to have multimorbidity prior to (aOR 1.71, 95%CI 1.69-1.74) and after the diagnosis of OA (aHR 1.29, 95%CI 1.28-1.30) than people without OA. Musculoskeletal (MSK), gastrointestinal (GI), cardiovascular (CV) and psychological conditions were associated with OA before and after the diagnosis of OA, whereas dementia and systemic lupus erythematous (SLE) were only associated with OA after its diagnosis. Other conditions that showed significant associations with OA both before and after diagnosis, were anaemia, inflammatory bowel disease (IBD), benign prostatic hypertrophy (BPH), gall stones, liver diseases, cancer, and hearing impairment.

Five multimorbidity clusters were identified in OA. These clusters were led by both pain and hypertension, hypertension only, depression, back pain only, and relative healthy group (lowest number of any conditions).

Over time, comorbidity clusters changed after the diagnosis of OA. About 30% of people changed from the cluster driven by either back pain or hypertension to the cluster driven by both back pain and hypertension. The accumulation of multimorbidity in people with OA happens in five different ways, and 17.5% of people develop multimorbidity quicker compared to relative healthy group. Obesity, smoking and alcohol use during the diagnosis of OA are strongly associated with the faster development of multimorbidity.

People with OA were 1.2 times more likely to consult with general practitioners (GP), 1.1 times more likely to be hospitalised, 3.25 times likely to get higher DALYs and 1.9 times more likely to die. Within OA, people with multimorbidity had higher mortality, burden, and health utilisations.

Conclusions

OA affects one in ten people aged 20 years or more in the UK. The burden of both GP diagnosed OA and joint pain in primary care is consistently high and increasing further. People with OA are more likely to develop other chronic conditions. Five different comorbidity clusters have been identified. While younger people are likely to have pain and depression, the elderly are likely to have CV-MSK comorbidities. The growth of multimorbidity in people with OA differs with 17.5% developing it faster than others. People with OA and CV-MSK and CV comorbidity have worse health outcomes. This information from this study can be used to develop personalised care in primary care. Further research is needed to understand the causality between OA and comorbidity.

List of publications

Publications based on PhD work (Research Articles)

Swain S., Sarmanova A., Coupland C., Doherty M. and Zhang W., Comorbidities in Osteoarthritis: A Systematic Review and Meta-Analysis of Observational Studies. Arthritis Care Res, (2020), 72: 991-1000. doi:10.1002/acr.24008

Swain S., Sarmanova A., Mallen C., Fu Kuo C., Coupland C., Doherty M. and Zhang W. Trends in incidence and prevalence of osteoarthritis in the United Kingdom: findings from the Clinical Practice Research Datalink (CPRD). Osteoarthritis Cartilage. 2020;28(6):792-801. doi:10.1016/j.joca.2020.03.004

Swain S., Coupland C., Mallen C., Fu Kuo C., Sarmanova A., Bierma-Zeinstra S.M.A, Englund M., Prieto-Alhambra D., Doherty M., and Zhang W.; Temporal relationship between osteoarthritis and comorbidities: a combined case control and cohort study in the United Kingdom primary care setting. (*Rheumatology*, keab067, https://doi.org/10.1093/rheumatology/keab067)

Swain S., Coupland C., Mallen C., Fu Kuo C., Sarmanova A., Bierma-Zeinstra S.M.A, Englund M., Prieto-Alhambra D., Doherty M., and Zhang W.; Clustering of comorbidities in people with osteoarthritis and associations with general practice consultations, hospitals admissions, and all-cause mortality rates: - a study from UK primary care. (Submitted to ARD)

Publications based on PhD work (Conference abstracts)

Swain S, Sarmanova A, Mallen C, Kuo C, Coupland C, Zhang W. Trends in Incidence and Prevalence of Osteoarthritis in the United Kingdom: Findings from the Clinical Practice Research Datalink (CPRD) [abstract]. *Arthritis Rheumatol.* 2019; 71 (suppl 10). American College of Rheumatology Congress 2019.

Swain S, Coupland C, Mallen C., Fu Kuo C., Sarmanova A, Doherty M., and Zhang W; Temporal relationship between osteoarthritis and comorbidities: a combined case control and cohort study in the United Kingdom primary care setting. OARSI, 2020 World Congress.

Swain S, Coupland C, Strauss V, Mallen, Fu Kuo C., Sarmanova A., Doherty M, Zhang W. Multimorbidity clusters, determinants, and trajectories in osteoarthritis in the UK: findings from the clinical practice research datalink. Annals of the Rheumatic Diseases 2020;79:49-50. EULAR 2020

Publications during PhD period (non-PhD work)

Zeng C, Zhang W, Doherty M, Persson MSM, Mallen C, **Swain S**, Li X, Wei J, Lei G, Zhang Y. Initial analgesic prescriptions for osteoarthritis in the United Kingdom, 2000-2016. Rheumatology (Oxford). 2020 Jun 28:keaa244. doi: 10.1093/rheumatology/keaa244.

Wang H, **Swain S**, Luo J, Blake H, Chattopadhyay K. Barriers, and facilitators to physical activity among ethnic Chinese children: a qualitative systematic review. JBI Evid Synth. 2020 Aug 21. doi: 10.11124/JBISRIR-D-19-00154.

Abdulrahim H, Jiao Q, **Swain S,** Sehat K, Sarmanova A, Muir K, Zhang W, Doherty M. Constitutional morphological features, and risk of hip osteoarthritis: a case control study using standard radiographs. (Minor revision from Annals of Rheumatic Diseases)

Acknowledgments

Firstly, I would like to express my sincere gratitude to all my supervisors, Professor Weiya Zhang, Professor Carol Coupland and Professor Michael Doherty. This journey so far is impossible without their support, guidance, and confidence in me throughout the period. Every possible opportunities, challenges, autonomies, feedback, and advice received from supervisors have enabled to change my life truly. I consider myself very lucky to develop and strengthen the research skills under their belief in me.

Further, I would like to thank other wonderful researchers and collaborators I worked with during this period. Heartiest thanks to Professor Christian Mallen for his thoughtful critical feedbacks on my work consistently. Dr Aliya Sarmanova, for being a friendly mentor always. Dr Georgina Nakafero, for never disappointing when asked for help with big data. Dr Victoria Strauss, for sharing the knowledge of complex methodologies and the hand-holding supports to perform them. Professor Chang-Fu Kuo, for helping with the interpretations. The School of Medicine and the Department of Academic Rheumatology, of University of Nottingham, UK for providing conducive and collaborative environment for the research.

Also, I would like to thank my friends, staffs, and colleagues in the Department of Academic Rheumatology for enriching my life with myriad experiences. The administrative supports, social gathering and those fun moments made the journey lighter and memorable. Special thanks to Rose-Farrands-Bentley for being an amazing supporter with a helping attitude and smile. Thank you Onosi Ifesemen for those wonderful discussions on complex statistical modelling and the help with software, and other PhD colleagues for all the moments spent supporting each other through listening to the presentations and sharing the feedbacks and creating a life beyond the PhD.

With the University, my distinct thanks to the University accommodation tutor team and the sports centres for the psychological supports, promoting health and encouraging the cross-disciplinary research.

It is worth mentioning, this journey would have been impossible without the funding support from the Vice-chancellor scholarship for International Students in Post Graduate Research by the University of Nottingham, UK and the international young researcher scholarship from the Beijing Joint Care Foundation, China.

Finally, my special thanks to my family for their support and motivation from thousands of miles away. Especially, my parents, sisters, brothers-in-law, and cousins, all your belief, encouragement, and pride in me during many stressful times motivated me enough to sail through the lengthy ‘away from home’ study period.

Table of contents

List of Figures

Page

List of Tables

Page

Appendix Tables

Appendix Figures

Appendix Methods

List of Abbreviations

AAPC	Average Annual Percentage Change
aBIC	Sample size adjusted BIC
ACL	Anterior Cruciate Ligament
AIC	Akaike Information Criteria
BIC	Bayesian Information Criteria
BMI	Body Mass Index
BNF	British National Formulary
BPH	Benign Prostatic Hypertrophy
CCI	Charlson Comorbidity Index
CI	Confidence Interval
CNS	Central Nervous System
COPD	Chronic Obstructive Pulmonary Disease
CPRD	Clinical Practice Research Datalink
CRD	Current Registration Date
CVD	Cardio-vascular Diseases
CV	Cardiovascular
DALY	Disability Adjusted Life Years
DM	Diabetes Mellitus
FDR	False Discovery Rate
ECI	Elixhauser Comorbidity Index
FHSA	Family Health Services Appeal Authority
GBD	Global Burden of Diseases
GI	Gastrointestinal
GP	General Practice
HALE	Healthy Life Expectancy at Birth
HES	Hospital Episode Statistic
HIV	Human Immune Virus
IBD	Inflammatory Bowel Disease
IBS	Irritable Bowel Syndrome
IL	Interleukins
ISAC	Independent Scientific Advisory Committee
JSN	Joint Space Narrowing
K-L	Kellgren & Lawrence
LCA	Latent Class Analysis
LCD	Last Collection Date
LLRT	Log-likelihood ratio test
LOD	Length of data
MH	Mental Health
MINAP	Myocardial Ischaemia National, Audit Project
MRI	Magnetic Resonance Imaging
MSK	Musculoskeletal
NGF	Nerve Growth Factors
NHANES	National Health and Nutrition Examination Survey
NHS	National Health Services

NICE	National Institute for Health and Care Excellence
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
OA	Osteoarthritis
ONS	Office for National Statistics
OXMIS	Oxford Medical Information system
PVD	Peripheral Vascular Diseases
QOF	Quality and Outcome Framework
RA	Rheumatoid Arthritis
RCT	Randomised Controlled Trial
SLE	Systemic Lupus Erythematous
TJR	Total Joint Replacement
UK	United Kingdom
UTS	Up-to Standard
VAMP	Value Added Medical Products
WHO	World Health Organisation

1 Chapter 1. Introduction

1.1 Osteoarthritis

1.1.1 Definition

Osteoarthritis (OA) is the most common arthritis and a major cause of chronic pain and disability in developing as well as developed countries (Vos *et al.*, 2012). Despite being the most common arthritis in the middle-aged and elderly population, the definition of the condition is still under active research (Zhang and Jordan, 2010; Kraus *et al.*, 2015).

According to the National Institute for Health and Care Excellence UK, (NICE, UK) "Osteoarthritis is characterized pathologically by localized loss of cartilage, remodelling of adjacent bone and associated inflammation. Osteoarthritis includes a slow but efficient repair process that often compensates for the initial trauma, resulting in a structurally altered but symptom-free joint (Osteoarthritis - NICE CKS, n.d.)."

Achieving consensus on globally accepted definitions of disease and standards for classifying OA would help in better understanding across both the clinical and research domains. Also, from public health perspectives, a uniform epidemiological definition is essential for better estimation and comparison of disease burden and associated risk factors, and for designing effective interventions (Martel-Pelletier *et al.*, 2016).

The site specificity of OA shows higher inclination toward certain synovial joints (Doherty, 2001). One of the interesting hypotheses to explain this is linked to human evolution. Joints that have undergone major changes in orientation and function to adapt to bipedal gait and altered fore-leg (arm) usage may not be fully adapted and are relatively under-designed, so more commonly fail to compensate for adverse mechanical stresses and present with clinical signs and symptoms.

1.1.2 Classification criteria

Diagnosis and classification of OA can be done in several ways, for example, according to pathological features, radiographic features, or physical signs and symptoms (Abhishek A and Doherty M, 2013). One common problem in finding a single definition for OA is the involvement of different joints such as hips, knees, hands, or foot joints.

In general, clinical, and radiographic OA diagnostic criteria are the most accepted and efficient in clinical settings. Radiographs are the most common method of classification because of widespread availability, low cost, good standardisation and reasonable reproducibility (Kinds *et al.*, 2011). The Kellgren and Lawrence (K-L) system was the first suggested, globally adopted method which measures OA through an ordinal scale (0 to 4), being used over the past half century. K-L depends on the presence of osteophytes, narrowing of joint space, subchondral sclerosis, cysts and deformity (Schiphof, Boers and Bierma-Zeinstra, 2008). Details of K-L are given in Table 1.1-1.

Table 1.1-1. K-L grading system for classification of OA

Scale	**Grade and characteristics**				
Kellgren-Lawrence(Kellgren and Lawrence, 1957)	0: No JSN or reactive changes	1: Doubtful JSN, possible osteophytic lipping	2: Definite osteophytes, possible JSN	3: Moderate osteophytes, definite JSN, some sclerosis, possible bone-end deformity	4: Large osteophytes, marked JSN, severe sclerosis, definite bone ends deformity

Reprinted with permission from Wright RW. Osteoarthritis classification scales: interobserver reliability and arthroscopic correlation. *J Bone Joint Surg Am.*2014;96:1145–1151; * JSN = joint space narrowing

1.1.3 Clinical diagnosis

OA can be diagnosed alternatively using clinical signs and symptoms.

The commonest algorithm for clinical diagnosis is the American College of Rheumatology (ACR) criteria (Altman *et al.*, 1986), specifically:

- Aged 40 years or more
- Crepitus or bony swelling
- Pain most of the days of the month

- Morning stiffness of less than half an hour
- Plain radiographic changes
- Negative Rheumatoid Factor

Not always, the presence of radiographic changes in joint are manifested as symptomatic (Bedson and Croft, 2008; Neogi *et al.*, 2009). The sign and symptoms reported by a person is influenced by individual factors such as disease status, presence of other chronic conditions, perceived severity, socio-economic factors and altered pain physiology (Wise *et al.*, 2010; Luong *et al.*, 2012; Neogi, 2013). So, identifying people with symptomatic OA is more useful than radiographic screening only, where the latter categorises a non-symptomatic individual with OA and is relatively insensitive at showing milder early OA.

1.2 Epidemiology of Osteoarthritis

1.2.1 Prevalence of Osteoarthritis

OA can develop in any synovial joint, but the most frequently affected joints are knees, hips, hands, spinal facet joints and feet. Prevalence and incidence data largely vary according to the site, study area, definition of OA (symptomatic versus radiographic), age group and gender (Abhishek A and Doherty M, 2013). Table 1.2-1 describes the prevalence of OA reported in different studies. In the year 2005, 26 million people in the United States had OA.(Lawrence *et al.*, 2008) According to Versus Arthritis one third of people in the United Kingdom aged 45 years and over have sought treatment for OA (Versus Arthritis, 2019). In total, 8.75 million people in the UK have visited any health facility for treatment and by 2035, 8.3 million people in the UK aged 45 years or over could have knee OA ("Osteoarthritis in General Practice; Data and perspectives," 2013). (Table 1.2-1) Studies from different countries report the overall prevalence of OA among those aged 45 years to vary between 20% to 35%. The Framingham community cohort study reported a higher prevalence of radiographic hip OA among men (Kim *et al.*, 2014). In Sweden, among people aged 56–84 years the prevalence of radiographic and

symptomatic knee OA were 25.4% and 15.4%, respectively (Turkiewicz *et al.*, 2015). In England, among 26,000 adults aged 50 years or more, 50% reported having OA in at least one of four joints (hand, hip, foot, knee) (Thomas, Peat and Croft, 2014).

Table 1.2-1 Prevalence of OA from different studies

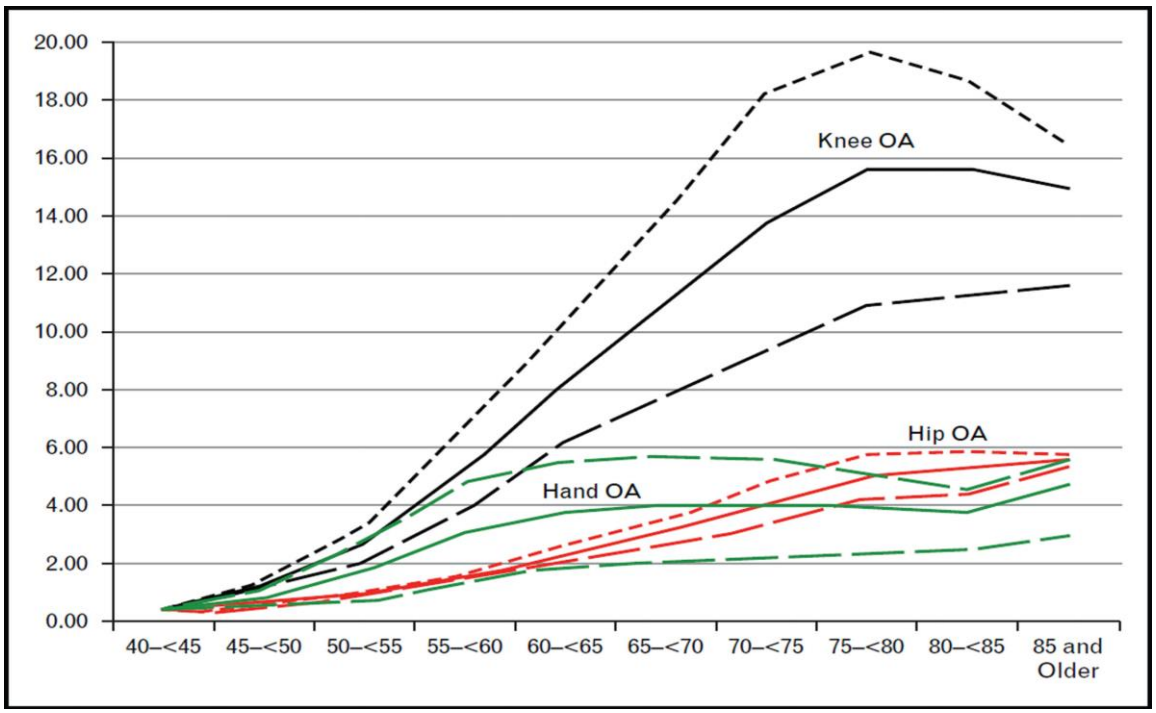
		Sample characteristics					Prevalence (%)			
Study	Country	Sex	Age (year)	Source of database	Method of diagnosis	Sample size	Overall	Hip	Knee	Hand
(Plotnikoff *et al.*, 2015)	Canada	M&F	≥18	Community survey	Self-reported	4733	14.8	10.5	8.5	
Framingham (Kim *et al.*, 2014)	USA	M&F	≥45	Cohort database	Radiographic	1424	19.2		33	M-13.2 W-26.2
Johnston Country (J. M. Jordan *et al.*, 2007)	USA	M&F	≥45		Radiographic and symptomatic	3018	27.8	36	43	
NHANES III (Dillon *et al.*, 2006)	USA	M&F	>60	Community survey	Radiographic	6913			37	8
WHO Study on global AGEing and adult health (Brennan-Olsen *et al.*, 2017)	China, Ghana, India, México, Russia, South África,	M&F	18-49	Community survey	Self-reported	44747	17.8			
National Health and Wellness Survey (NHWS) (Kingsbury *et al.*, 2014)	UK, France, Germany, Spain, and Italy	M&F	≥65	Community	Self-reported	3750		30.1	54.7	34.7
National Health Survey, 2014-15(Statistics, 2015)	Australia	M&F	≥18	Community	Self-reported	19000	20.4			
Korean NANHES (Lee and Kim, 2017)	South Korea	M&F	≥50		Radiographic	9512	34.10			
Zoetermeer (van Saase *et al.*, 1989)	Netherlands	M&F	≥19	Community	Radiographic	6585				27-80

M- Men; W-Women

1.2.2 Incidence of Osteoarthritis

Few studies have used large databases to describe OA incidence. The crude incidence rate of OA in Canadian adults was 14.6 per 1000 person-years in 2000/2001(Rahman *et al.*, 2014). Primary care records of more than 3 million patients in Spain reported the incidence rates (per 1000 person-years) for knee, hip, and hand osteoarthritis as 6.5, 2.1, and 2.4, respectively (Prieto-Alhambra *et al.*, 2014).(Figure 1.2-1) Ascertaining OA through administrative records has limitations, still it supports the potential of large databases for estimation of population burden and trends.

Figure 1.2-1 Incidence of joint specific OA



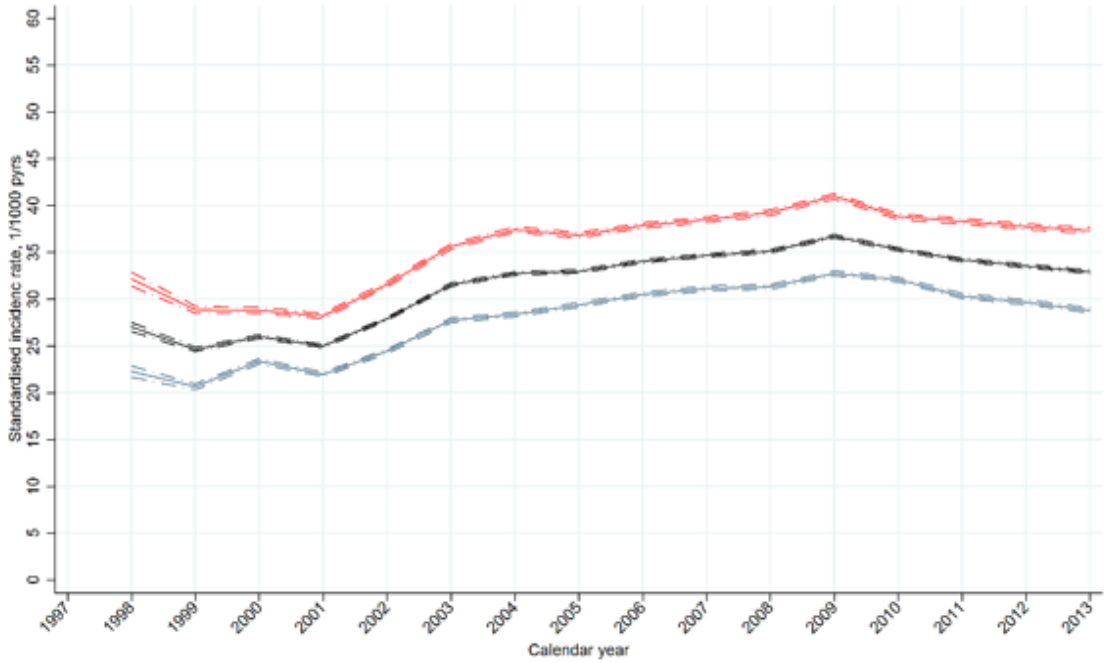
Age and sex-specific incidence rates (/1000 person-years) of knee osteoarthritis (black), hip osteoarthritis (red), and hand osteoarthritis (green). Solid, all population; short dash line, women; long dash line, men. (Prieto-Alhambra *et al.*, 2014) Produced with permission (copyright number 4897681007765)

In the UK during the year 2013, the age-sex standardized incidence rate of clinical OA (symptomatic) was 40.5 per 1000 person-years and was higher in women than men. Joint specific rates were higher for the knee (19.7) followed by the hip (10.4) and the hand (4.3) with progressive increases with age (Yu *et al.*, 2017).

1.2.3 Trends in incidence of OA

Very few countries have explored the trends in incidence of OA. According to a UK CPRD data analysis, the trends for any clinical OA are increasing gradually over the last 15 years (1998-2013), as illustrated in Figure 1.2-2. In Sweden, age-standardized hospitalization rates due to OA have increased from 1998 to 2014 for the hip and knee (Kiadaliri *et al.*, 2018). In Canada, during 2000-01 to 2008-09, crude incidence rates changed from 11.8 to 14.2 per 1000 person-years for men, and from 15.7 to 18.5 per 1000 person-years for women (Rahman *et al.*, 2014). The increase in crude rates per annum was about 2.5-3.3% for both men and women (Rahman *et al.*, 2014).

Figure 1.2-2 Trends of clinical OA incidence in UK



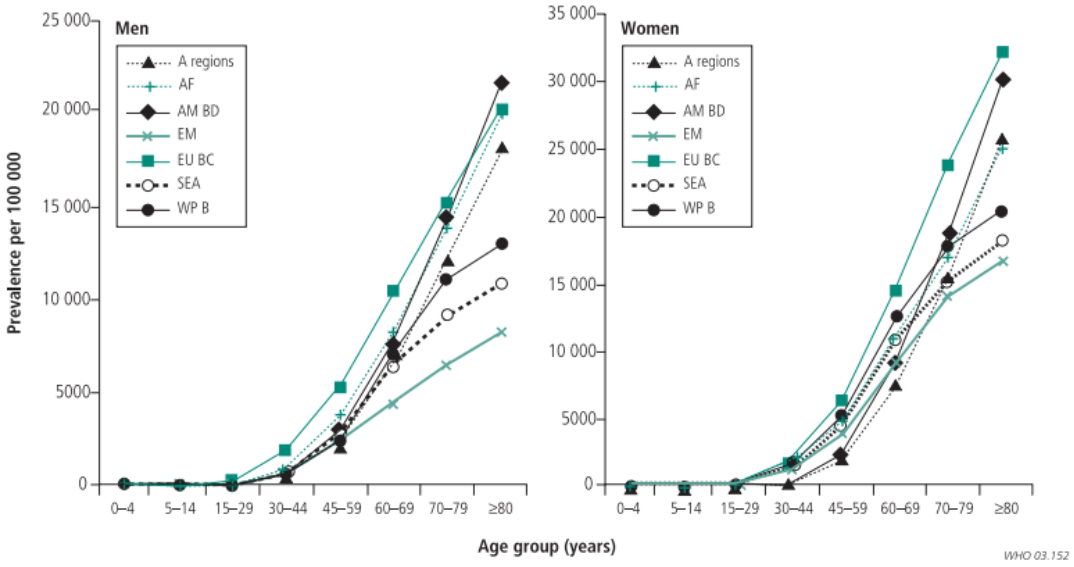
Red: women; Blue: men; Black: men and women combined. Solid lines represent incidence estimation; dash dot line represents the 95% confidence interval. (Yu *et al.*, 2017) Produced with permission (copyright number 4897681161059)

1.3 Burden of Osteoarthritis

1.3.1 Disability Adjusted Life Years (DALYs) in OA

According to one estimate, 27.6% of the older population (>60 years) in the world have OA (Symmons, Mathers and Pfleger, 2003) (Figure 1.3-1). Global burden of disease (GBD 2010) ranked hip and knee OA as the 11th highest contributor to global disability and the 38th highest in disability adjusted life years (DALYs). Increasing life expectancy and ageing populations are expected to make OA the fourth leading cause of disability by the year 2020 (Woolf and Pfleger, 2003). Years of life with disability (YLDs) for hip and knee OA increased by 6.6 million over the period 1990-2010 (10.5 million in 1990 to 17.1 million in 2010). The United Nations estimates that by 2050 more than 20% of the world's population will be aged 60 years or more ('World Population Prospects: The 2017 Revision', 2017). By 2050, it is estimated that 130 million people worldwide will suffer from OA, 40 million of whom will be severely disabled by the condition (WHO Scientific Group on the Burden of Musculoskeletal Conditions at the Start of the New Millennium, 2003).

Figure 1.3-1 Prevalence of OA in the world



A region= Developed countries in North America, Western Europe, Japan, Australia and New Zealand, AF= Countries in sub-Sharan Africa, AM BD- Developing countries in Americas, EM- Eastern Mediterranean and North African regions, EU BC- Developing Countries in Europe, SEA- South East Asia, WP B- Western Pacific region, (Symmons, Mathers and Pfleger, 2003). Produced with permission (copyright number 4897680659617)

1.3.2 Economic burden of OA

As for economic considerations, the direct costs for topical and oral NSAIDs used for management of OA in UK were estimated at £19.2 million and £25.65 million, respectively (Chen *et al.*, 2012). Higher costs were incurred for hip and knee replacement surgery (costing more than £850 million) and arthroscopic surgery for OA was about £1.34 million. Similarly, the loss of economic production (indirect cost) was over £3.2 billion. OA results in increasing economic burden for all countries, both from direct and indirect costs.

1.4 Risk factors for Osteoarthritis

OA is a common complex disorder with multiple genetic, constitutional, and environmental risk factors.

1.4.1 Individual level risk factors

1.4.1.1 Ageing

Studies have documented that OA increases with age (Figure 1.2-1). OA is uncommon in people aged less than 40 years , although recent studies have started to document knee pain symptoms and/or symptomatic OA in quite a number of younger adults (Plotnikoff *et al.*, 2015). Age influences risk of OA differently according to joint site. For example, the Fallon Community Health Plan reports higher incidence of hand OA in both men and women with increasing age (Oliveria *et al.*, 1995) whereas, in the Spanish population higher incidence of knee and hip clinical OA with age was reported compared to hand OA (Prieto-Alhambra *et al.*, 2014).

1.4.1.2 Gender

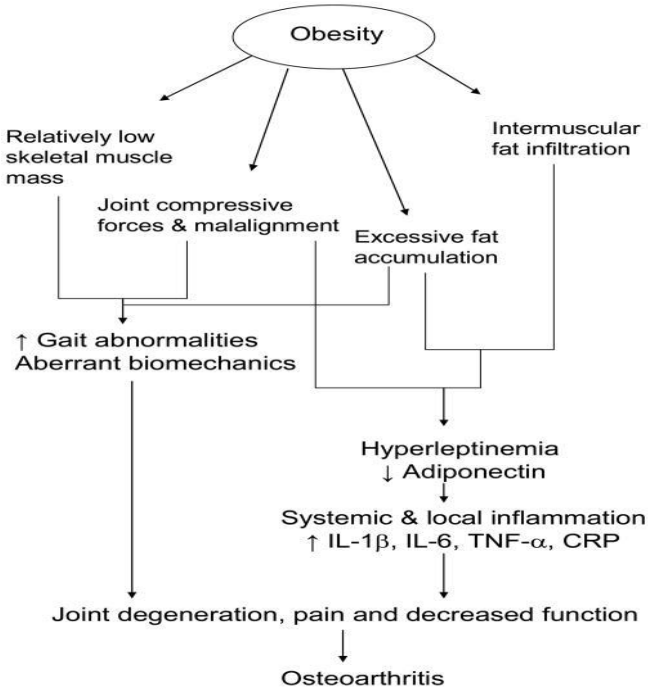
Gender has a strong relationship with OA mostly towards the latter half of life. Prevalence of OA is higher among women (Felson, 2000). Not only OA diagnosis, also the severity and involvement of multiple joints become more common after the age of 50 years. In women the interphalangeal joints, first carpometacarpal joints and knees are most

commonly affected, whereas, men are more likely to have OA at metacarpophalangeal joints and hips (Moskowitz, 2007). There are theoretical explanations to describe the role of oestrogen, and the reduced levels following the menopause, in influencing and mediating the risk factors for OA (Spector and Campion, 1989).

1.4.1.3 Obesity and metabolic syndrome

Worldwide the prevalence of obesity has nearly tripled since 1975 (WHO report, 2019). In 2016, more than 1.9 billion adults aged 18 years and more, were overweight, of which 650 million were obese(WHO report, 2019). A strong association of obesity with OA has been established and reported in various systematic reviews (Zhou *et al.*, 2014; Zheng and Chen, 2015). However, most of the studies reported associations with knee OA and hip OA. This association is explained through higher/altered loading of weight-bearing joints and possible systemic low-grade inflammation (Runhaar *et al.*, 2011). Higher BMI also associates with OA of non-weight bearing joints, such as in the hand, suggesting a role for biomechanical, metabolic, or inflammatory aspects of obesity.

Figure 1.4-1 Potential obesity related pathways that contribute to OA



Produced with permission (copyright number 4897690041039). (Vincent *et al.*, 2012)

The exact metabolic relationship between obesity and OA is not clear. It is thought that adipokines released into the systemic circulation from excess fat cells deteriorates the health of chondrocytes and other joint tissues through degradative enzymes (Gómez *et al.*, 2011). Local intra-capsular fat (e.g. retro-patellar fat in the knee) has also been implicated in releasing local cytokines that may damage cartilage (Simopoulou *et al.*, 2007). (Figure 1.4-1)

1.4.1.4 Nutritional factors

Amongst, nutritional factors, vitamin D has been studied most extensively. Lower intake of vitamin D and increased risk of OA has been suggested because of impairment in cartilage metabolism (Garfinkel, Dilisio and Agrawal, 2017). However, the findings remain conflicting for vitamin D and other nutrients/vitamins (E, K and C). An inconsistent protective effect of vitamin C and E with osteoarthritis risk has been suggested (McAlindon *et al.*, 1996; Wluka *et al.*, 2002). Similarly, low vitamin K has been reported in people with OA aged 50 years or more (Neogi *et al.*, 2006).

1.4.1.5 Bone density and bone mass

High bone density has been recognised as a risk factor for incident OA (Nevitt *et al.*, 2010), although a causal relationship and mechanism remain unclear. High bone mass is linked with subchondral bone sclerosis rather than with joint space narrowing. This might be an indication of a hypertrophic OA phenotype.

1.4.1.6 Smoking

Smoking, a major risk factor for cardio-vascular disease, does not have a significant association with OA (Hui, Doherty and Zhang, 2011). Another systematic review by the same group found similar findings on smoking having no significant role in incidence and progression of OA (Pearce *et al.*, 2013).

1.4.1.7 Low birth weight

Low birth weight and pre-term birth have been associated with higher risk of hip arthroplasty as an adult (Hussain *et al.*, 2018). Even hip osteophytes in later life is linked with lower birth weight (Clynes *et al.*, 2014).

1.4.1.8 Sports and physical activity

Individuals engaged in heavy activities or high contact sports are at increased risk of developing OA. Reviews have reported conflicting evidence on physical activities and OA, mostly because of non-uniformity of measuring physical activities (Tran *et al.*, 2016; Timmins *et al.*, 2017). It is not merely the activities, but joint loading, injury, physical strength, and other factors that determine the risk of OA. Moderate levels of physical activity do not cause OA through improving the muscle balance and controlling overweight and obesity. Overall a moderate level of physical activity is necessary to maintain joint and overall health.

1.4.1.9 Occupational factors

There is a strong evidence of association of OA with occupation. McWilliams et al reported in a systematic review the risk of knee OA was 1.6 times high with occupational activities. The weight bearing joints of lower limbs are exposed to excessive load, stress, trauma due to different activities in occupation such as weightlifting, prolong standing and lift walking (Klussmann *et al.*, 2010; Schram *et al.*, 2020). A genetic study among co-twins supported the mentioned occupational activities relation with knee OA (Skousgaard *et al.*, 2018). Not only the weight bearing joints, but small joints in hand have reported to develop OA due to occupation exposure (Fontana *et al.*, 2007).

1.4.1.10 Genetic

OA has a genetic component which may vary by joint site. More than 20 gene polymorphisms are known to be associated with OA (Bravatà *et al.*, 2015; Ren *et al.*,

2017). According to Twin and family studies, the heritable component of OA varies between 50 and 65%. It has greater genetic influences for hand and hip OA compared to knee OA (Spector *et al.*, 1996). However, having familial clustering of OA could be because of sharing similar living environment and lifestyles, which needs to be excluded. In addition, differences in the pattern of OA also exist, such as hip OA being more common in the Western countries, whereas knee OA is more common in the Asian and Chinese population (Allen, 2010).

1.4.2 Joint-Level Risk Factors

1.4.2.1 Bone/joint shape

Constitutional joint shape and risk of OA is gaining interest among researchers. A significant difference in femoral head shape and incident hip OA was reported among men in the Johnston County osteoarthritis project (J. M. Jordan *et al.*, 2007). The Nottingham musculoskeletal research group also found the pistol-grip deformity of the hip is associated with hip OA (Doherty *et al.*, 2008). Neogi et al predicted knee OA using joint shape (full joint) rather than individual bones (Neogi *et al.*, 2009). Bone/joint shapes alter the biomechanics of the joint which then predispose to OA.

1.4.2.2 Injury

Joint injury includes meniscal damage, ligament rupture, or direct articular cartilage injury. Patients with anterior cruciate ligament (ACL) deficiency and reconstructed knees had altered synovial fluid biomarker levels indicative of OA (Riccardo, Fabio and Pietro, 2017). Apart from overt injuries, repetitive micro-trauma may also compromise joint tissues.

1.4.2.3 Muscle strength and mass

Muscle weakness predisposes to an increased risk of knee OA (Dell'isola *et al.*, 2018). Especially, in weight bearing joints, stronger muscles afford stability and protect the joint from undue loading, minimising the trauma to cartilage and joint tissues. However, the role

of muscle strength and mass in OA development and progression is still somewhat unclear for other joint sites.

1.4.2.4 Other joint-level risk factors

Limb length can be another joint level factor associated with OA. Unequal limb length causes more loading in the longer leg and increases the risk of knee OA. Similarly, bony malalignment such as varus or valgus changes the biomechanics of joint loading and predisposes to knee OA.

1.5 Clinical manifestation of Osteoarthritis

People with OA may present with pain, stiffness, reduced function, and participation restriction which all may cause reduced quality of life.

1.5.1 Pain

Pain reported in OA is multidimensional and can be influenced by both peripheral (local joint tissue) factors and central (nervous system) factors. Pain is typically usage-related, relieved by rest, and worsens towards the end of the day (Hawker *et al.*, 2008). Pain in OA cannot originate from cartilage as it is not innervated, but there are nociceptors in the joint synovium, capsule, subchondral bone, and periosteum. Stimulation of peripheral nociceptor sensory nerves can be via soluble inflammatory mediators such as cytokines and prostaglandins and by biomechanical trauma, and the ascending signals are frequently modified in OA at the spinal cord and brain level (Lluch *et al.*, 2014). Central sensitization is described as sensitization of nociceptive stimuli by active neurons which becomes hyperresponsive to subsequent stimuli to the neuron's receptor fields (O'Neill and Felson, 2018). So, even in the presence of low intensity nociceptive stimuli from the OA joint, higher perceived pain can be because of sensitization of the pain centres from chronic input and from other sources of pain. This suggests the possible presence of other

painful conditions such as; fibromyalgia or chronic fatigue syndrome in an individual, which alters the pain sensitivity (Woolf and Salter, 2000).

1.5.2 Stiffness

In contrast to inflammatory arthritis, morning stiffness on arising is a minor feature of OA and is generally short-lived (less than 30 minutes) and rapidly wears off with movement. Similarly, stiffness after inactivity is short-lived and minor.

1.5.3 Functional impairment

Both pain and reduced range of joint movement result in impairment of joint function. Such impairments can cause disability with difficulty undertaking activities of daily living, and restriction in participation in social and work activities. Disability and participation restriction often have a negative impact on quality of life.

1.6 Pathology of OA

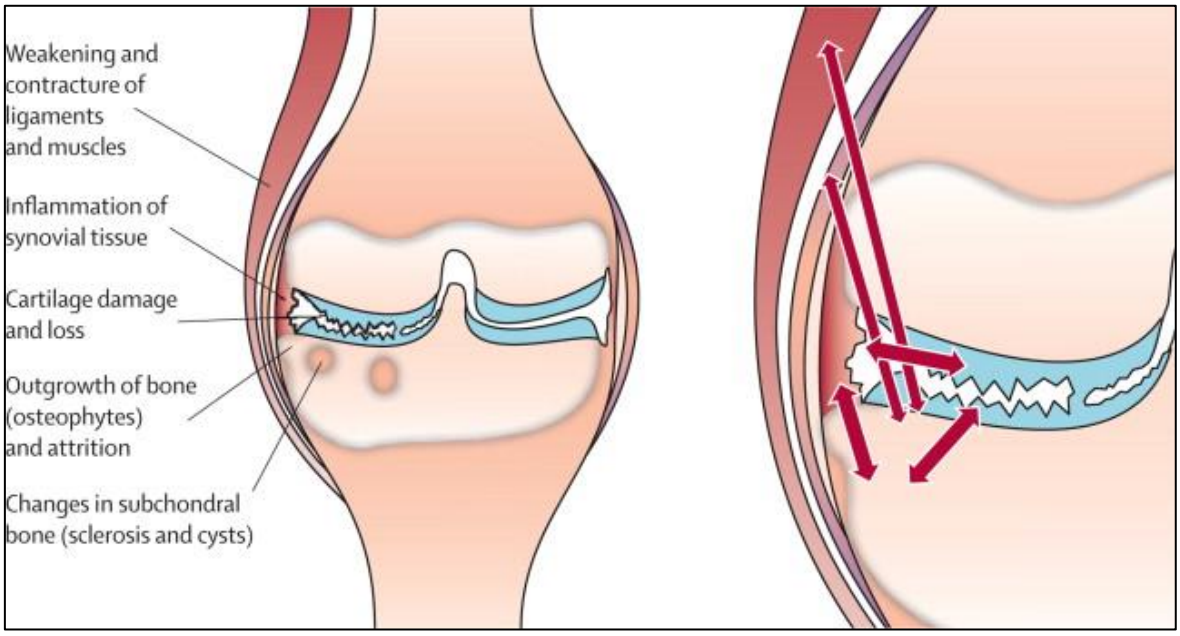
OA has a long history parallel to the evolution of man. As evidenced by osteoarthritic changes seen in a Comanchean dinosaur fossil, it appears to have remained pathologically unchanged for 100 million years (Dequeker and Luyten, 2008). The changes in osteoarthritic joints are relatively well recognised, although the causes for these changes are not. A variety of hypotheses aimed at explaining the changes have been voiced, including chronic mechanical overloading, matrix proteolysis, pro-inflammatory cytokine production, activation of cellular inflammatory signalling pathways, premature ageing of chondrocytes and cartilage matrix, and damage to the chondrocyte's deoxyribonucleic acid (DNA) (Hochberg *et al.*, 2015). The most robust hypothesis is the role of adverse biomechanical factors in the pathogenesis of OA (Hochberg *et al.*, 2015), however, it is by no means likely to be the only driving factor for OA. Most likely, a large

variety of factors contribute to the pathogenesis of OA and are responsible for the pathological changes witnessed.

1.6.1 Pathophysiology

The understanding towards the pathophysiology of OA is evolving. Though it was previously thought to be a consequence of normal ageing and damage due to mechanical factors, this theory is now thought to be inaccurate. Although much remains to be known regarding the causes of OA, different reasons apart from the 'wear and tear' have been postulated (Doherty *et al.*, 2016). OA causes biochemical changes to the cartilage, bone, and synovium altogether (Hochberg *et al.*, 2015). The metabolic dynamic process of OA involves both attrition and synthesis of tissues responsible for pathogenesis (Doherty *et al.*, 2016). Figure 1.6-1 shows the possible changes in joints during OA.

Figure 1.6-1 Changes in joints in OA



Different tissues like cartilage, synovial tissue, and subchondral bone, osteophytes involved in the clinical and structural changes of the disease shown on left. right hand side shows the bi-directional interplay between cartilage, bone and synovial tissue involved in OA. (Bijlsma et al., 2011). Produced with permission (copyright number 4897680659617)

1.6.1.1 Non-vascular pathology

The production of new tissue in synovial joints with OA with increased metabolic activity at many joint sites are evident. These suggest a potentially regenerative process, that is a repair/modelling phase after the 'wear and tear'. In many cases this slow repair process may be successful, thus leaving a structurally abnormal joint without pain. This may explain the evolutionary advantage and preservation of OA in species with synovial joints. However, with overwhelming insult and/or a poor repair process, the joint may continue to remodel to try to keep pace with the insults and progress towards joint failure with associated symptoms and functional impairment.

Normal articular hyaline cartilage forms a smooth covering for the joint surfaces and is responsible for the biomechanical properties of joints (Hochberg *et al.*, 2015). Articular cartilage undergoes a variety of changes in OA. Initially, disruption to the type II collagen scaffolding causes the cartilage volume to increase as the water content increases and proteoglycan swelling occurs (Hochberg *et al.*, 2015). This is later followed by cartilage loss secondary to the formation of superficial cracks (fibrillation) and continued protein-degrading activity (Hochberg *et al.*, 2015). Deep to these changes in the non-calcified cartilage, the calcified zone becomes thicker. In addition, blood vessels and nerves break through the tidemark separating the calcified and non-calcified cartilage, leading to neovascularisation and neo-innervation of the cartilage (Hochberg *et al.*, 2015). As articular cartilage is aneural, these changes do not produce any pain until innervated tissue gets involved (Bijlsma et al., 2011).

Deeper within the joint, at the junction between cartilage and bone, periosteal and synovial mesenchymal stem cells are induced to proliferate and form fibrocartilage which then undergoes endochondral ossification to form bone (Doherty *et al.*, 2016). The bony spurs appear on the marginal aspects of joints and are known as osteophytes. Some changes are seen directly within the subchondral bone in OA, although it's cause is

unclear (Hochberg *et al.*, 2015). Subchondral sclerosis, cysts, and bone marrow lesions (BMLs) are the morphological changes in subchondral bone (Hochberg *et al.*, 2015).

Recent evidence suggests low-grade inflammation of the synovial membrane (synovitis) in OA (Doherty *et al.*, 2016). Cartilage debris is released from the damaged cartilage surface, causing pathological synovial changes and the release of cytokines, growth factors, and enzymes that can further disrupt articular cartilage homeostasis (Hochberg *et al.*, 2015). Finally, the pathological changes in synovium, cartilage, and bone are often accompanied by changes in the menisci, joint capsule, intra-articular ligaments, extra-articular connective tissues, and peri-articular muscle (Doherty *et al.*, 2016).

The role of inflammation in the pathogenesis of OA is widely debated. The presence of inflammation in OA is generally accepted, as evidenced by the presence of synovitis, effusion, and stiffness (Robinson *et al.*, 2016). Changes indicative of inflammation, such as effusion, synovial hypertrophy , and Power Doppler signals, can be visualised using ultrasound examination. Ultrasound markers of inflammation are raised in those with knee pain compared to those without knee pain (Sarmanova *et al.*, 2017). In addition, the intensity of ultrasound inflammation increases with worsening structural changes (Sarmanova *et al*., 2017). Others believe inflammation to be a primary driver of OA (Berenbaum, 2013; Robinson *et al.*, 2016). Low-grade local and systemic inflammation are thought to cause OA through a multitude of mechanisms, including innate and adaptive immune mechanisms and inflammatory mediators (Robinson *et al*., 2016). Post-traumatic OA is thought to cause OA through local inflammation which causes synovitis and activates mechanoreceptors (Berenbaum *et al.*, 2013). Similarly, OA is thought to be driven by low-grade systemic inflammation in metabolic syndrome, by a secretory inflammatory phenotype in increased age, and by innate immunity in crystal OA (Berenbaum *et al*., 2013).

The vicious cycle of mechanical factors leading to OA has been recently emphasized (Felson, 2013). Abnormal mechanical loading may result from various insults such as malalignment, congenital dysplasia, meniscal tears, or chronic excessive loading (Felson, 2013). The affected cartilage is damaged, and the underlying bone may also undergo remodelling, further driving the abnormal loading (Felson, 2013). Consequently, cartilage debris may cause a secondary inflammatory response in the synovium, evidenced by synovitis and excess fluid secretion (Felson, 2013). The basis for the hypothesis is grounded primarily in studies that have found associations between risk factors that cause abnormal mechanical forces in the joint, such as obesity, injury, and occupational overuse, and the development of OA (Felson, 2013).

1.6.1.2 Vascular pathology

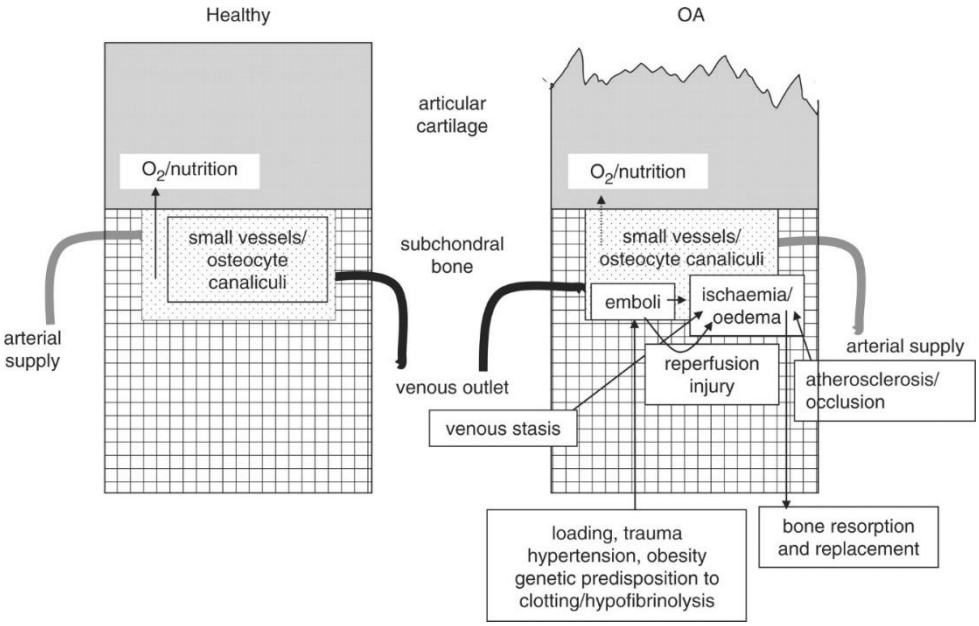
In the context of increasing reporting of metabolic syndromes and cardiovascular diseases in OA, vascular pathology has gained recent attention. Even though cartilage itself is avascular, the vascular pathologies in subchondral regions drive the possible hypothesis (Imhof *et al.*, 2000). Bone remodelling is accomplished by the co-ordinated action of osteoclasts and osteoblasts. The initiating event for sites of bone remodelling is not known, but it seems likely that these sites are targeted and the basis of this targeting may be regions of loss of osteocyte viability (Noble, 2003).

These osteocytes are prone to apoptosis due to reduced blood flow around them, triggering the osteoclast activities and excavation of non-viable bones. Bone marrow oedema, a consequence of bone trauma is known to be responsible for reduced blood flow (Mandalia *et al.*, 2005). It is also found to be associated with structural deterioration in knee OA (Hunter *et al.*, 2013). The reasons for bone marrow oedema are not well understood, but probably it is related to local trauma and injury.

Another vascular mechanism for OA is subchondral bone ischaemia, which leads to reduced nutrients and oxygen supply to the cartilage. The increased turnover of the

subchondral bone in OA could be secondary to episodic ischaemia, in turn due to vascular pathology in the subchondral bone. A detailed mechanism of subchondral bone vascularisation role is provided in Figure 1.6-2.

Figure 1.6-2 Subchondral vascular pathology of OA



Role of the subchondral vasculature in the initiation and/or progression of OA.

The left panel shows a representation of healthy articular cartilage overlying the subchondral trabecular bone. In addition to structural support and absorption of shock offered by the subchondral bone, its small vessels and probably the interstitial bone fluid in osteocyte canaliculi, provide important nutrition to the cartilage. The right panel shows some cartilage erosion, as seen in OA. Typically, the subchondral bone would also be altered in OA, with areas of bone marrow oedema (BMO), increased bone turnover, and sclerosis. BMO may be due to episodes of ischaemia, perhaps due to occlusion of the supply vessels by atherosclerosis, or venous stasis due to loading and/or increased intra-articular pressure, or to embolus formation in the small vessels of the subchondral bone. The latter could be due to trauma, obesity, or increased propensity to clot, perhaps exacerbated by reperfusion injury, at sites where blood flow has been lost and then recovers into hypoxic tissues. One result of local ischaemia in the bone may be to deny the overlying cartilage of nutrition, causing catabolic and reparative events in the cartilage. Osteocyte death in areas of bone affected by hypoxia will also be targeted for resorption and replacement, increasing subchondral bone turnover. The support to the articular cartilage and the shock absorption provided by the subchondral bone may be compromised during episodes of bone repair, leading to articular cartilage damage. (Findlay, 2007) Produced with permission (copyright no 4897680137150).

It is not only the inflow of blood, but also obstruction to the outflow that may decrease the cellular nutrient and oxygen supply (Wang C. -C *et al.*, 2011). Episodes of venous stasis in OA may lead to loss of osteocyte viability in regions of the bone. This is likely to occur especially in the highly vascular subchondral region of long bones. In vascular conditions such as hypertension, there is evidence of impaired capacity of vascular growth and

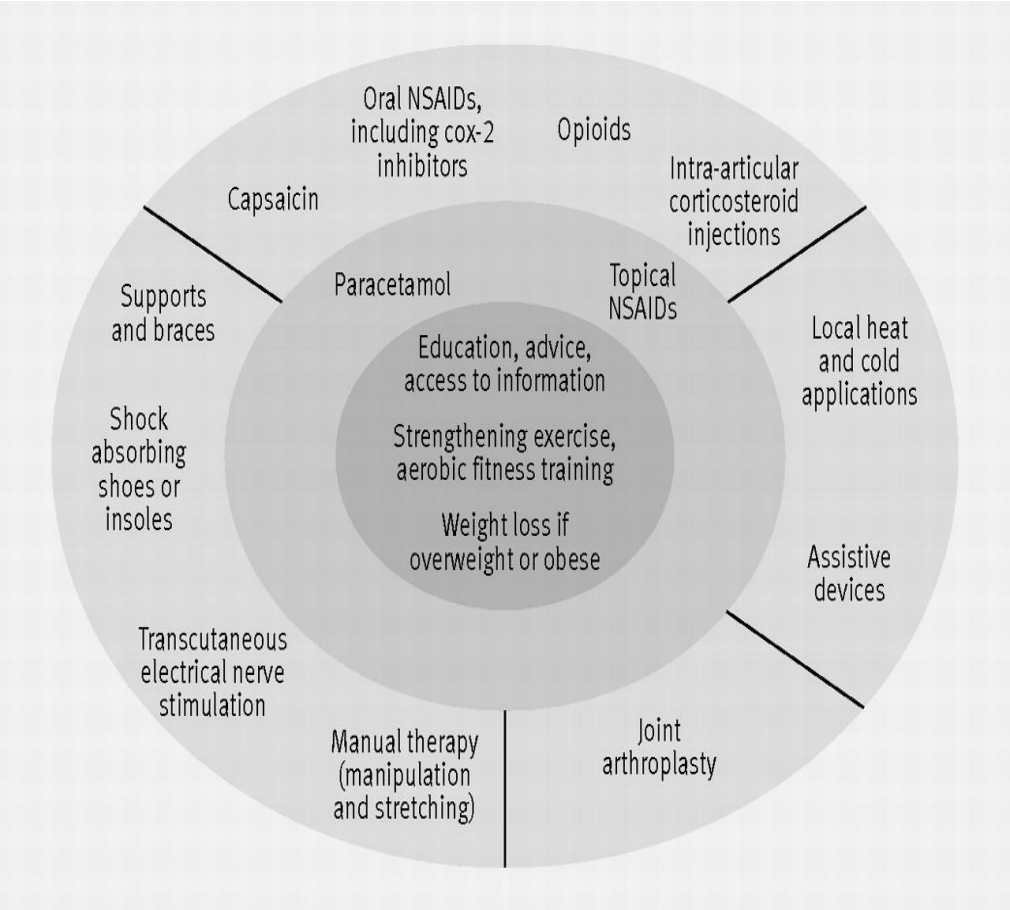
angiogenesis, which in turn may lead to endothelial cell damage or dysfunction. The factors responsible for such changes in hypertension are seen in OA too, though this is in a preliminary research stage (Enomoto *et al.*, 2003). The reduced microcirculation in diabetes and other peripheral vascular diseases might be explained also under the above mechanisms.

Listing the pathophysiological changes and the factors responsible is not complete without the multifactorial explanation. There is complex interaction between adverse mechanics, inflammation, and the systemic vascular changes, as potentially other factors. For large joints, the adverse mechanical loading of the joint may be the preliminary and primary driver of OA, proceeding to secondary inflammation. However, in smaller joints with greater inflammatory features, such as in erosive hand OA, there may be more extreme inflammation that may be an important driver of tissue damage. Also, the commonality of the vascular changing factors seen at subchondral regions could make the pathogenesis process proceed faster.

1.7 Management of Osteoarthritis

Management of OA has always been challenging because of the diversity of joints involved and symptoms reported. An individual's pain severity and illness perception towards the condition varies because of individual and health system factors. Thus, assessment and management of people with OA is recommended to be individualised 'person-care' rather than just 'disease-specific'. There is no single specific "cure" for symptomatic OA and the management aims to improve symptoms and reduce further joint insult in order to halt or retard the progression of OA (Anandacoomarasamy and March, 2010). Major guidelines support a package of care that comprises core non-pharmacological and adjunctive pharmacological modalities with an emphasis on individualised management and patient engagement (Doherty and Dougados, 2001; Zhang *et al.*, 2008). (Figure 1.7-1)

Figure 1.7-1 NICE guideline for management of OA (Conaghan *et al.*, 2008)



The core and options approach emphasised by NICE recommended treatments. Treatment option that are to be considered for everyone are in the centre, 'First line' analgesic to try in the second ring and other interventions are in the outer ring. From NATIONAL INSTITUTE OF CARE AND EXCELLENCE (NICE) GUIDELINES (COPYRIGHT NUMBER 4897690227029)

1.7.1 Core treatment

The core treatments of the NICE guideline include, education and information, exercise, and reduction of adverse mechanical factors (e.g. weight loss if overweight or obese) (NICE, 2014).

It is a primary responsibility of the healthcare professionals to provide tailored information and education to the patients about the nature of condition, the causes, prognosis, diagnostic options, and the available treatments, including their possible advantages and disadvantages. Enquiry should be made about the individual's illness perceptions of OA and incorrect perspectives should be discussed and changed. All information should be given in terms that can be understood by the individual patient. Such education underpins

the success of any treatment especially for patients with hip, knee, or multiple joint OA (Mazzuca et al., 1997).

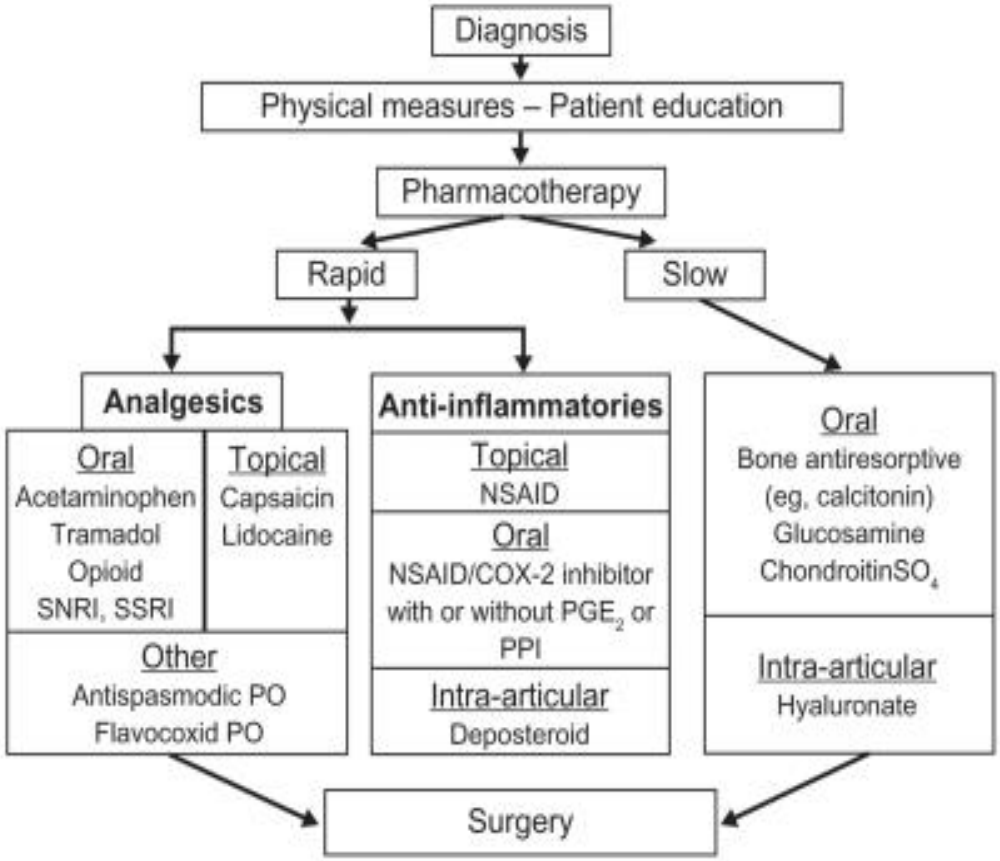
Continuing physical activity in people with OA, whilst pacing activities and avoiding undue mechanical stress, is important to maintain joint health. Although many perceive activity as a damaging factor for the joints, lack of activity is deleterious and regular appropriate activity and exercise (both local strengthening exercises, and aerobic exercise) is essential for effective management of OA(Doherty *et al.*, 2016). It is recommended that, people with OA should continue with neuro-muscular training, strengthening, and aerobic exercise within optimal limits to improve general fitness, muscle strength and maintain joint range of motion (Hunter and Eckstein, 2009). It is evident that both local strengthening and aerobic exercises reduce pain in OA (Ettinger *et al.*, 1997). Also physiotherapy and exercise used as therapeutic purpose is more effective in OA management (Fransen *et al.*, 2015).

Another important core treatment option is reduction of modifiable mechanical risk factors such as obesity. Reduction and maintenance of weight in overweight and obese people has been shown to significantly improve function and prevent OA (Miller *et al.*, 2006; Schlenk *et al.*, 2011) and is strongly recommended. Other mechanisms to reduce joint loading are the use of walking aids, splints, environmental modification (e.g. raised toilet seats, walk-in showers instead of a bath etc.), modification of footwear (e.g. thick compressible sole), local heat or cold applications, and transcutaneous nerve stimulation.

1.7.2 Pharmacological

Present pharmacological therapy in OA aims solely to improve the symptoms, mostly joint pain, and stiffness. Currently there is no drug that is licensed as a disease-modifying OA drug (DMOAD). FFigure 1.7-2 outlines the pharmacological management plan for OA.

Figure 1.7-2 Treatment algorithm for symptomatic OA



COX-2, Cyclo-oxygenase-2; NSAID, Nonsteroidal anti-inflammatory drug; PGE2, prostaglandin E2; PPI, proton pump inhibitor; PO, oral; SNRI, selective norepinephrine reuptake inhibitors; SSRI, selective serotonin reuptake inhibitor. (Argoff and Gloth, 2011) Produced with permission

Topical non-steroidal anti-inflammatory drugs (NSAIDS) can be considered as first-line treatment for readily accessible joints such as hands, knees, feet (but not deep joints such as hips). These drugs have better efficacy and minimum side-effects compared to oral NSAIDs and paracetamol (Derry *et al.*, 2016). A recent systematic review reports the good efficacy and excellent safety of topical capsaicin and topical NSAIDS used in licensed dosage in OA (Persson *et al.*, 2018).

The second preferred drug of choice is oral paracetamol. This has long been considered as the safest systemic analgesic and older trials confirmed its efficacy in reducing pain in OA (Zhang, Jones and Doherty, 2004; Towheed *et al.*, 2006). More recent trials, however,

show little or no benefit compared to placebo, and there are growing concerns over its side-effect profile (Roberts *et al.*, 2016).

Oral NSAIDs have analgesic, anti-inflammatory, and anti-pyretic effects and are effective for OA when used with safe dosage (da Costa *et al.*, 2017). Oral opioid analgesics are reserved for marked pain resistant to other analgesics. Oral NSAIDS and specific COX-2 inhibitors, particularly in older people, have several potentially serious gastrointestinal, renal, and cardiac side-effects in long term use. They should always be prescribed with a PPI, and there are numerous absolute or relative contraindications to their use in patients with comorbidity and on other medications (which apply to many people with OA).

According to a Cochrane meta-analysis, glucosamine, chondroitin, and their combination show slight pain reduction compared to placebo in hip and knee OA patients (Effect size -0.5, 95% confidence interval -0.9 to 0.0) (Wandel *et al.*, 2010). However, this effect is considered not clinically significant and, although popular as over-the-counter self-medications, they are not recommended by NICE (NICE, 2014).

A meta-analysis of 27 trials concluded that intra-articular corticosteroids helps in pain reduction (Jüni *et al.*, 2015) and they are recommended for consideration in people with pain resistant to other simple analgesics. The recommendation of hyaluronic acid use in OA is still debatable because of great heterogeneity of evidence (Fernández López and Ruano-Ravina, 2006) and NICE recommends to not use it within the NHS (NICE, 2014).

1.7.3 Surgical

If conservative management fails to give enough improvement, surgical treatment may need to be considered. Of the surgical options available, joint replacement for knee or hip OA are the most successful. Total joint replacement (TJR) is one of the preferred options for end-stage OA (Choong and Dowsey, 2014). However, although post-surgical improvement in pain and function has been reported by many observational studies, up to

20-30% of people still experience pain following TJR (Wylde *et al.*, 2017). The trend of TJR in UK is seen to be consistent decline over the past 20 years (Yu, Jordan and Peat, 2018). (Figure 1.7-3 and Figure 1.7-4)

Figure 1.7-3 Trends of primary hip replacement in UK among clinical and diagnosed OA (% of total diagnosed cases)

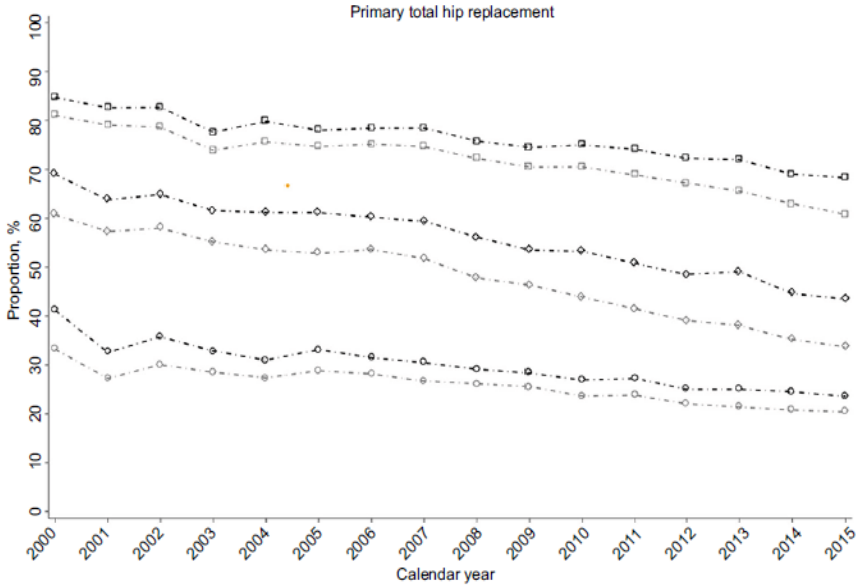
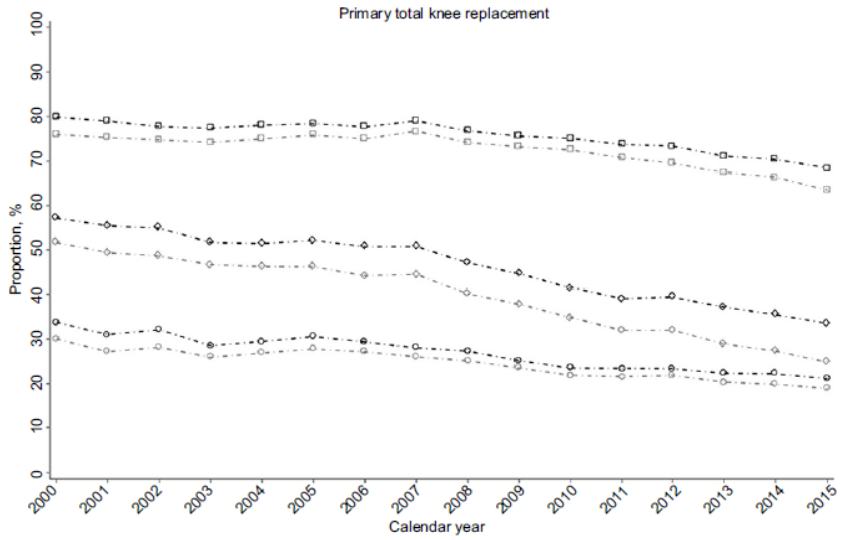


Figure 1.7-4 Trends of primary knee replacement in UK



Notes: Square, diamond, and circle line represents proportion of clinical OA, diagnosed OA (any joint), and diagnosed OA (joint-specific), respectively. Black and grey lines indicate the proportion with diagnosis in 10 years and 3 years prior to the index joint replacement, respectively. (Yu, Jordan, and Peat, 2018), produced with permission

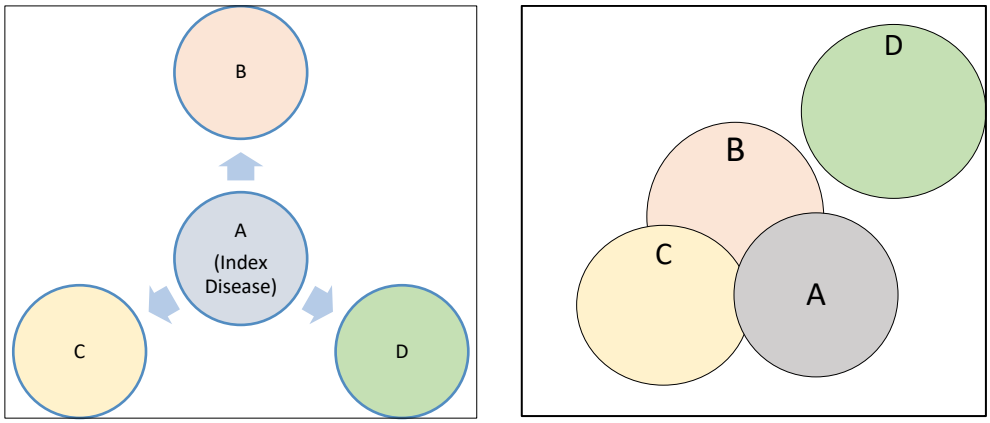
1.8 Comorbidities or Multimorbidity

1.8.1 Definitions and difference

The presence of multiple diseases in any given individual is becoming increasingly frequent. It can be termed as comorbidity or multimorbidity. Comorbidity is defined as "existence or occurrence of any distinct additional entity during the clinical course of a patient who has the index disease under study"(Feinstein, 1970), while multimorbidity denotes "the coexistence of two or more chronic diseases in one individual" (Akker, Buntinx and Knottnerus, 1996). A diagram showing the difference between the two is given in Figure 1.8-1.

Figure 1.8-1 Comorbidity and Multimorbidity

Comorbidity Multimorbidity



Comorbidity has been defined broadly in two ways:

1. Two or more medical conditions existing simultaneously but independent of each other.

2. Two or more medical conditions existing simultaneously and linked with each other.

Throughout this thesis, I used 'comorbidity' to OA irrespective of a causal relationship between the conditions. I also used the term multimorbidity when there were 2 or more other chronic conditions in OA and control groups without the index disease.

1.8.2 Measuring comorbidity

Apart from inconsistencies in definition of OA, the measurement of comorbidity has been challenging. Authors have developed and validated various tools for comorbidity measurement as listed in Table 1.8-1. Of these, commonly used tools are the Charlson comorbidity index (Charlson *et al.*, 1987) and the cumulative illness rating scale (Linn, Linn and Gurel, 1968). These tools differ in methods of capturing information on disease and very few tools have been designed to measure the burden and severity of the disease.

Table 1.8-1 Measuring comorbidity and multimorbidity

Tool	Author (year)	Measurement	Population	Data sources	System/ Condition	Items
Cumulative Illness Rating Scale (CIRS)	(Linn, Linn and Gurel, 1968)	Physical impairment		Clinical record	System	13 or 14 systems
Kaplan-Feinstein index KFI	(Kaplan and Feinstein, 1974)	Comorbidity among diabetics	188 men	Clinical record	System	12 systems
Charlson Comorbidity Index	(Charlson *et al.*, 1987)	Developing 'prognostic taxonomy' for comorbid conditions	608 general medical patients	Clinical record	Condition	17 conditions in 19 categories
Ambulatory Care Group	(Weiner *et al.*, 1991)	Predicting resource use in Health Maintenance Organization	16,000	Administrative data	Condition	93 mutually exclusive
Chronic Disease Score /Rx-Risk	(Von Korff *et al.*, 1992; Clark *et al.*, 1995)	To predict resource in Health Maintenance Organization (HMO)	122,911	Pharmaceutical data	Condition	Open
Index of Co-existent Disease	(Greenfield *et al.*,	To measure impact of comorbidity and physical functioning	356 hip replacement	Clinical records	System	14 systems

(ICED)	1993)		patients			
Elixhauser	(Elixhauser *et al.*, 1998)	To measure comorbidity using administrative data	1,779,167acute care hospital patients	Administrative data	Condition	30
Barnett	(Barnett *et al.*, 2012)	Selective chronic conditions	1.75 million	Administrative data	Condition	40

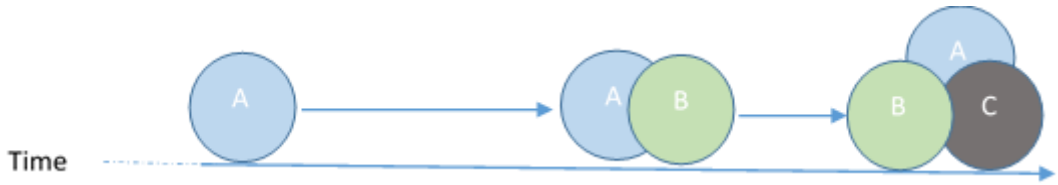
1.8.3 Types of comorbidity

The extensive list of possible comorbidities makes it difficult to group the conditions. Still, attempts have been made to classify comorbidities as per their relationship and aetiopathogenesis.

1.8.3.1 Primary vs Secondary (Feinstein, 1970)

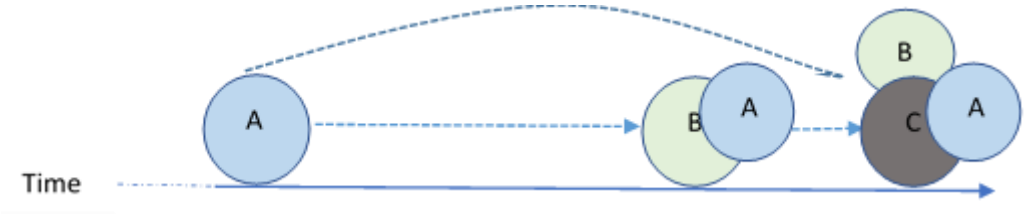
According to Feinstein, comorbidities can be grouped as primary or secondary based on chronological sequence and causal inference. Chronological comorbidity (Figure 1.8-2) is time dependent and develops in sequence. For example, in a person having chronic condition A, B develops in later life and C develops even later in life. There could be a linkage between the conditions, but this need not be so. In this case B is the primary comorbidity of A and C is the secondary comorbidity of A and can also be the primary comorbidity of B.

Figure 1.8-2 Chronological comorbidity model



In causal inference, condition B developed later than A but occurs because of A. Similarly, C can develop in late life which is induced by A or B or both. The causal link could be direct because of the disease, associated medication uses or other factors. Similarly, B is the primary comorbidity compared to C. (Figure 1.8-3)

Figure 1.8-3 Causal inference comorbidity model



1.8.3.2 Concordant and discordant comorbidity (Piette and Kerr, 2006)

Conditions of similar pathophysiologic risk profile and which are more likely to share a common management plan are grouped as concordant comorbidity. For example, diabetes and cardio-vascular diseases (CVDs) follow similar management programs and are classified as concordant comorbidity. Discordant conditions do not share similar pathophysiology or management plan. For example, osteoarthritis and asthma have different management plans but can exist together.

1.8.4 Comorbidity in OA

Abundant studies have been done on comorbidity in cardio-vascular diseases and psychiatric and cancer conditions, but there is a paucity of data on comorbidities in OA. This appears surprising, given that OA is one of the leading chronic conditions in older people, in whom multiple conditions are common. Two reasons could be that OA is not perceived to have a higher fatality index as does CVD, and it is considered an inevitable accompaniment of ageing. Other possible reasons could be the underexamined and perception of the unshared risk factors and no association with other conditions. However, recent research on comorbidity in OA is increasing, especially with respect to pain severity in OA which might be linked to central sensitisation mechanisms and coexistence of other painful conditions. Recently some reviews have examined the existing literature on individual comorbidities in OA, though on a separate individual basis.

1.8.4.1 OA and Cardio-Vascular Diseases (CVD)

A systematic review reported, among 358,944 participants, that the risk of CVD was significantly increased by 24% in patients with OA (n=80,911) compared with the general population (H. Wang *et al.*, 2016). Another review published in the same year by Hall et al reported a pooled prevalence of overall CVDs of 38.4% in people with OA and the risk of having heart failure was three times higher in OA compared to non-OA (Hall *et al.*, 2016). Similar findings have been documented by Parkinson et al (Parkinson, Waters and Franck, 2017).

1.8.4.2 OA and Diabetes

Louati et al, included 49 studies in a meta-analysis of the association of diabetes with OA. The prevalence of diabetes in people with OA was 14.4% and the risk of diabetes was 1.4 times higher compared to non-OA controls (Louati *et al.*, 2015). In 2016, Williams et al did another review of the association of OA with diabetes and reported in people with diabetes an odds ratio of 1.2 of developing OA (Williams *et al.*, 2016). One hospital based study reported the odds of having hand or knee OA was higher in female diabetics compared to males (Nieves-Plaza *et al.*, 2013).

1.8.4.3 OA and depression

Evidence for an association between OA and depression is inconclusive. A recently published review found a prevalence of depressive episodes in OA of 19.9%, and the prevalence of anxiety was 21.3%. However, the risk of having depression and anxiety in people with OA compared to non-OA population was not statistically significant (Stubbs *et al.*, 2016a).

1.8.4.4 OA and Chronic Obstructive Pulmonary Diseases (COPD)

The only review on OA with COPD reported a prevalence of OA in COPD of 35.5% (Wshah *et al.*, 2018).

1.8.4.5 OA with Fatigue and sleep disturbances

Fatigue is defined as weakness and tiredness. Fatigue in OA is not well researched compared to rheumatoid arthritis (RA) and other inflammatory diseases. People with OA often report tiredness, irritability and low mood and depression (Snijders *et al.*, 2011). More frequent in older age, it is associated with disturbances in sleep, early morning wakening and difficulty in falling asleep (Parmelee, Tighe and Dautovich, 2015). These disturbances could be due to OA pain and deprived sleep could amplify the pain and cause fatigue and can have fibromyalgia syndrome.

1.8.4.6 OA and musculoskeletal disorders

Joint pain and other musculoskeletal comorbidities in OA is well documented, especially, the relationship of back pain with OA (Suri *et al.*, 2010; Bollegala, Perruccio and Badley, 2011). Another primary care database study in UK reported positive associations of OA with other arthropathies, upper limb sprain, synovial and tendon disorders and other joint disorders (Kadam, Jordan and Croft, 2004).

1.8.4.7 OA and multiple comorbidities

Versus arthritis has published a report on multimorbidity in OA (Loftis, Ellis and Margham, 2014). According to this, one in five people with OA in the UK have at least one other chronic condition. Three of ten people aged 45 years or more with multimorbidity have musculoskeletal problems (Versus Arthritis, 2016). Even though musculoskeletal conditions are extremely common and cause impaired quality of life, their associations and outcomes have not been studied in detail.

1.8.5 Systematic review of OA and comorbidities

I did a systematic review and metanalysis of the comorbidities in OA in the beginning of my PhD (Swain et al. 2019).

Four databases for observational studies on comorbidities in individuals with OA were searched. Studies of OA only or in comparison with non-OA controls were included. The risk of bias and study quality were assessed using the Newcastle-Ottawa Scale. The prevalence of comorbidities in the OA group and the prevalence ratio (PR) and 95% confidence interval (95% CI) between OA and non-OA groups were calculated.

In all, 42 studies from 16 countries (27 case-only and 15 comparative studies) met the inclusion criteria. The mean age of participants varied from 51 to 76 years. The pooled prevalence of any comorbidity was 67% (95% CI 57–74) in individuals with OA versus 56% (95% CI 44–68) in individuals without OA. The pooled prevalence ratio (PR) for any comorbidity was 1.21 (95% CI 1.02–1.45). The PR increased from 0.73 (95% CI 0.43–1.25) for 1 comorbidity to 1.58 (95% CI 1.03–2.42) for 2, and to 1.94 (95% CI 1.45–2.59) for ≥3 comorbidities. The key comorbidities associated with OA were stroke (PR 2.61 [95% CI 2.13–3.21]), peptic ulcer (PR 2.36 [95% CI 1.71–3.27]), and metabolic syndrome (PR 1.94 [95% CI 1.21–3.12]).

Heterogeneity in the prevalence estimates observed in this review, stemming from diversity of methodologies, may have caused uncertainty of the results. There was ambiguity in disease definitions, for example over whether peptic ulcer, gastritis, and acidity should be considered separate entities. Suboptimal information about OA reported in studies made it difficult to differentiate between structural OA and symptomatic OA and to determine whether associations were linked primarily with structural OA or with pain experience. Similarly, the count of chronic conditions and the definition used varied considerably between studies and may have influenced the estimates. The comparative

groups included any non-OA cases, so the comorbidity pattern might have been different because of the selection of comparative/control groups, which needs to be interpreted with caution. Furthermore, the unavailability of joint-specific OA within comparative studies limited the estimation of joint-specific comorbidities. The study also compiles data from different study designs and thus has limitations for understanding the time sequences of OA with comorbidities. Unfortunately, there were not enough studies in each subgroup (only 1 in the cohort design) in comparative studies to perform subgroup analysis as per the study design.

Individuals with OA are more likely to have other chronic conditions. The association is dose-dependent in terms of the number of comorbidities, suggesting multimorbidity. Further studies on the causality of this association and clinical implications are needed. The published paper is attached as an appendix (Appendix- Publication 1 354).

1.9 Aim and objectives

1.9.1 Rationale

OA is a common chronic condition with subsequent significant detrimental impact on daily activities and quality of life (Szoeke *et al.*, 2006; Litwic *et al.*, 2013). Four out of five people with OA have at least one other long-term condition such as hypertension, CVD or depression (Breedveld, 2004). A few studies have reported the association of OA with multiple chronic diseases such as long-term widespread pain, CVD and diabetes, but the pattern and distribution of these have not been fully explored (Kadam, Jordan and Croft, 2004; Hoogeboom, Broeder, *et al.*, 2012; Zambon *et al.*, 2015a).

However, most research on comorbidity in OA has focused on CVD and metabolic syndrome. Musculoskeletal conditions such as OA, despite being one of the leading chronic conditions, have often been neglected in comorbidity/multimorbidity research. Because of a lack of clear understanding of the distribution and causal relationship between OA and its comorbidities, optimal management of the disease and its comorbidities remains undefined (de Rooij *et al.*, 2014). For example, although the association between OA pain and widespread pain has been investigated, the temporal direction of the causal association between these two conditions is unknown.

Aim: The overall aim of the thesis is to understand and explore the comorbidities occurrence in people with OA.

1.9.2 Objectives

The specific objectives are:

- To examine the trends in prevalence and incidence of OA in the UK population over the last 20 years
- To examine the temporal association between OA and comorbidities

- To identify the common clusters of comorbidities in OA and controls
- To understand the transition of people from one cluster to another with time in both OA and controls
- To explore the trajectory of the growth of multimorbidity and associated factors in people with OA and controls
- To explore the health care utilisations in people with OA and associated comorbidities

2 Chapter 2 Materials and methods

2.1 Population based studies using the Clinical Practice Data-Link (CPRD) GOLD database

A series of observational studies were carried out using the CPRD database to investigate the incidence and prevalence of OA, to identify the most common comorbidities associated with OA, their temporal associations with OA, and resultant healthcare utilization and mortality rates in people with OA compared to those with OA alone.

2.1.1 Definition of the Study population

I used data contained within the CPRD database (Nada F Khan, Harrison and Rose, 2010; Herrett, Arlene M. Gallagher, *et al.*, 2015). For this study CPRD GOLD was used. CPRD GOLD contains prospective healthcare data on around 17 million people from over 736 general practices throughout the UK and is a nationwide primary care database. For this study, data available for registered people from 1st January 1997 to 31st December 2017 were used.

General inclusion criteria:

- aged 20 years or more during the study year,
- have had minimum active registration for at least 12 months with the up-to-standard practice prior to the study,
- acceptable quality (decided by the CPRD based on certain parameters to select data suitable for research purpose)

2.1.2 Clinical Practice Research Datalink

The CPRD is one of the largest general practice based electronic databases created for health research. Started in 1987 in London, initially it was named as Value Added Medical Products (VAMP) research databank (Kousoulis, Rafi and de Lusignan, 2015). Later in 1993 it was renamed as the general practice research database (GPRD) and since 2012 it has been known

as the CPRD (Williams *et al.*, 2012). The database depends on the large general practices network in the UK. General practitioners in the UK are the gatekeepers of care of the National Health Service (NHS). The NHS covers around 98% of the UK population and patient data are routinely recorded on computers at each health centre (*National Health Services, UK*, no date). The anonymised electronic health record data of participating practices are routinely collated by the CPRD (*Clinical Practice Research Datalink* | *CPRD*, no date). CPRD has linkage to other health databases such as office of the national statistics (ONS) death registration, CPRD mother baby link, hospital episode statistics (HES) inpatient and outpatient data, national cancer registration and analysis service (NCRAS) cancer registration data, and data for index of multiple deprivation which enables access to multi-linked longitudinal data, thus enhancing the data available for health care research.

2.1.2.1 Contents of CPRD

Currently, CPRD data are available in two forms namely, CPRD GOLD and CPRD Aurum based on the software used at practice level. CPRD GOLD includes data from practices using Vision software, whereas CPRD Aurum includes data from practices which use EMIS-Web electronic patient record software. The basic differences in two databases are because of the different structure and coding system in use. The details of CPRD GOLD are explained later (next page).

Vision software is an electronic system for GPs to manage patients, which records information primarily as "events". The records of patient information exist from their first until their last contact with the participating practice. Practice level data are uploaded to CPRD monthly. The collected data are processed through the quality check to ensure internal consistency of patient data, complete longitudinal records, complete practice level data and compliance with CPRD recording guidelines. As of December 2017, the CPRD contained data on 17,480,766 patients of which 12.80% were considered as ‘non-acceptable’ for research after the quality check. With such a large amount of patient data and good representation of the UK population, the CPRD provides an excellent opportunity for research in areas such as health service research, drug

utilisation, clinical epidemiology, disease management, outcomes research, drug safety, health outcomes and pharmaco-economics.

2.1.2.2 Structure of CPRD

The CPRD collates data broadly in two categories, namely practice data and patient data against masked identifiers. In practice data, the geographical regions are recorded according to the 13 regions in the UK with ten from England and one each for Wales, Scotland, and Northern Ireland.

The database is further separated into clinical, referral, immunisation, test, and therapy data. It contains information on demographics, medical staff and practices, and clinical data, which records extensive clinical information on consultation, diagnoses, laboratory test/examination data, referral details, immunisation, and therapy. Descriptions of the CPRD structure are shown in the following Table 2.1-1.

Table 2.1-1. Details of the CPRD dataset

Data files	Demographic/clinical details	Registration details
Patient	Gender, birth year, birth month, marital status, family number, child health surveillance registration details, prescription exemption	Patient identifiers, VAMP identifier, first registration date, current registration date, registration status, registration gaps, internal transfer, transfer out date, transfer out reason, death date, acceptable
Practice		Practice identifier, region. Last collection date, up to standard date
Staff	Gender and role	Staff ID,
Consultation	consultation type,	Patient identifier, event date, system date, consultation identifier, staff identifier, duration
Clinical	Event date, consultation type, medical code, episode, additional details	Patient identifier, consultation identifier, staff identifier, entity type
Referral	Event date, medical code, NHS speciality, FHSA speciality, inpatient, attendance type, urgency	Patient identifier, consultation identifier, staff identifier, system date, source
Immunisation	Immunisation details	Patient identifier, consultation identifier, staff identifier, system date, source
Test	Medical code, consultation type, event date	Patient identifier, consultation identifier, staff identifier, entity type
Therapy	Details of the prescription	Patient identifier, consultation identifier, staff identifier, event date, system date

	(product code, dosage, BNF, quantity, duration)

FHSA- Family Health Services Appeal Authority; BNF- British National Formulary; VAMP- Value Added Medical Services; ID- Identification number

2.1.2.3 External linkage

Leading external databases are the Hospital Episode Statistic (HES), Office for National Statistics (ONS) mortality data, Cancer Registry, Myocardial Ischaemia National, Audit Project (MINAP), mother-baby link, national joint registry, and socioeconomic deprivation. Not all the practices are linked to external data source. ONS mortality data and HES are only available for English practices. Researchers must request these external linkages if needed with appropriate justification.

2.1.2.4 Diagnostic codes and validation

During the initial period of the CPRD, diagnoses were based on the Oxford Medical Information system (OXMIS), which was later replaced by Read codes in 1995. A substantial amount of research has been done to examine the validity and completeness of the CPRD, which provides satisfactory results. HES records data regarding hospital admissions, but uses ICD-10 for diagnosis, rather than the Read coding system.

2.1.2.5 CPRD Coverage

The CPRD GOLD had information for nearly 17 million patients by December 2017, representing 12% of the UK population. Of these, nearly 5 million were active registered patients. The database covers nearly 735 practices from 13 regions of the UK. However, the participation of practices across the regions is not uniform and some regions contribute more than others. According to a publication in 2013, the contribution of patients to the total database was highest from London (13%) and lowest from the East Midlands (0.7%) (Herrett, Arlene M Gallagher, *et al.*, 2015). This was validated later in 2016. According to Kontopontelis E. et al the Vision clinical computer system is used by less than 10% of practices and is heavily concentrated in three major conurbations and the Southern region (Kontopantelis *et al.*, 2018).

Table 2.1-2. Number of participating practices and CPRD GOLD-registered patients in 13 areas in the UK in 2017

	Practices (LCD)	**Number of patients (CRD)**			
Region	2017 n (%)	2014 n (%)	2015 n (%)	2016 n (%)	2017 n (%)
North East	3(0.96%)	3120(0.7%)	2493(0.7%)	2297(0.83%)	884(0.4%)
North West	23(7.4%)	33895(7.7%)	25717(7.2%)	15511(5.6%)	13690(6.0%)
Yorkshire and The Humber	2(0.64%)	2879(0.65%)	2408(0.7%)	1779(0.6%)	2315(1.0%)
East Midlands	0(0%)	271(0.06%)	0(0%)	0(0%)	293(0.1%)
West Midlands	21(6.7%)	39795(9.0%)	32683(9.1%)	22293(8.1%)	19357(8.5%)
East of England	9(2.9%)	26858(6.1%)	20363(5.7%)	12937(4.7%)	10582(4.6%)
South West	13(4.2%)	42418(9.6%)	26425(7.4%)	15821(5.7%)	9436(4.1%)
South Central	13(4.2%)	57954(13.1%)	43435(12.2%)	22663(8.2%)	14306(6.3%)
London	38(12.2%)	66860(15.1%)	44699(12.5%)	34029(12.3%)	29200(12.8%)
South East Coast	35(11.2%)	57098(12.9%)	50052(14.0%)	43212(15.6%)	28779(12.6%)
Northern Ireland	20(6.4%)	9230(2.1%)	9297(2.6%)	8730(3.1%)	7979(3.5%)
Scotland	68(21.8%)	46282(10.5%)	45812(12.8%)	43941(15.8%)	39468(17.3%)
Wales	67(21.5%)	55519(12.6%)	53801(15.1%)	53743(19.4%)	52163(22.8%)
Total	312 (100%)	442179 (100%)	357185 (100%)	276956 (100%)	228452 (100%)

LCD: Last Collection Date from practices; CRD: Current Registration Date of the patient. Dates for each year interval include 1st January of the year till 31st December of that year.

During 2017, data was collected from 312 practices. Table 2.1-2 describes the contribution of practices to 2017 data. Nearly 43% of the data are from practices in Scotland and Wales, whereas no practices were contributing from the East Midlands region. Similarly, from 2014 onwards, the number of current registered patients for each year in the East Midlands region was nearly zero. This validates the findings of other authors describing non-uniform distribution of the CPRD data in the UK.

2.1.2.6 Strengths and limitations of the CPRD

The CPRD is one of the largest primary care datasets, which includes information on morbidity and lifestyle factors with linkage to secondary care and mortality data.

A major strength of the database is the large number of patients and the longitudinal data, which allows researchers to investigate disease associations and outcomes. The CPRD is broadly representative of the general population of UK and makes population-based studies feasible. However, the representativeness at the practice level is debatable. The quality of the data is maintained by the Quality and Outcomes Framework (QOF), which was introduced in 2004. QOF has certain indicators for measuring the quality of services provided at practice level. Internal audits that exclude patients with non-continuous follow-up or poor recording that questions the validity of the patient's record, maintains data quality. With approval, investigators can obtain access to original medical records and contact GPs for additional questionnaire surveys of enrolled individuals. Other strengths include the availability of laboratory and examination results data, multiple external linkage, and high recording rate of secondary care events in GP records.

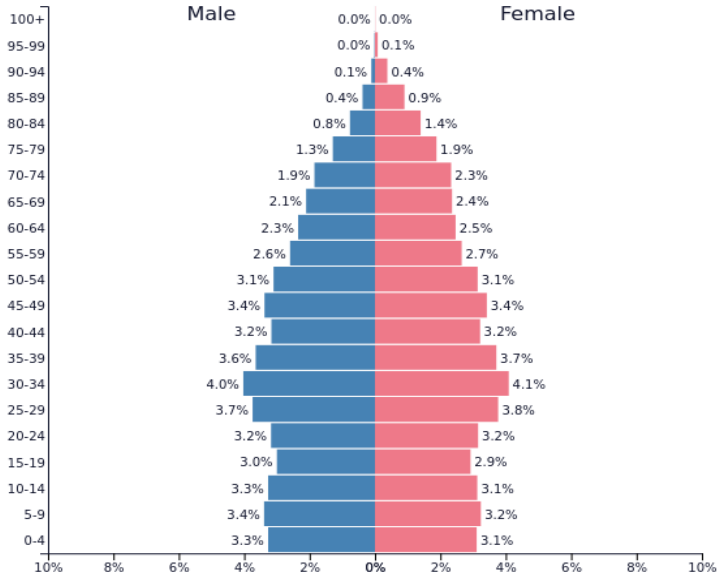
One of the key limitations of the CPRD is completeness of the data on every patient with missing values for BMI and other lifestyle factors. Also, information on socio-economic data, such as occupation and employment, are generally limited. However, recent external linkage to the Townsend score, an index of deprivation, helps to compensate for this. However, some of the indicators are available for England only. Missingness of lifestyle factors is important from the epidemiological research perspective. The CPRD records events within General Practice but information on hospital events may not be so complete. The addition of HES helps, though this is limited to data from England and Wales. There is no consistent definition for each disease, which needs to be developed and validated by the researchers. The size and complexity of the CPRD requires technical expertise.

2.1.3 Population structure of UK (1997, 2017)

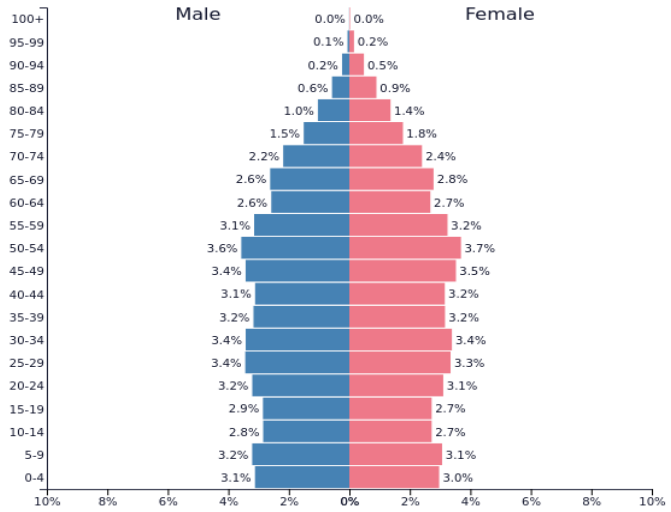
Being that OA predominates in older people, understanding the population structure is important in interpreting the findings. Because OA is strongly related with ageing, the trends of the incidence and prevalence are dependent on the changes in population structure of the country. Therefore, it is essential to understand the changes in the UK, which is depicted below.

Figure 2.1-1 Population structure of UK in 1997 and 2017

1997



2017



Source: Office For National Statistics, National Records Of Scotland and Northern Ireland Statistics And Research Agency, 2016

Over the twenty years from 1997 to 2017, there has been a change in population structure. In 2017 the percentage of the population aged 60-64 years and above is higher than in 1997. In 1997, there was a higher contribution from young adults (25 -40 years), which shifted higher in

the pyramid during 2017 (Office For National Statistics, National Records Of Scotland, and Northern Ireland Statistics And Research Agency, 2016). (Figure 2.1-1)

2.1.4 Study Design

- A repeated cross-sectional study design was used to estimate the annual prevalence of OA and a cohort study to estimate the annual incidence and prevalence of OA.
- A combined case-control and cohort study design was used to examine the temporal associations between OA and comorbidities.
- Latent class analysis was undertaken to identify common clusters of comorbidities ever diagnosed in OA and controls
- Latent transition analysis was performed to identify the movement of people within clusters after the index date (first recording of OA diagnosis)
- Latent trajectory analysis was used to estimate the trajectory of multimorbidity in OA and controls after the index date
- A cohort study was undertaken to examine the association of OA and the identified clusters with mortality and other health utilisation

2.1.5 Case definition

In this study OA is broadly identified in two categories; (1) physician-diagnosed; and (2) OA-related joint pain.

Read codes (clinical terminology used in general practice in the UK) was used to identify people with a diagnosis of incident OA and joint pain from the CPRD between 1st January 1997 and 31st December 2017. The available Read code list for these diagnoses (www.keele.ac.uk/mrr) was updated and adapted according to inclusion and exclusion criteria (e.g., physician-diagnosed OA and patients having total knee/hip replacement). For joint specific OA, separate codes for each of hip, knee, ankle/foot, wrist/hand, unspecified and generalised were extracted and used in further analysis.

2.1.5.1 Physician-diagnosed OA

- at least one recorded physician diagnosis of OA for hip, knee, ankle/foot, wrist/hand, elbow, shoulder joints, or recorded as ‘generalised’ and ‘unspecified’
- any recording of joint replacement

2.1.5.2 OA related joint pain

OA related joint pain was identified as a minimum of one consultation record of peripheral joint pain symptom of any of the joints (knee, hip, hand/wrist, and ankle/foot).

2.1.6 Exposures and outcomes

- For the case control study, the exposure was previous comorbidity, and the outcome is OA.
- For the cohort study, incident OA cases within 1st January 1997 until 31st December 2017 were considered as the exposed group, and matched control group of non-OA participants as a non-exposed group, the outcome of interest being subsequent comorbidity and mortality.
- For latent class analysis all the comorbidities were included as exposures and the clusters identified were the outcomes
- For outcomes of OA and the comorbidity clusters, both OA and clusters were included as the exposures, and mortality and other health utilisations were the outcomes

2.1.7 Extraction of comorbidity list

2.1.7.1 Comorbidity definition and extraction

Comorbidity was defined as the presence of any chronic conditions other than OA in individuals of both OA and non-OA control. An extensive list of 49 chronic conditions was prepared from the chronic diseases included in the Quality Outcome Framework (QOF) (*Quality and Outcome Framework (QOF)*, no date), chronic conditions listed in the US Department of Health and

Human Services Initiative on Multiple Chronic Conditions (Medicare, Baltimore and Usa, 2017) and the Charlson comorbidity index (Charlson *et al.*, 1987). The list was updated with findings from the systematic review and a previous community-based knee pain study (Sarmanova *et al.*, 2018; Swain *et al.*, 2019). Multimorbidity was measured as the presence of two or more chronic conditions in an individual, for further analysis (Akker, Buntinx and Knottnerus, 1996).

The 49 comorbidities in this study were further categorised into eight groups namely, musculoskeletal, respiratory, genitourinary, neuropsychiatric, cancer, circulatory, metabolic/endocrine, and digestive. In addition, a list of six conditions were grouped as 'other' category. The definition of all these conditions was based on physician diagnoses recorded as Read codes.

Comorbidity was defined as the presence of any chronic conditions (in OA group, any additional diseases other than OA) in individuals of both groups. An extensive list of 49 chronic conditions was prepared from the chronic diseases included in the Quality Outcome Framework (QOF) (*Quality and Outcome Framework (QOF)*, no date), chronic conditions listed in the US Department of Health and Human Services Initiative on Multiple Chronic Conditions (Medicare, Baltimore and Usa, 2017) and the Charlson comorbidity index (Charlson *et al.*, 1987). The list was updated with findings from the systematic review and a previous community-based knee pain study(Sarmanova *et al.*, 2018; Swain *et al.*, 2019). I preferred this approach over the commonly used Charlson comorbidity index alone (Charlson *et al.*, 1994; Quan *et al.*, 2011) because, although it is a useful predictor of mortality, the Charlson index summarises only 17 diagnostic categories to represent health status (specifically, myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, rheumatological disease, peptic ulcer disease, mild liver disease, moderate or severe liver disease, diabetes mellitus (DM), DM with chronic complications, renal diseases, any malignancy (including leukaemia and lymphoma), metastatic solid tumour and HIV infection). The Charlson index does not consider conditions like fibromyalgia, chronic fatigue syndrome, depression or back pain which are relevant to my study interests. Polymyalgia

included in the study was recorded as polymyalgia only not ‘polymyalgia rheumatica’. Multimorbidity was defined as the presence of two or more chronic conditions in an individual excluding OA, for further analysis (Akker, Buntinx and Knottnerus, 1996).

2.1.8 Data/ Statistical Analysis

All data analysis was done using R (version 3.5) and STATA (version 15 and 16) software. Details of the methods are given in respective chapters.

2.1.9 Study approval

This study was approved by the independent scientific advisory committee for CPRD research (protocol reference: 19_030 R).

3 Chapter 3 Incidence and prevalence of OA in the UK from 1997-2017: a trend analysis

3.1 Introduction

To date, very few studies are available on the trends of OA incidence and prevalence using national representative cohort data. The lack of such information creates challenges in reliable estimation of the burden of OA. Worldwide the incidence of OA has varied from 14.6 per 1000 person-years in Canada (Rahman *et al.*, 2014) to 40.5 per 1000 person-years in the UK (Yu *et al.*, 2017). Only three countries have reported increasing trends of the incidence of OA, whereas none have published prevalence trend data. In Sweden age-standardized hospitalisation rates due to OA increased from 1998 to 2014 for the hip and knee (Kiadaliri *et al.*, 2018). In Canada, during January 2000 to September 2008, crude OA incidence rates increased from 11.8 to 14.2 per 1000 person-years for men, and from 15.7 to 18.5 person-years for women (Rahman *et al.*, 2014). However, one recent UK study using the CPRD reported no change in the incidence of physician-diagnosed OA between 1992 to2013 (Yu *et al.*, 2017). According to Versus Arthritis UK, one third of people in the UK aged 45 years and over have sought treatment for OA (Versus Arthritis, 2019). In total, 8.75 million people in the UK have visited any health facility for OA treatment, and by 2035 it is estimated that 8.3 million people in the UK aged 45 years or over could have knee OA (Versus Arthritis, 2013).

Most available studies have used incidence to describe the burden of OA, but prevalence of any chronic disease is thought to be a better measure of the disease burden explaining the potential health resource users. This study aimed to explore the trends in both incidence and prevalence of OA (overall and joint specific) in the UK during the period 1997-2017 using this large nationally representative primary care database.

3.2 Methods

3.2.1 Source of data

CPRD GOLD data was used for registered people from the start of the database (1994) until 31st December 2017 for this study. Details of the selection of study participants is given in a flow diagram at Appendix Figure 1 (page 302).

3.2.1.1 Case Definition of OA

Read codes were used to identify people with a diagnosis of incident OA and/or joint pain from the CPRD between 1st January 1997 and 31st December 2017. The available Read code list (www.keele.ac.uk/mrr) was updated and adapted according to inclusion and exclusion criteria (e.g., physician-diagnosed OA or patients having total knee/hip replacement). Using the Read code list, the possible codes for OA diagnosis and OA-related joint pain were extracted separately. The extracted codes were matched with the obtained Read codes from the repository of Keele University (www.keele.ac.uk/mrr). After removing any duplication, a comprehensive list of the codes was prepared separately for both OA diagnosis and joint pain. Both lists were shared with the researchers (having expertise with Read codes), clinical experts and a general practitioner for their comment. After receiving inputs from all experts, the final list was prepared and used for the study. For joint-specific OA, separate codes for each of hip, knee, ankle/foot, wrist/hand, unspecified joint and generalised OA were extracted and used in further analysis. The codes are provided in Appendix Table 1 and Table 2 (pages 298-299).

OA was defined according to two definitions: (1) General Practitioner (GP) diagnosed; and (2) OA-related joint pain.

GP-diagnosed OA was defined as either:

- at least one recorded physician-diagnosis of OA for hip, knee, ankle/foot, wrist/hand, elbow, shoulder joints, or recorded as ‘generalised’ or ‘unspecified’, or

- any recording (in clinical file) of joint replacement in the absence of recording of GP-diagnosed OA during the study year

OA related joint pain was defined as:

- a minimum of one consultation record of peripheral joint pain of any of the joints (knee, hip, hand/wrist, and ankle/foot) without any record of OA before the date of joint pain diagnosis

For diagnosed OA and joint pain, the index date was defined as the first date of diagnosis recorded in the database.

3.2.1.2 Study population

General criteria for inclusion were:

- patients aged 20 years or more during each study year of 1997 to 2017
- patients who had active registration for at least 12 months with the up-to-standard (UTS) practice prior to the study start date (determined by CPRD database standards),
- flagged as 'Acceptable' in the database following the quality check

Different exclusion criteria were used for calculating incidence and prevalence. The exclusion criteria were used to avoid misclassification.

Exclusion criteria for incidence:

Patients meeting the following criteria were excluded from incidence estimation:

- For the 'GP-diagnosed' OA definition, any previous history of diagnosis of OA or any listed joint diseases (rheumatoid arthritis, psoriatic arthritis, systemic lupus erythematosus, ankylosing spondylitis, septic arthritis, spondyloarthropathy, crystal disease and human parvovirus B19 infection in the same joint recorded with joint pain/OA) before or within three years of the index date

- For the joint pain definition, any previous history of diagnosis of rheumatoid arthritis, psoriatic arthritis, systemic lupus erythematosus, ankylosing spondylitis, septic arthritis, and human parvovirus B19 infection in recorded joint pain consultation before the index date
- Any record of specific non-OA diagnosis (soft-tissue disorders, other bone/cartilage diseases) at the same joint within +/- 12 months the recorded OA/joint pain consultation.
- Any history of severe joint injury within one year prior to the index date

3.2.1.3 Estimation of incidence and prevalence

Annual prevalence of OA was calculated by dividing the number of people ever diagnosed with OA between 30th June previous year to 1st July of each calendar year, by the total number of eligible people in the population at the same time point of the calendar year.

Annual incidence rate for OA was calculated by dividing the number of incident (new) cases between 1st January to 31st December of each year, by the number of person-years at risk during that calendar year. Person-years of follow-up were calculated for eligible people at risk (i.e., no previous diagnosis of OA) from the latest of 1st January to the first diagnosis of OA, date of transfer-out, death, last data collection or 31st December of the study year whichever came first. Inclusion of study participants using the dates is depicted in detail in Appendix Fig-2 (page 303) and Appendix Table 3 (page 303).

3.2.1.4 Statistical analysis

The incidence and prevalence for each year from 1997 to 2017 were standardised according to age (5 years band), sex and length of data contribution (observation period) using the CPRD population structure in the year 2017 as reference. This method of adjustment for the observation period has been used previously (Kuo *et al.*, 2014b). The length of data

contribution of each patient was defined as the period from the UTS date for each participant to 1st July of each calendar year for prevalence and 1st January of each calendar year for incidence. The UTS date is always later than the registration date. The length of data contribution was then categorised in four groups 0-3 years, 4-6 years, 7-9 years and >=10 years. Standardization by length of data contribution was done because higher estimates were observed for longer lengths of data contribution. (Appendix Fig-3, page 304) For 1997, no data contribution was seen for >=10 years. (Appendix Fig 3 and 4, pages 304-305) Because, even though the first registration date with the database was traced back before 1987, the UTS practice data started recording in 1988, which is acceptable as a quality data, as per CPRD. For sex specific estimation, only age and length of data contribution standardisation was done. Age-sex standardized incidence and prevalence were calculated for 13 regions of the UK. Choropleth maps were used to represent the geographical variations of OA in the UK using QGIS (*QGIS Geographic Information System. Open Source Geospatial Foundation Project.*, 2016).

Age, sex, and length of data standardised trends (overall and sex specific) of the incidence and prevalence of OA were calculated for any-OA, joint specific and unspecified OA for 1998-2017. Unspecified OA cases are coded as 'unspecified' in the database without any mentioning of the site involved. The incidence and prevalence were estimated across each age group for both sexes only for the year 2017. The 95% CIs were derived based on the assumption of a Poisson distribution for the observed cases. The trends were tested using Joinpoint regression analysis (Kim *et al.*, 2000) and Joinpoint software (*Joinpoint Regression Program*, 2018). Bayesian Information Criterion (BIC) values were used to provide 'join points', which describe the significant change across the trend line and best-fit data series. Using BIC, a maximum of three joinpoints were selected. Annual percentage changes (APC) for each segment and average annual percentage changes (AAPC) for the entire study period were calculated at the significance level of 0.05.

Both incidence and prevalence trends were modelled as a function of age at diagnosis, period (year of diagnosis) and birth (year of birth) cohort. To assess the cohort effect, age-period-cohort (A-P-C) analysis was undertaken (Keyes *et al.*, 2010). For visual clarity incidence and prevalence were aggregated in five-year age groups for period and birth cohort graphs. The A-P-C analysis was performed in R using the package 'Epi' and 'APC' (Carstensen, 2005; Nielsen, 2015; Carstensen *et al.*, 2019). Statistical analyses were performed using STATA (SE v 15, STATA corp, Texas) and R(V 5.2, R software, Austria) (*R: A language and evironment for statistical computing.*, no date; 'StataCorp. 2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.', no date).

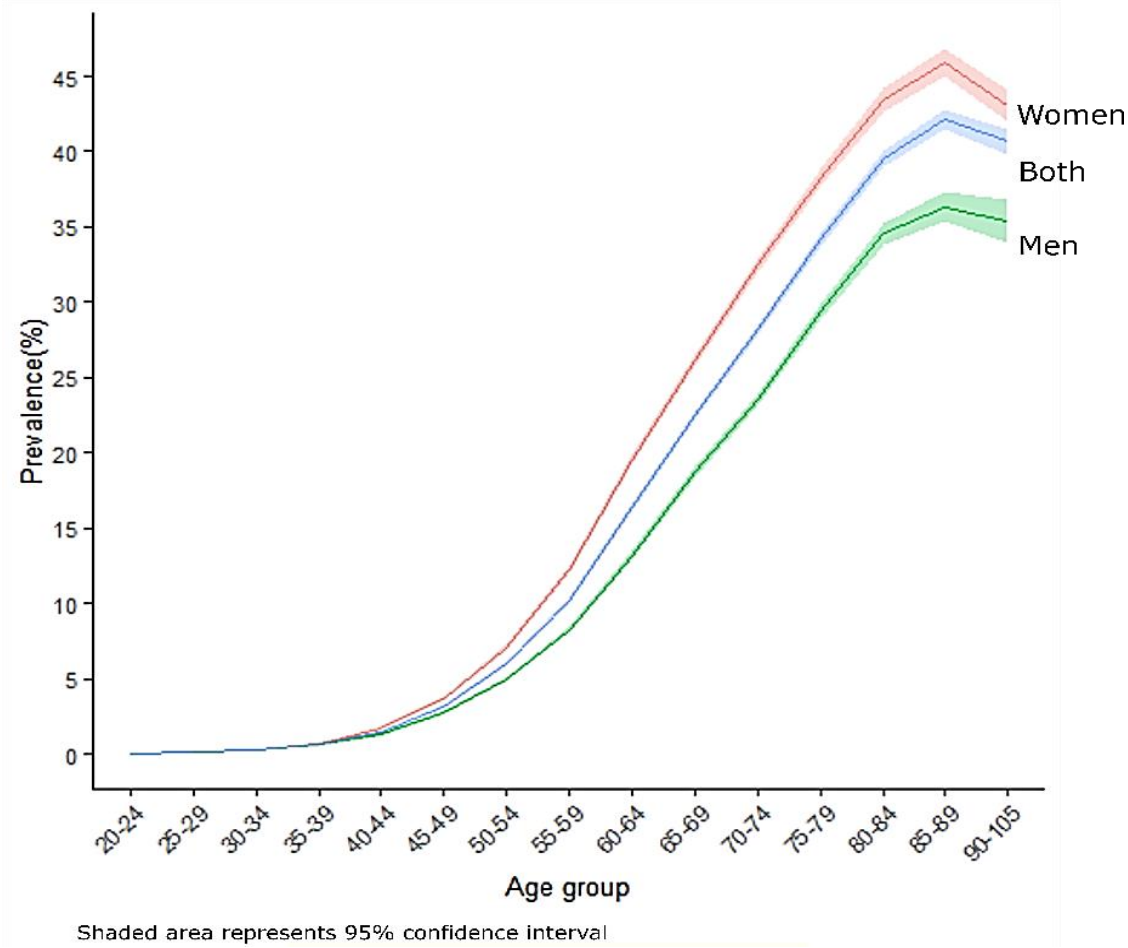
3.3 Results

3.3.1 GP diagnosed OA

3.3.1.1 Prevalence in 2017

Of 1,690,618 eligible individuals in 2017, 181,464 had a recorded diagnosis of OA at any site. The prevalence in 2017 was 10.78% (95% CI: 10.73%-10.83%). It was higher in women (12.79%; 95% CI 12.71%-12.86%) than in men (8.58%; 95%CI 8.52%-8.65%) across all age groups. The prevalence in younger age groups (less than 30 years) was very low, but it increased sharply at age 40-44 years in women and 45-49 years in men. In both men and women, the increasing trend continued until the age group of 80-84 years, reaching a peak of 47% for women and 35% for men. After ages 80-84, it plateaued and then started to decline after age 85-89 years. (Figure 3.3-1and Figure 3.3-1)

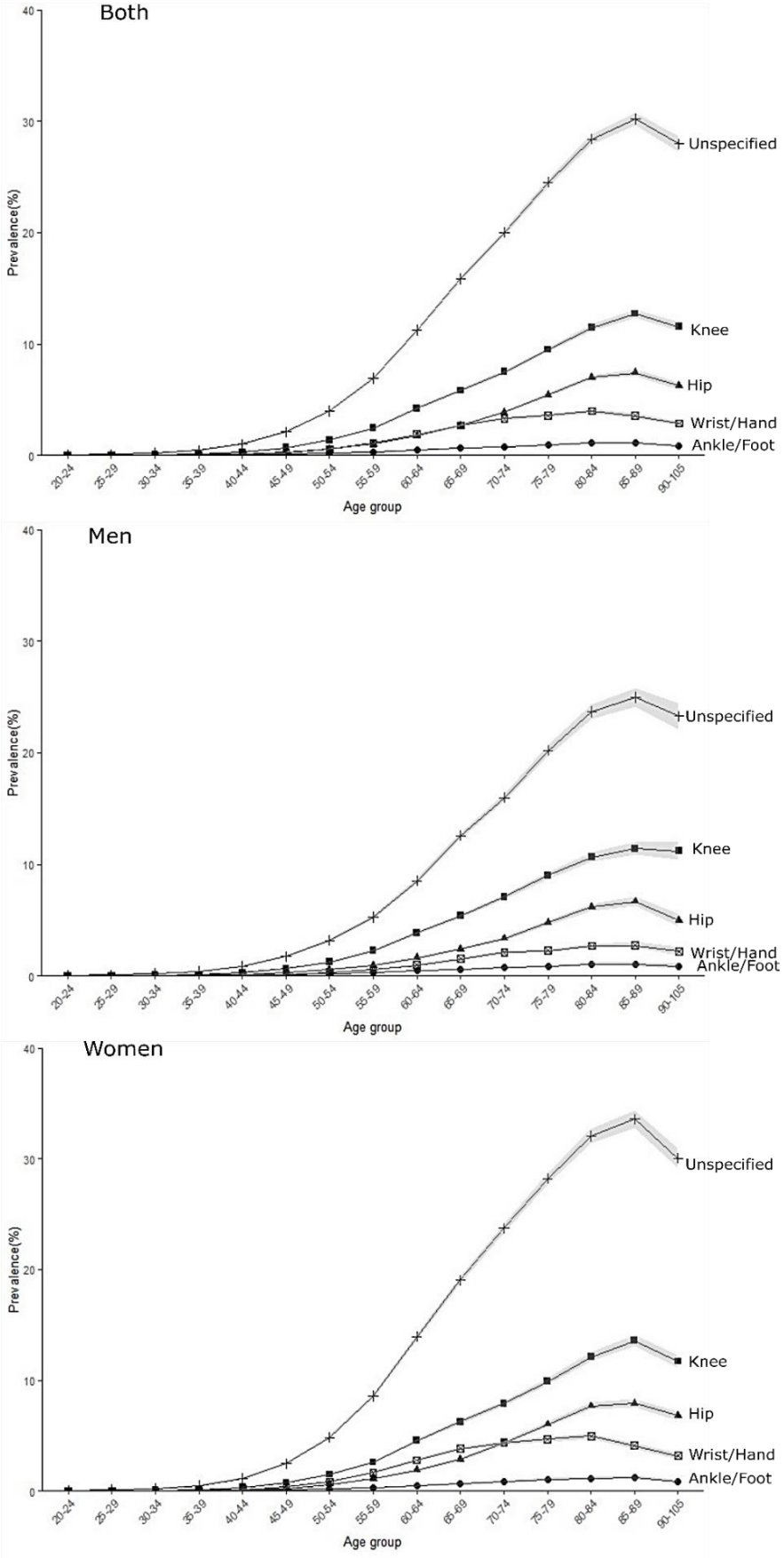
Figure 3.3-1 Prevalence of GP diagnosed OA by age and sex in 2017



Joint specific distribution of OA shows that the overall standardised prevalence estimates in descending order were as unspecified site (7.62%, 95%CI 7.58-7.65%), knee (2.86%, 95% CI 2.83-2.89%), hip (1.47%, 95%CI 1.45-1.49%), wrist or hand (0.52%, 95%CI 0.51-0.53%) and ankle or foot (0.29%, 95% CI 0.28-0.30%).

Joint specific OA prevalence across age in both men and women is shown in Figure 3.3-2. Among men, unspecified OA (6%) was the leading form of recorded OA followed by knee (2%), hip (1%), wrist and hand (0.7%), and ankle/foot (0.4%) OA. A similar pattern was seen in women, unspecified OA (8%) being the most recorded, followed by knee (2.5%), hip (1.2%), wrist and hand (1.1%), and ankle/foot (0.5%) OA. Across the age groups, the prevalence of all joint specific OA was seen to rise until age group 85-89 in both men and women, and then fall. In younger women, the prevalence of hand and wrist OA was higher than hip OA until the age group 70-74 years, after which hip OA became more common than hand and wrist OA.

Figure 3.3-2 Joint specific prevalence of GP diagnosed OA by age and sex in 2017



3.3.1.2 Trends of the prevalence (1997-2017)

Table 3.3-1 shows the temporal trend of any GP-diagnosed OA. Both the crude and standardised prevalence increased from 1997 to 2017. The standardized estimates were slightly higher than the crude estimates until 2014 after which they were slightly less for years 2016 and 2017. (Table 3.3-1) The age standardised rates were seen to rise in both men and women across the years.

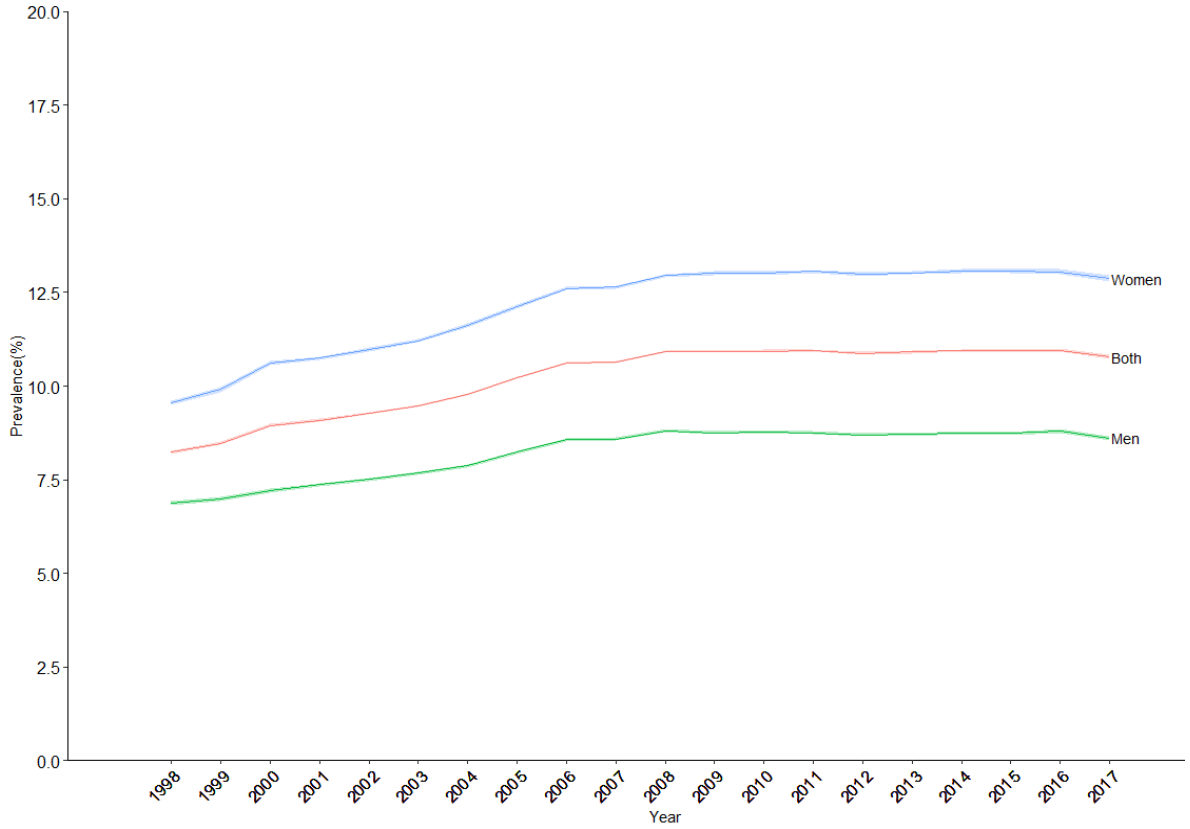
Table 3.3-1 Prevalence of GP diagnosed OA in the UK (1997-2017)

			Prevalence (%)		
Year	Eligible population	Cases	Crude [95% CI]	Age-sex Standardized [95% CI]	Age-sex-LOD standardized [95% CI]
1997	5711501	195362	3.42 [3.40-3.44]	6.15 [6.11-6.19]	
1998	5781677	215113	3.72 [3.70-3.74]	7.20 [7.16-7.24]	8.23 [8.06-8.40]
1999	5848216	234835	4.01 [3.98-4.03]	7.41 [7.37-7.45]	8.47 [8.39-8.55]
2000	5896329	255264	4.32 [4.30-4.35]	7.41 [7.37-7.44]	8.94 [8.88-9.00]
2001	5900383	276091	4.77 [4.74-4.80]	7.87 [7.83-7.90]	9.08 [9.03-9.13]
2002	5862771	296445	5.05 [5.02-5.08]	7.98 [7.95-8.01]	9.27 [9.22-9.32]
2003	5788957	317611	5.48 [5.45-5.51]	8.19 [8.16-8.22]	9.47 [9.42-9.52]
2004	5705620	339718	5.95 [5.92-5.98]	8.55 [8.52-8.58]	9.77 [9.73-9.82]
2005	5615033	363534	6.47 [6.43-6.52]	9.06 [9.03-9.09]	10.21 [10.16-10.26]
2006	5467107	378799	6.92 [6.90-6.94]	9.44 [9.42-9.47]	10.62 [10.57-10.66]
2007	5294313	388708	7.34 [7.30-7.38]	9.73 [9.71-9.76]	10.64 [10.60-10.68]
2008	5112496	398003	7.78 [7.74-7.82]	10.07 [10.04-10.10]	10.91 [10.87-10.95]
2009	4924529	405402	8.23 [8.20-8.26]	10.35 [10.32-10.38]	10.91 [10.88-10.95]
2010	4689058	403343	8.60 [8.56-8.64]	10.54 [10.51-10.57]	10.93 [10.90-10.96]
2011	4421201	398434	9.01 [8.96-9.06]	10.69 [10.66-10.72]	10.94 [10.91-10.97]
2012	4165371	391691	9.40 [9.36-9.44]	10.76 [10.73-10.79]	10.87 [10.84-10.90]
2013	3812788	374298	9.82 [9.78-9.86]	10.87 [10.84-10.90]	10.90 [10.87-10.93]
2014	3314992	337168	10.17 [10.14-10.20]	10.96 [10.93-10.99]	10.95 [10.92-10.98]
2015	2761702	290020	10.50 [10.47-10.53]	10.94 [10.90-10.97]	10.93 [10.90-10.96]
2016	2100061	223948	10.66 [10.63-10.69]	10.96 [10.93-11.00]	10.95 [10.92-10.99]
2017	1690618	181464	10.77 [10.72-10.82]	10.77 [10.72-10.82]	10.77 [10.72-10.82]

Age-sex and length of data contribution (LOD) standardization was done using 2017 CPRD population as standard population. For 1997, LOD standardisation was not calculated because of absence of data for >=10 years. (See Appendix Figure 3 and 4, page 304-305)

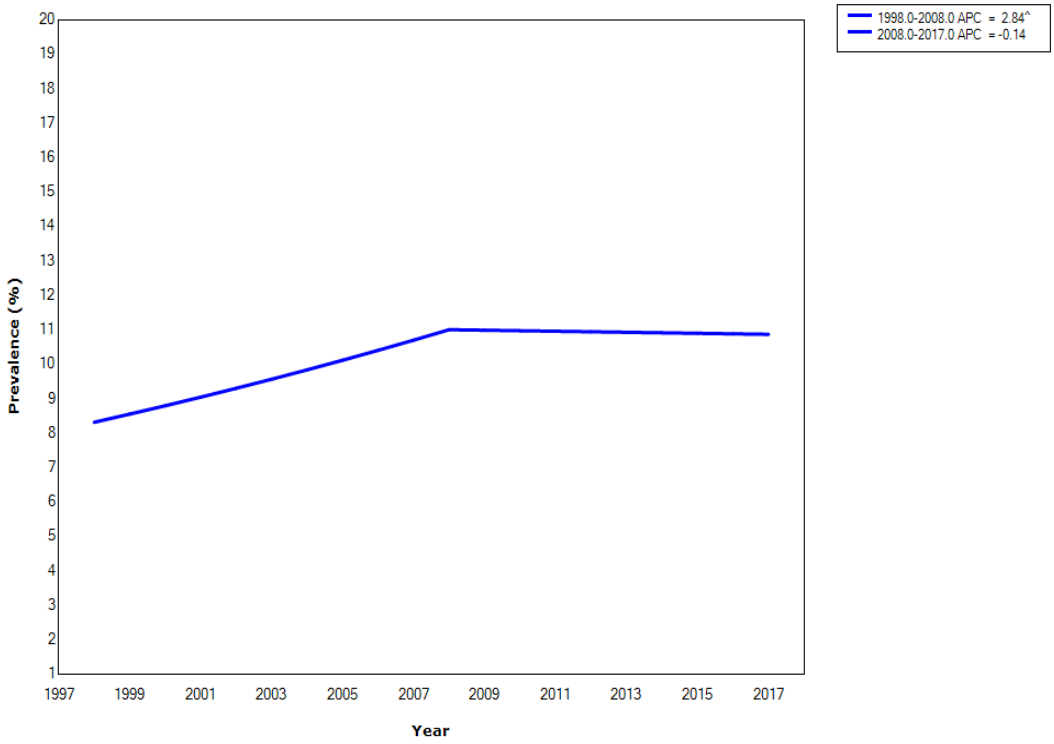
The overall standardized prevalence of people with any OA in 2017 was found to increase to 10.7% from the 8.2% reported in 1998, a two-fold increase in prevalence over this period. (Figure 3.3-3)

Figure 3.3-3 Trends of standardized prevalence of GP diagnosed OA in the UK 1997-2017



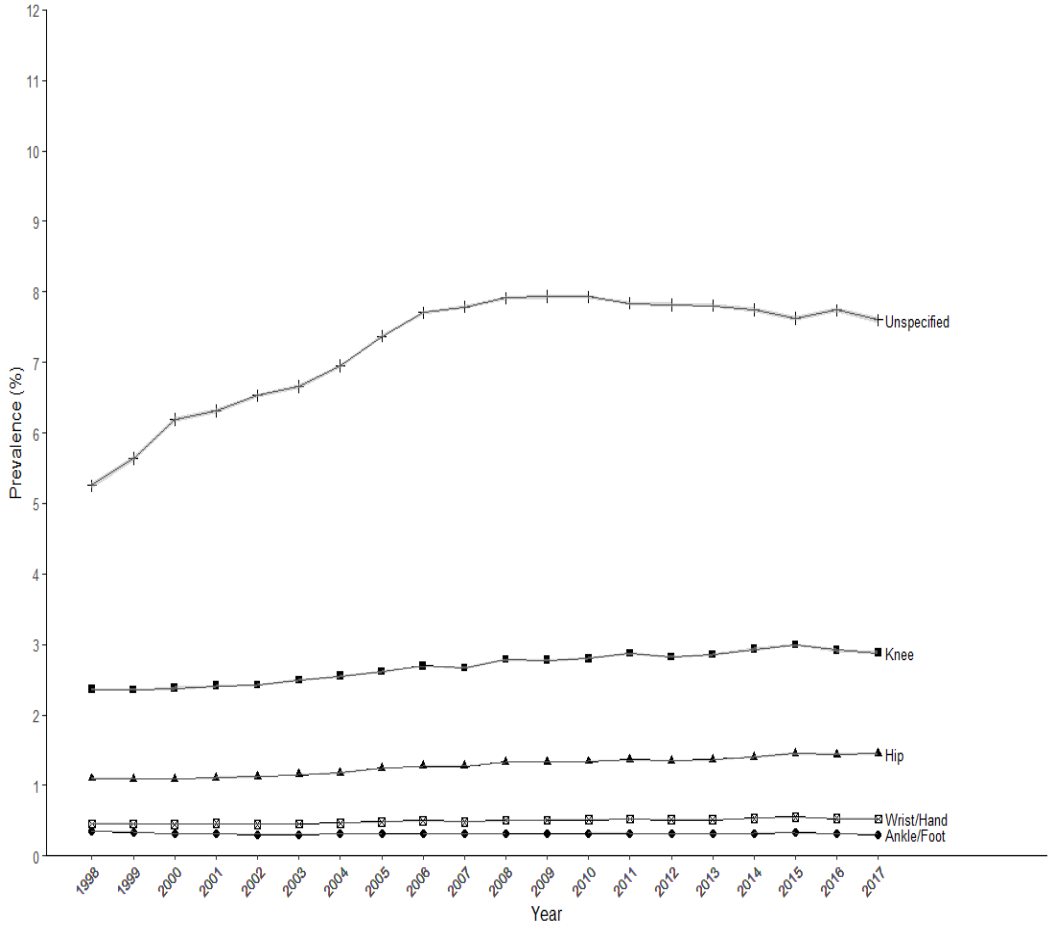
The temporal trends of prevalence for men and women are not parallel across the study period. (Figure 3.3-3) In 1998 the difference in prevalence between men and women was 2.5% which increased to nearly 5% in 2017. The average annual percentage change was 1.4% (95% CI 1.3-1.6%) overall, being 1.6% (95% CI 1.4-2.8%) in women and 1.3% (95% CI 1.1-1.4%) in men. The joinpoint trend analysis showed a statistically significant increase in prevalence during the years 1998-2008 and non-significant change during the period 2008-2017. (Figure 3.3-4)

Figure 3.3-4 Joinpoint trend analysis



The trend in prevalence has increased for all the joint specific OA sites since 1998 (Figure 3.3-5). The prevalence of unspecified site OA increased from 5.2% in 1998 to 7.8% in 2017. However, after 2014 the trend line was seen to increase for all the sites, except for ankle and foot OA.

Figure 3.3-5 Joint specific trends for GP diagnosed OA in the UK from 1997-2017



As Table 3.3-2 and Table 3.3-3 show, the temporal trends were higher in women compared to men. In men, the trend of wrist and hand OA was seen to be parallel with hip OA, whereas, in women, the gap started to narrow from 2005. The highest increase was seen for unspecified OA (APC 1.9%; 95%CI 1.6-2.2%), followed by hip OA (APC 1.7%; 95%CI 1.3-2.0%) in both sexes. Details of sex differences are given in Table 3.3-2 and Table 3.3-3.

Table 3.3-2 Trend in prevalence of GP diagnosed OA by joints in men

	Standardised Prevalence (%) [95% CI]				
Years	Hip	Knee	Wrist/Hand	Ankle/Foot	Unspecified
1998	0.95 [0.92-0.98]	2.19 [1.15-2.24]	0.30 [0.27-0.33]	0.28 [0.24-0.32]	4.18 [4.15-4.21]
1999	0.86 [0.83-0.89]	2.14 [2.11-2.15]	0.28 [0.24-0.31]	0.31 [0.29-0.33]	4.48 [4.45-4.51]
2000	0.90 [0.87-0.93]	2.11 [2.06-2.16]	0.27 [0.26-0.29]	0.27 [0.25-0.29]	4.82 [4.75-4.89]
2001	0.92 [0.89-0.95]	2.13 [2.09-2.18]	0.27 [0.26-0.29]	0.27 [0.25-0.29]	4.92 [4.86-4.98]
2002	0.93 [0.91-0.96]	2.14 [2.10-2.18]	0.27 [0.25-0.28]	0.27 [0.25-0.28]	5.08 [5.02-5.14]
2003	0.96 [0.93-0.98]	2.21 [2.18-2.25]	0.27 [0.26-0.29]	0.27 [0.26-0.28]	5.17 [5.12-5.23]
2004	0.98 [0.95-1.01]	2.26 [2.22-2.30]	0.28 [0.26-0.29]	0.27 [0.26-0.29]	5.37 [5.31-5.43]
2005	1.04 [1.01-1.06]	2.32 [2.28-2.36]	0.28 [0.27-0.30]	0.28 [0.26-0.29]	5.70 [5.65-5.76]
2006	1.05 [1.03-1.08]	2.39 [2.35-2.42]	0.29 [0.28-0.31]	0.28 [0.27-0.30]	5.97 [5.92-6.02]
2007	1.05 [1.03-1.07]	2.36 [2.33-2.39]	0.28 [0.27-0.30]	0.27 [0.26-0.29]	6.03 [5.98-6.08]
2008	1.10 [1.08-1.12]	2.46 [2.43-2.49]	0.30 [0.28-0.31]	0.29 [0.28-0.30]	6.13 [6.09-6.18]
2009	1.09 [1.07-1.11]	2.42 [2.40-2.45]	0.29 [0.28-0.30]	0.28 [0.28-0.29]	6.11 [6.07-6.15]
2010	1.10 [1.08-1.12]	2.44[2.41-2.47]	0.29 [0.28-0.30]	0.29 [0.28-0.30]	6.12 [6.08-6.16]
2011	1.12 [1.11-1.14]	2.50 [2.48-2.52]	0.30 [0.29-0.31]	0.29 [0.28-0.30]	6.03 [5.99-6.06]
2012	1.11 [1.09-1.12]	2.46 [2.43-2.48]	0.29 [0.29-0.30]	0.28 [0.27-0.29]	5.99 [5.96-6.03]
2013	1.12 [1.11-1.14]	2.49 [2.46-2.51]	0.29 [0.28-0.30]	0.28 [0.27-0.29]	5.97 [5.93-6.00]
2014	1.16 [1.14-1.18]	2.56 [2.53-2.58]	0.30 [0.30-0.31]	0.29 [0.28-0.30]	5.92 [5.88-5.95]
2015	1.20 [1.18-1.22]	2.60 [2.58-2.63]	0.32 [0.31-0.32]	0.30 [0.29-0.31]	5.79 [5.76-5.83]
2016	1.19 [1.17-1.21]	2.55 [2.53-2.58]	0.30 [0.29-0.31]	0.28 [0.27-0.29]	5.91 [5.87-5.95]
2017	1.20 [1.18-1.22]	2.51 [2.48-2.54]	0.29 [0.28-0.31]	0.27 [0.26-0.28]	5.77 [5.73-5.82]
AAPC	1.5 [0.9 to 1.9] *	0.9 [0.4 to 1.4] *	0.2 [-0.4 to 0.9]	0.1 [-0.2 to 0.4]	1.7 [1.3 to 2.1] *

AAPC: Average annual percentage change; CI- Confidence interval; Age and length of data standardized using 2017 CPRD population; *P-value <0.05

Table 3.3-3 Trend in prevalence of GP diagnosed OA by joint in women

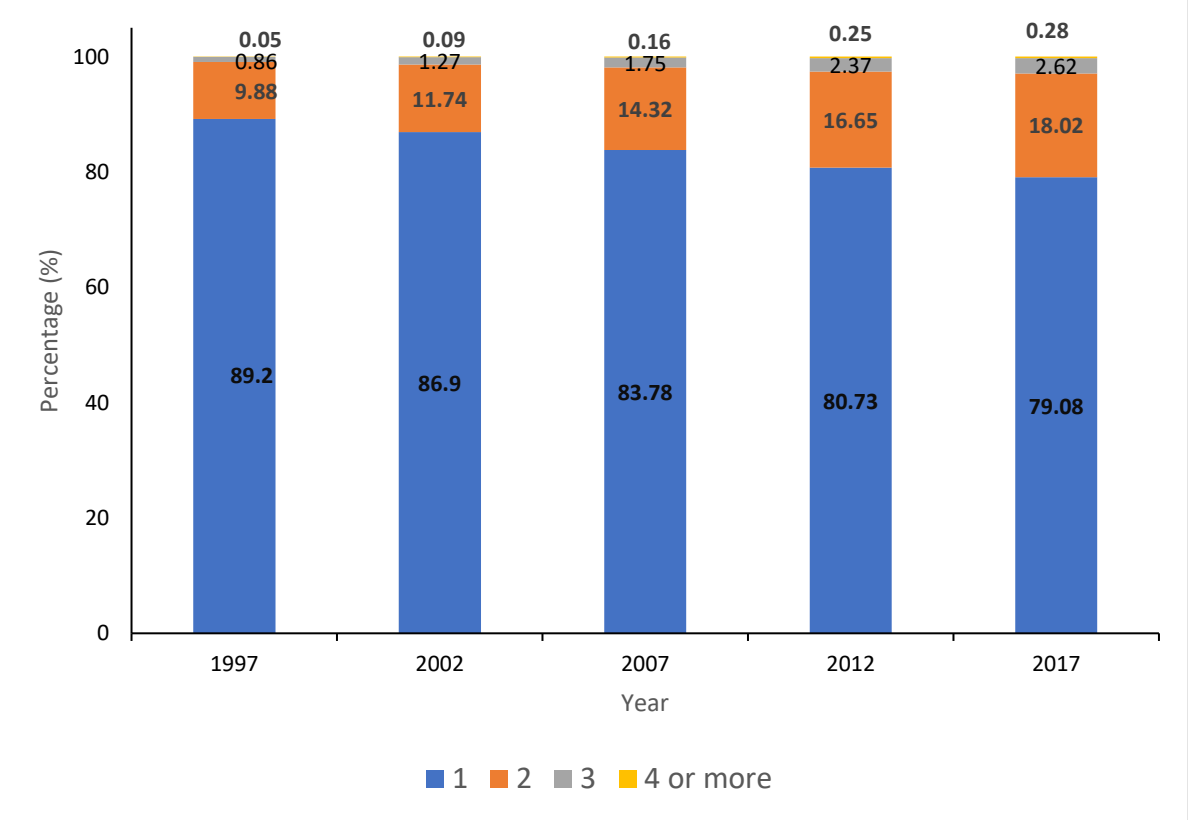
	Standardised Prevalence (%) [95% CI]				
Years	Hip	Knee	Wrist/Hand	Ankle/Foot	Unspecified
1998	1.24 [1.21-1.27]	2.51 [2.48-2.54]	0.60 [0.56-0.64]	0.41 [0.38-0.44]	6.27 [6.24-6.30]
1999	1.18 [1.15-1.21]	2.56 [2.52-2.60]	0.62[0.59-0.65]	0.43 [0.40-0.47]	6.75 [6.71-6.79]
2000	1.25 [1.22-1.29]	2.63 [2.58-2.68]	0.61 [0.58-0.63]	0.35 [0.33-0.36]	7.51 [7.43-7.59]
2001	1.28 [1.25-1.31]	2.67 [2.63-2.72]	0.63 [0.60-0.65]	0.34 [0.32-0.36]	7.64 [7.57-7.72]
2002	1.30 [1.27-1.33]	2.69 [2.65-2.73]	0.61 [0.59-0.63]	0.33 [0.32-0.35]	7.92 [7.85-7.99]
2003	1.33 [1.30-1.36]	2.76 [2.72-2.80]	0.62 [0.60-0.64]	0.33 [0.31-0.34]	8.08 [8.02-8.15]
2004	1.36 [1.33-1.39]	2.81 [2.77-2.85]	0.64 [0.62-0.66]	0.34 [0.32-0.35]	8.46 [8.40-8.53]
2005	1.43 [1.40-1.46]	2.89 [2.85-2.93]	0.66 [0.64-0.68]	0.34 [0.33-0.35]	8.97 [8.90-9.03]
2006	1.48 [1.45-1.51]	2.98 [2.94-3.01]	0.68 [0.67-0.70]	0.35 [0.34-0.36]	9.36 [9.30-9.42]
2007	1.48 [1.45-1.50]	2.96 [2.92-2.99]	0.67 [0.65-0.68]	0.33 [0.32-0.34]	9.47 [9.41-9.53]
2008	1.55 [1.52-1.57]	3.09 [3.06-3.13]	0.69 [0.68-0.71]	0.35 [0.33-0.36]	9.62 [9.57-9.67]
2009	1.55 [1.53-1.57]	3.10 [3.07-3.13]	0.71 [0.69-0.72]	0.34 [0.33-0.35]	9.68 [9.64-9.73]
2010	1.56 [1.54-1.58]	3.13 [3.11-3.16]	0.71 [0.70-0.72]	0.34 [0.33-0.35]	9.66 [9.62-9.71]
2011	1.59 [1.57-1.61]	3.22 [3.19-3.24]	0.73 [0.71-0.74]	0.35 [0.34-0.36]	9.58 [9.54-9.62]
2012	1.57 [1.55-1.59]	3.16 [3.13-3.19]	0.71 [0.70-0.73]	0.33 [0.32-0.33]	9.57 [9.52-9.61]
2013	1.58 [1.57-1.60]	3.20 [3.17-3.23	0.72 [0.70-0.73]	0.33 [0.32-0.34]	9.56 [9.52-9.60]
2014	1.63 [1.61-1.65]	3.29 [3.26-3.31]	0.74 [0.73-0.76]	0.34 [0.33-0.35]	9.50 [9.46-9.54]
2015	1.68 [1.66-1.70]	3.36 [3.33-3.39]	0.77 [0.76-0.79]	0.35 [0.34-0.36]	9.38 [9.34-9.42]
2016	1.65 [1.63-1.68]	3.26 [3.23-3.29]	0.74 [0.73-0.76]	0.33 [0.32-0.34]	9.49 [9.44-9.54]
2017	1.68 [1.65-1.70]	3.23 [3.19-3.26]	0.74 [0.72-0.75]	0.32 [0.31-0.34]	9.35 [9.30-9.41]
AAPC	1.8 [1.5 to 2.1] *	1.3 [1.0 to 1.7] *	1.3 [1.1 to 1.5] *	-1.2 [-2.0 to -0.3] *	2.2 [1.9 to 2.5] *

AAPC: Average annual percentage change; CI- Confidence interval; Age and length of data standardized using 2017 CPRD population; **P*-value <0.05

3.3.1.3 Prevalence of GP diagnosed multiple joint OA

The prevalence of multiple joint OA, i.e., OA affecting more than one joint in the same individual was examined. Five types of OA reported in different joints (hip, knee, ankle/foot, wrist/hand and unspecified) were included for this count. The trend was calculated at five year intervals starting from 1997. In 2017 nearly 21% of people with OA had it reported in more than one joint, this prevalence having increased by 10% from 1997. (Figure 3.3-6) While the prevalence of single joint OA declined, the crude prevalence of multiple joint OA increased.

Figure 3.3-6 Trends in number of sites involved according to GP-diagnosed OA (1997-2017)



Further analysis of pattern of joint involved among people having OA diagnosed at any-two joint revealed knee, hip, and wrist/hand in combination with unspecified OA contributing more than 70% of the total pattern. (Figure 3.3-7)

Figure 3.3-7 Proportion of pattern of any two-sites involved in 2017

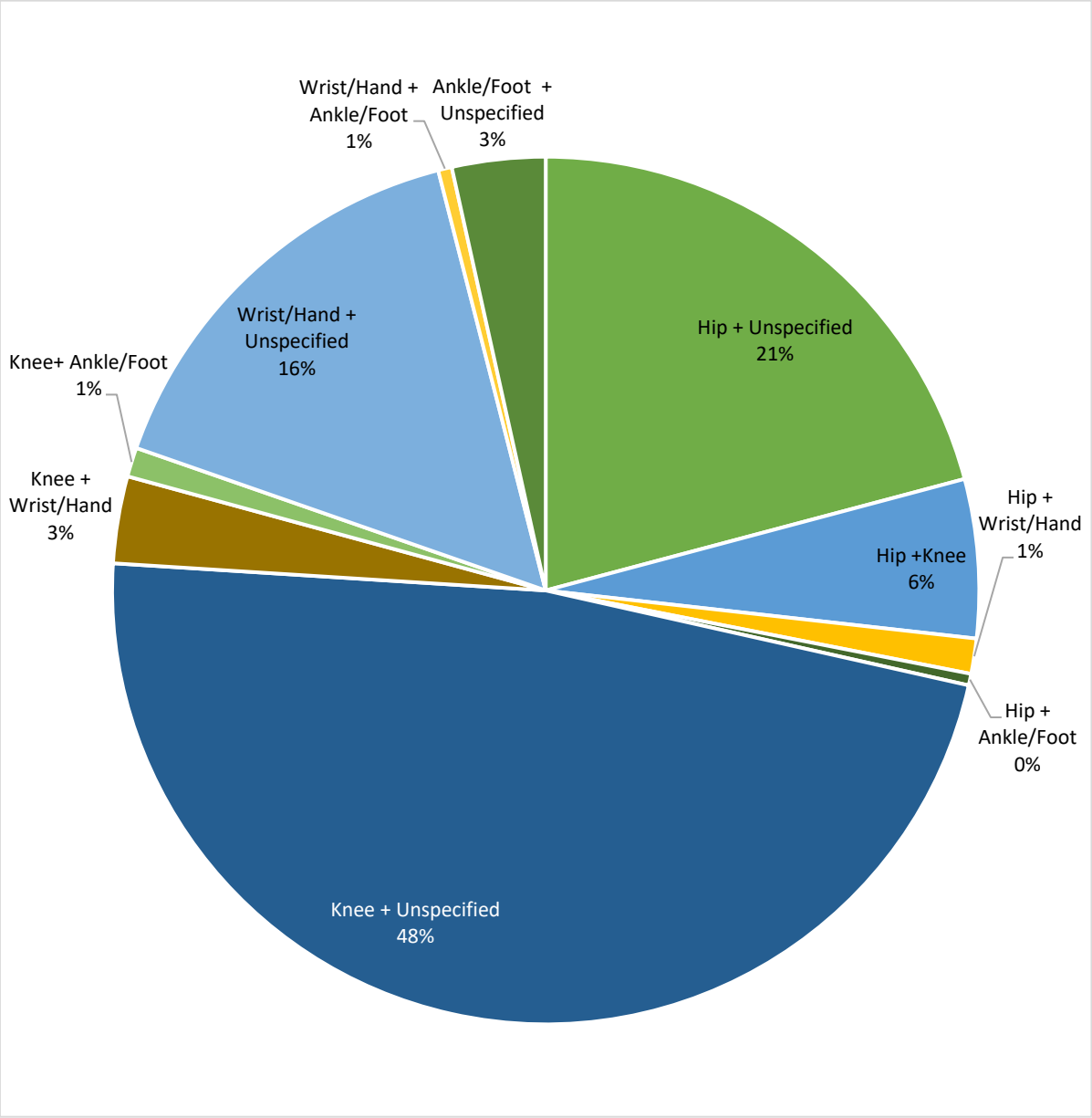
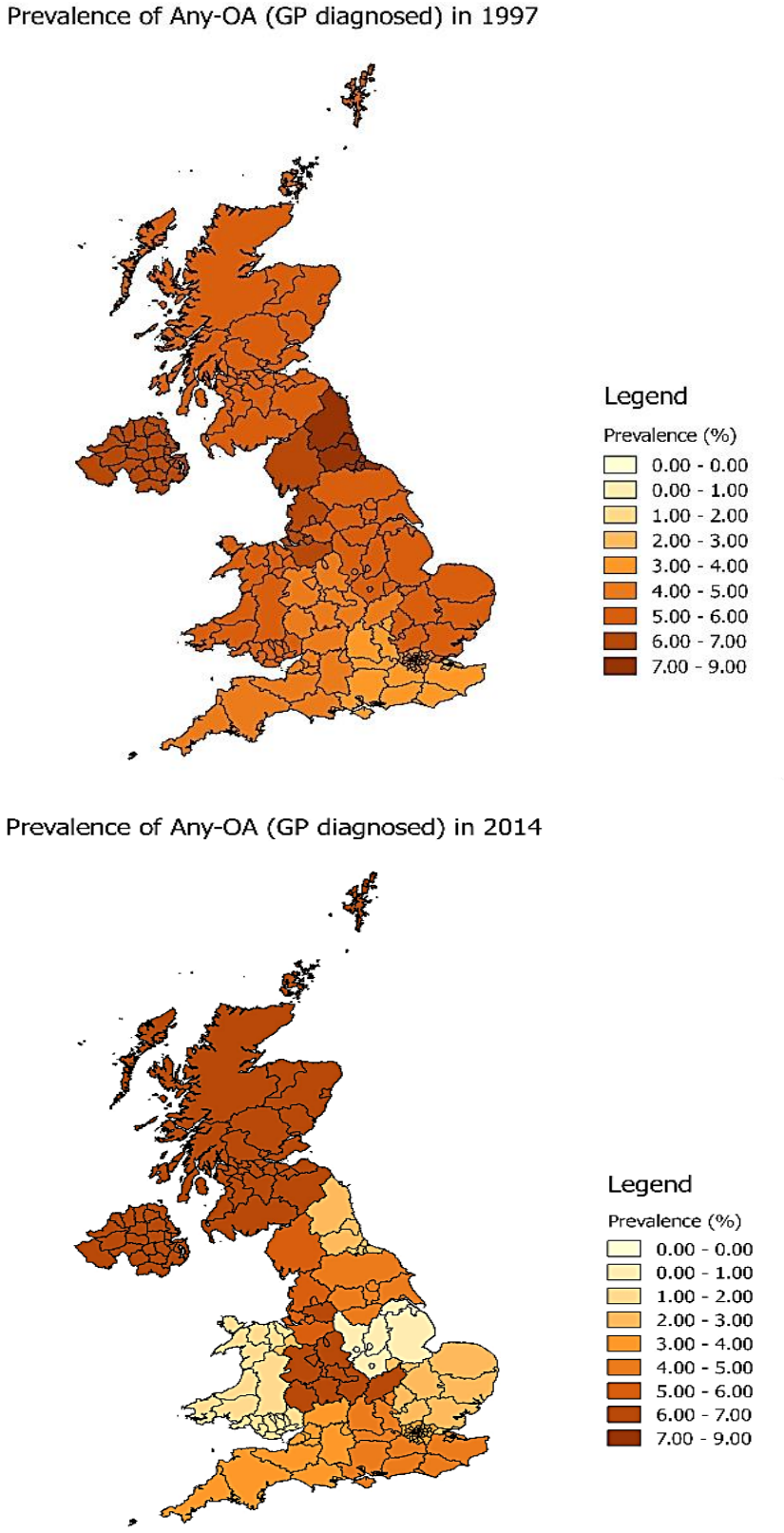


Figure 3.3-8 Prevalence of GP-diagnosed OA in different regions of the UK in 1997 and 2014

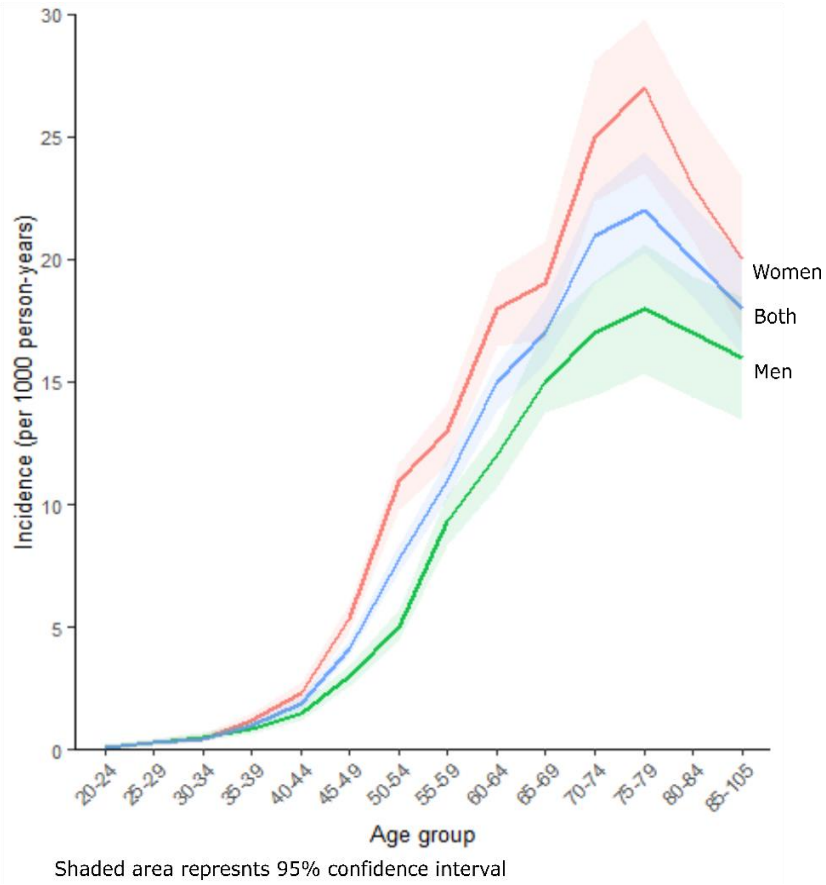


The contribution from East Midlands to the CPRD-GOLD was nearly zero, which did not allow to have whole country representation for 2017.

3.3.1.4 Incidence of GP diagnosed OA in the UK

In 2017 the total person-years of follow up for any OA was 1,495,497 with 10147 incident cases, and the incidence was 6.8 per 1000 person-years [95% CI 6.7-6.9 person-years]. The incidence was higher in women (8.1; 95% CI 7.9 to 8.3) than in men (5.5; 95% CI 5.3 to 5.7 per 1000 person-years). Age specific incidence in 2017 shows that, at the younger age (30-34 years) the incidence was 0.08 per 1000 person-years in both sexes but this became increasingly higher in women than in men after 40 years of age. The incidence peaked at the age of 75-79 years to 27.0 [95% CI 23.5-29.8] per 1000 person-years in women and 18.0 [95% CI 15.4-20.6] per 1000 person-years in men. (Figure 3.3-9)

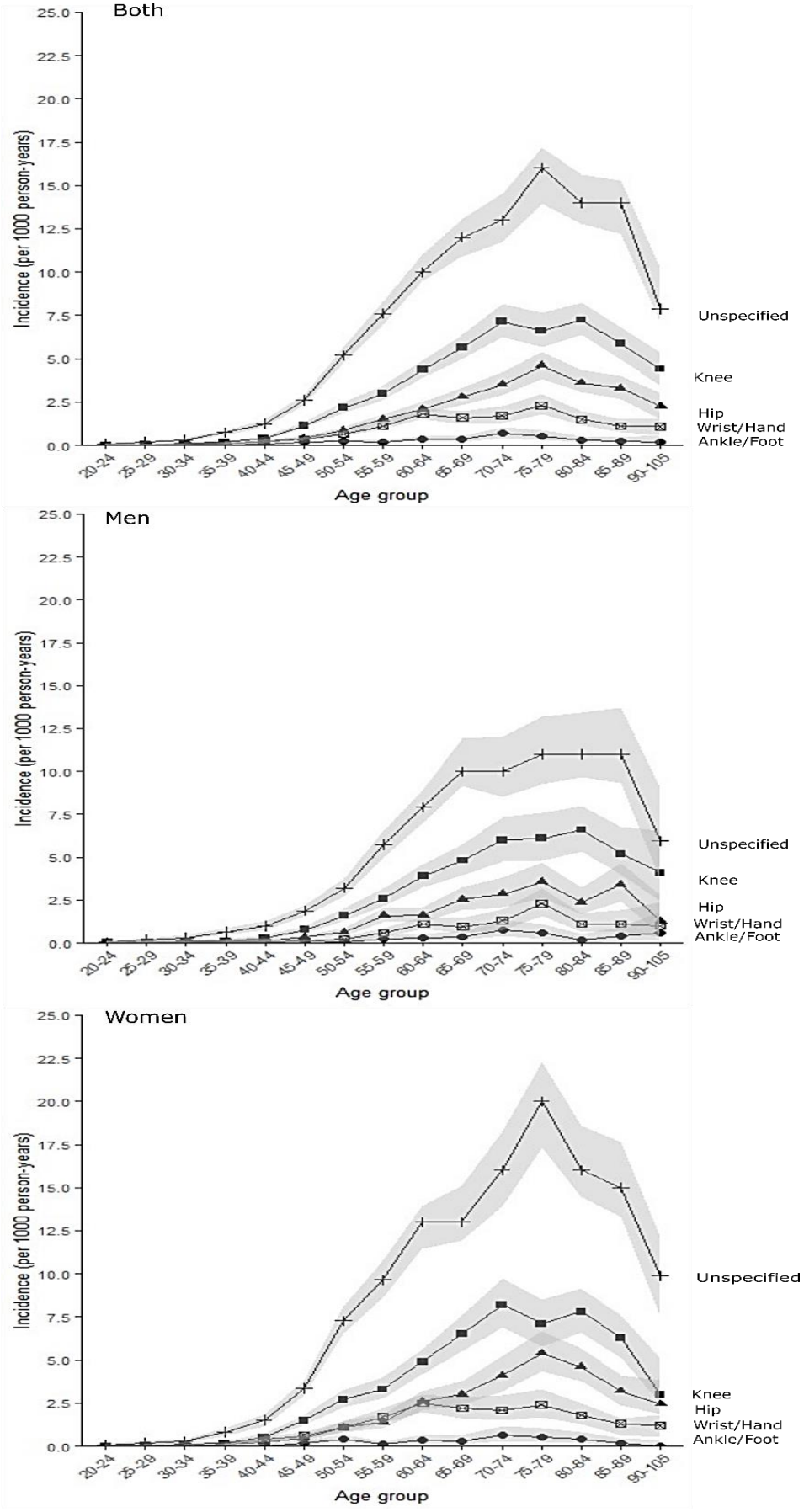
Figure 3.3-9 Incidence of GP diagnosed OA in 2017 in different age groups



The joint-specific incidence rate (per 1000 person-years) in 2017 was highest for unspecified OA (5.20; 95% CI 5.06-5.30), then knee OA (2.30; 95% CI 2.22-2.37), hip OA (1.15; 95% CI

1.10-1.20), wrist and hand OA (0.65; 95% CI 0.62-0.73) and ankle and foot OA (0.19; 95% CI 0.17-0.21). All the joint specific incidence rates were higher in women than in men. In men, incidence rates increased until the age group 70-74 years and in women the rate increased until 75-79 years after which it declined. Ankle and foot, and knee OA rates increased until 70-74 years in both the groups. (Figure 3.3-10)

Figure 3.3-10 Incidence of GP diagnosed OA according to joint in 2017



3.3.1.5 Trends in incidence of GP diagnosed OA in the UK (1997-2017)

Table 3.3-4 shows the temporal trends in incidence of any OA from 1997 to 2017. The standardised rates were slightly higher than crude rates.

Table 3.3-4 Incidence of GP diagnosed OA in the UK

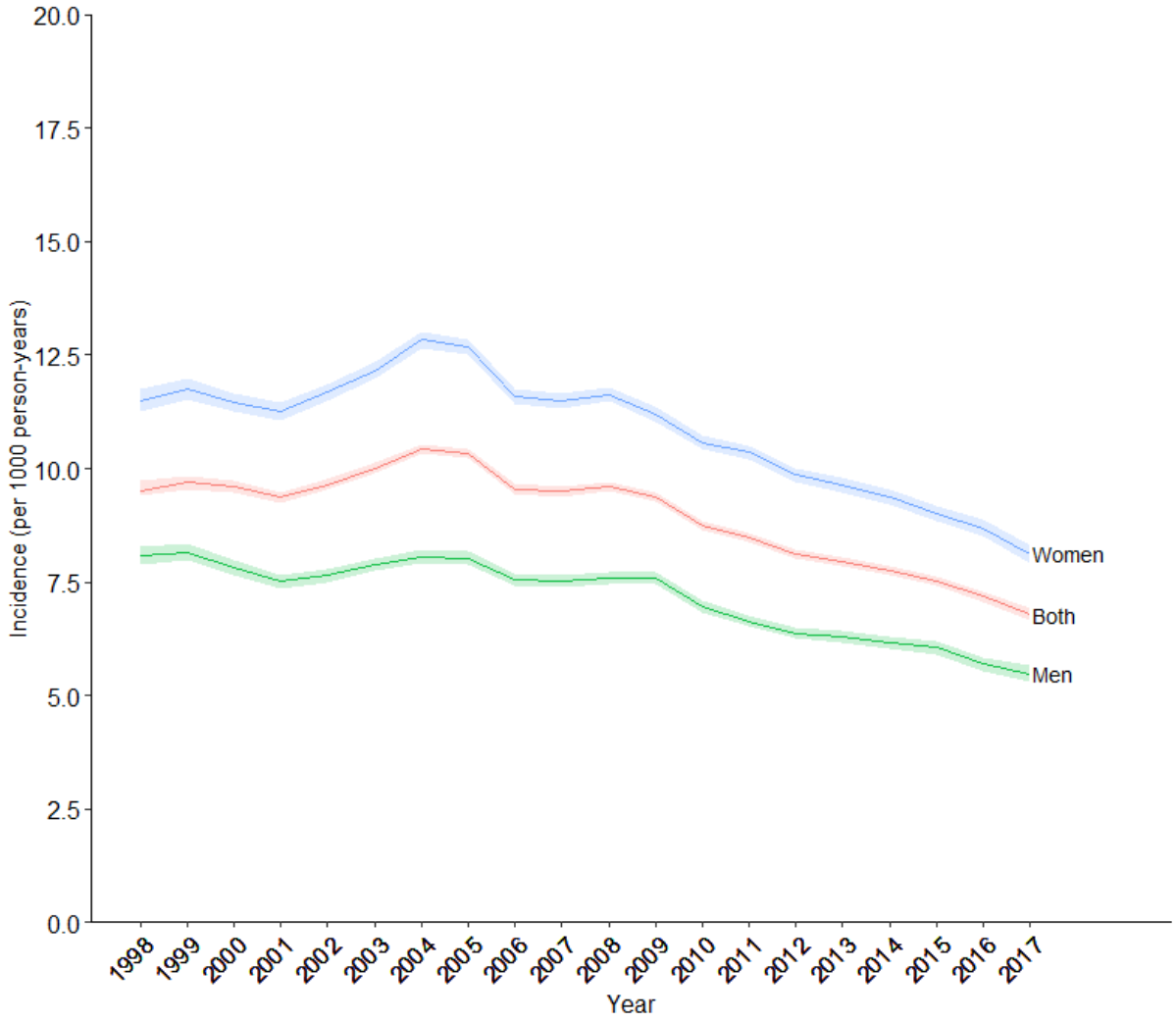
			Incidence (1000 person-years)		
Years	Person-Year	Cases	Crude [95% CI]	Age-sex standardized [95% CI]	Age-sex-LOD standardized [95% CI]
1997	1321487	12296	9.30 [9.14-9.47]	9.17 [9.00-9.34]	
1998	1509159	14817	9.81 [9.66-9.97]	9.05 [8.89-9.20]	9.50 [9.40-9.70]
1999	1831971	17216	9.39 [9.26-9.54]	8.87 [8.73-9.01]	9.69 [9.00-10.37]
2000	2262732	20599	9.10 [8.98-9.22]	8.97 [8.84-9.11]	9.61 [9.31-9.92]
2001	2534401	23615	9.31 [9.19-9.43]	9.20 [9.07-9.32]	9.36 [9.15-9.57]
2002	2858237	26597	9.30 [9.19-9.41]	9.37 [9.25-9.49]	9.64 [9.44-9.84]
2003	3046692	29358	9.63 [9.52-9.74]	9.63 [9.51-9.74]	10.00 [9.81-10.19]
2004	3247175	32543	10.02 [9.91-10.13]	10.06 [9.95-10.17]	10.42 [10.23-10.61]
2005	3317484	33093	9.97 [9.86-10.08]	10.15 [10.04-10.26]	10.33 [10.15-10.52]
2006	3346598	30840	9.21 [9.11-9.31]	9.39 [9.29-9.50]	9.55 [9.37-9.72]
2007	3374993	30236	8.95 [8.88-9.06]	9.15 [9.04-9.25]	9.49 [9.32-9.65]
2008	3381824	30261	8.94 [8.84-9.05]	9.20 [9.10-9.30]	9.59 [9.44-9.74]
2009	3362701	29387	8.73 [8.63-8.83]	8.99 [8.89-9.10]	9.36 [9.22-9.50]
2010	3314620	27133	8.18 [8.09-8.28]	8.42 [8.32-8.52]	8.74 [8.62-8.87]
2011	3235505	26100	8.06 [7.96-8.16]	8.30 [8.20-8.40]	8.48 [8.36-8.59]
2012	3196392	24727	7.73 [7.64-7.83]	7.95 [7.85-8.05]	8.10 [7.90-8.30]
2013	3030317	23409	7.72 [7.62-7.82]	7.87 [7.77-7.97]	7.94 [7.84-8.05]
2014	2758065	21113	7.65 [7.55-7.75]	7.74 [7.64-7.85]	7.75 [7.65-7.86]
2015	2360852	17690	7.49 [7.38-7.60]	7.52 [7.41-7.63]	7.51 [7.40-7.62]
2016	1889587	13540	7.16 [7.04-7.28]	7.18 [7.06-7.30]	7.17 [7.05-7.29]
2017	1495497	10146	6.78 [6.67-6.93]	6.78 [6.67-6.93]	6.78 [6.67-6.93]

Age-sex and length of data contribution (LOD) standardization was done using 2017 CPRD population as standard population. For 1997, LOD standardisation was not calculated because of absence of data for >=10 years.

In general, both crude and standardised estimates decreased over time during this period, changing from 9.5 per 1000 person-years [95% CI 9.4-9.7] to 6.8 per 1000 person-years [95% CI 6.7-6.9]. A similar trend was seen in both men and women. Figure 3.3-11 shows age and LOD standardised rates in men and women. The incidence of OA in men declined from 8.0 per 1000 person-years (95% CI 7.8 to 8.3 per 1000 person-years) in 1997 to 5.5 per 1000

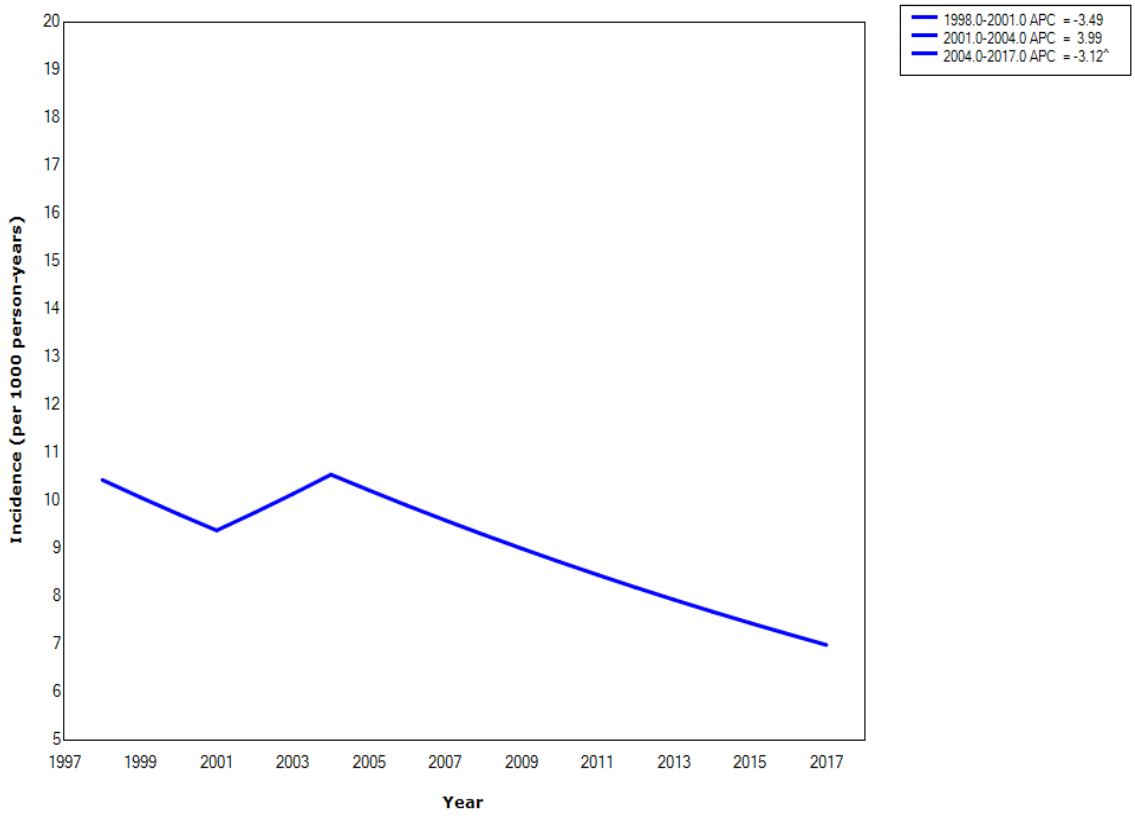
person-years (95% CI 5.3 to 5.7 per 1000 person-years) in 2017, whereas in women the incidence reduced from 11.5 per 1000 person-years (95% CI 11.2 to 11.7 per 1000 person-years) to 8.1 per 1000 person-years (95%CI 7.9 to 8.3 per 1000 person-years)

Figure 3.3-11 Trends in age-standardised incidence of GP diagnosed OA in the UK (1997-2017)



Joinpoint analysis identified two points of changes in overall trend in 2002 and 2005. The AAPC was -1.6% (95% CI -2.0 to -1.1%), indicating a slight decline in the incidence since 1998. Women (-1.9%; 95% CI -2.2 to -1.6%) had a higher decline in rates compared to men (-1.5%; -1.1 to -1.9%). (Figure 3.3-12)

Figure 3.3-12 Joinpoint analysis of trend in incidence of GP diagnosed OA



Trends in joint specific OA incidence are shown in Figure 3.3-13. No change in trend was observed for ankle and foot and wrist and hand sites. Unspecified OA showed a declining trend, whereas OA at the knee and hip showed slightly increasing trends. The incidence of unspecified OA reduced from 9.0 per 1000 person-years [95% CI 8.8 - 9.1 per 1000 person-years] in 1998 to 5.2 per 1000 person -years [95% CI 5.1-5.3 per 1000 person-years] in 2017. A similar trend was seen in men and women. (Table 3.3-5 and Table 3.3-6)

Figure 3.3-13 Trends in incidence of GP-diagnosed OA by joint in the UK (1997-2017)

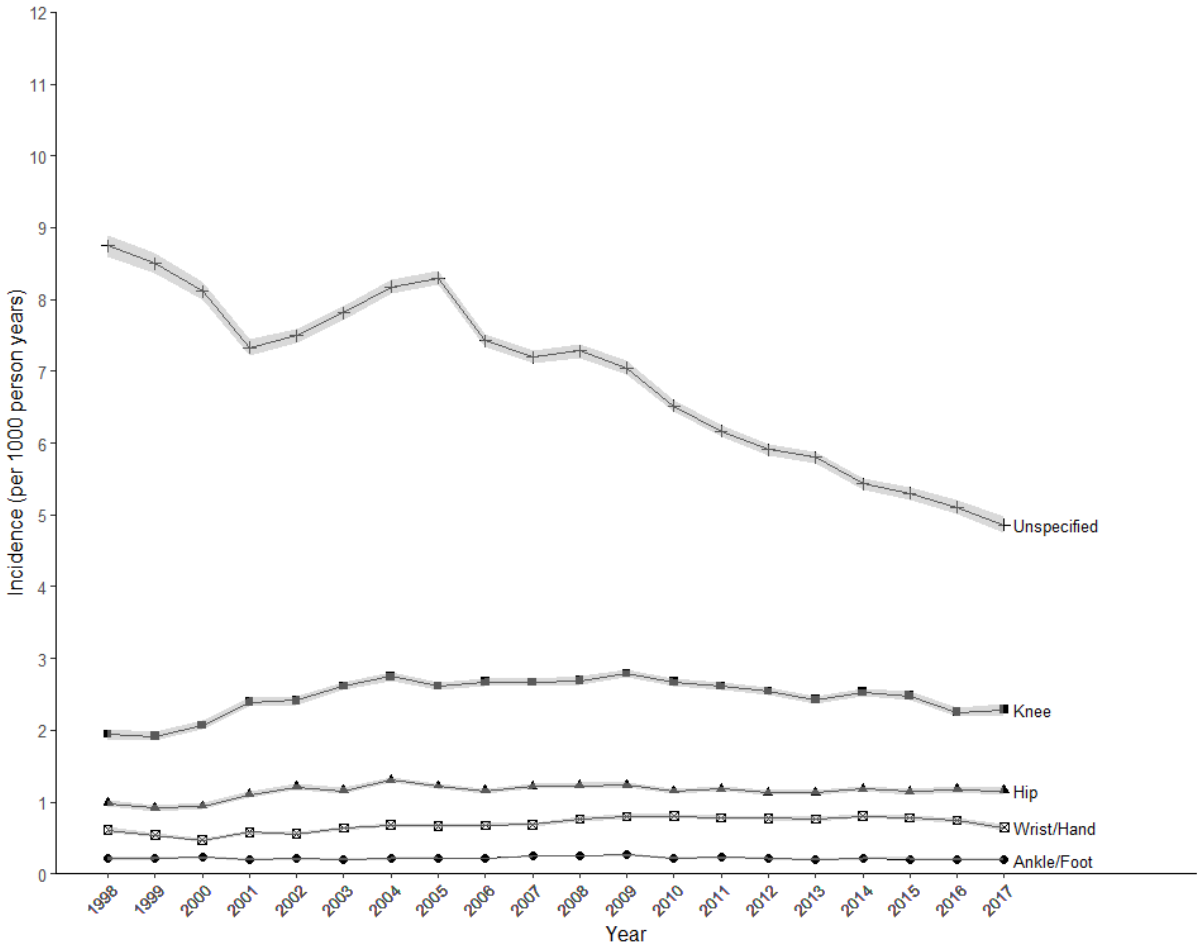


Table 3.3-5 Trends in incidence GP diagnosed OA by joint in men

	Standardised Incidence (per 1000 person-years) [95% CI]				
Years	Hip	Knee	Wrist/Hand	Ankle/Foot	Unspecified
1998	0.86 [0.74-0.98]	1.69 [1.26-2.02]	0.44 [0.34-0.54]	0.25 [0.19-0.28]	6.72 [3.73-9.71]
1999	0.74 [0.48-0.99]	1.70 [1.31-2.09]	0.35 [0.29-0.41]	0.29 [0.12-0.45]	7.22 [6.42-8.03]
2000	0.71 [0.60-0.83]	1.94 [1.74-2.13]	0.34 [0.27-0.42]	0.23 [0.16-0.29]	6.29 [5.94-6.65]
2001	0.96 [0.86-1.06]	2.00 [1.87-2.14]	0.32 [0.22-0.42]	0.15 [0.11-0.19]	5.64 [5.41-5.88]
2002	1.03 [0.93-1.12]	2.12 [1.98-2.25]	0.36 [0.30-0.42]	0.17 [0.13-0.21]	5.77 [5.55-5.99]
2003	0.98 [0.89-1.07]	2.26 [2.13-2.39]	0.36 [0.30-0.42]	0.19 [0.15-0.23]	5.91 [5.7-6.12]
2004	1.12 [1.04-1.21]	2.44 [2.31-2.57]	0.40 [0.32-0.49]	0.18 [0.15-0.22]	6.02 [5.82-6.23]
2005	1.00 [0.92-1.08]	2.21 [2.09-2.33]	0.40 [0.31-0.49]	0.16 [0.12-0.20]	6.18 [5.98-6.38]
2006	0.98 [0.90-1.06]	2.34 [2.22-2.47]	0.43 [0.39-0.46]	0.19 [0.15-0.24]	5.53 [5.33-5.73]
2007	1.02 [0.96-1.08]	2.30 [2.19-2.42]	0.41 [0.37-0.44]	0.24 [0.20-0.28]	5.43 [5.33-5.53]
2008	1.05 [1.00-1.10]	2.36 [2.25-2.46]	0.41 [0.38-0.45]	0.21 [0.17-0.25]	5.42 [5.33-5.51]
2009	1.0 [0.93-1.07]	2.42 [2.32-2.52]	0.44 [0.40-0.47]	0.24 [0.20-0.28]	5.41 [5.30-5.52]
2010	0.96 [0.90-1.02]	2.31 [2.22-2.40]	0.49 [0.46-0.54]	0.21 [0.18-0.24]	4.93 [4.85-5.02]
2011	0.97 [0.91-1.03]	2.24 [2.16-2.32]	0.48 [0.45-0.52]	0.22 [0.19-0.25]	4.62 [4.56-4.68]
2012	0.92 [0.87-0.97]	2.17 [2.09-2.25]	0.47 [0.43-0.51]	0.19 [0.16-0.22]	4.37 [4.29-4.45]
2013	0.94 [0.89-0.99]	2.04 [1.97-2.12]	0.46 [0.42-0.50]	0.18 [0.15-0.21]	4.33 [4.20-4.46]
2014	0.98 [0.91-1.05]	2.17 [2.09-2.25]	0.46 [0.42-0.50]	0.20 [0.14-0.26]	4.02 [3.92-4.12]
2015	0.97 [0.91-1.04]	2.19 [2.10-2.27]	0.49 [0.45-0.54]	0.19 [0.13-0.25]	3.94 [3.83-4.05]
2016	0.96 [0.90-1.02]	1.98 [1.89-2.06]	0.50 [0.45-0.55]	0.19 [0.13-0.25]	3.77 [3.63-3.90]
2017	0.93 [0.86-0.99]	1.93 [1.83-2.02]	0.42 [0.37-0.47]	0.17 [0.14-0.20]	3.71 [3.57-3.84]
AAPC	1.2 [0.0 to 2.3]*	1.1 [0.3 to 1.8]*	0.4 [-1.6 to 2.3]	-2.5 [-5.1 to 0.1]	-3.2 [-4.0 to -2.4]*

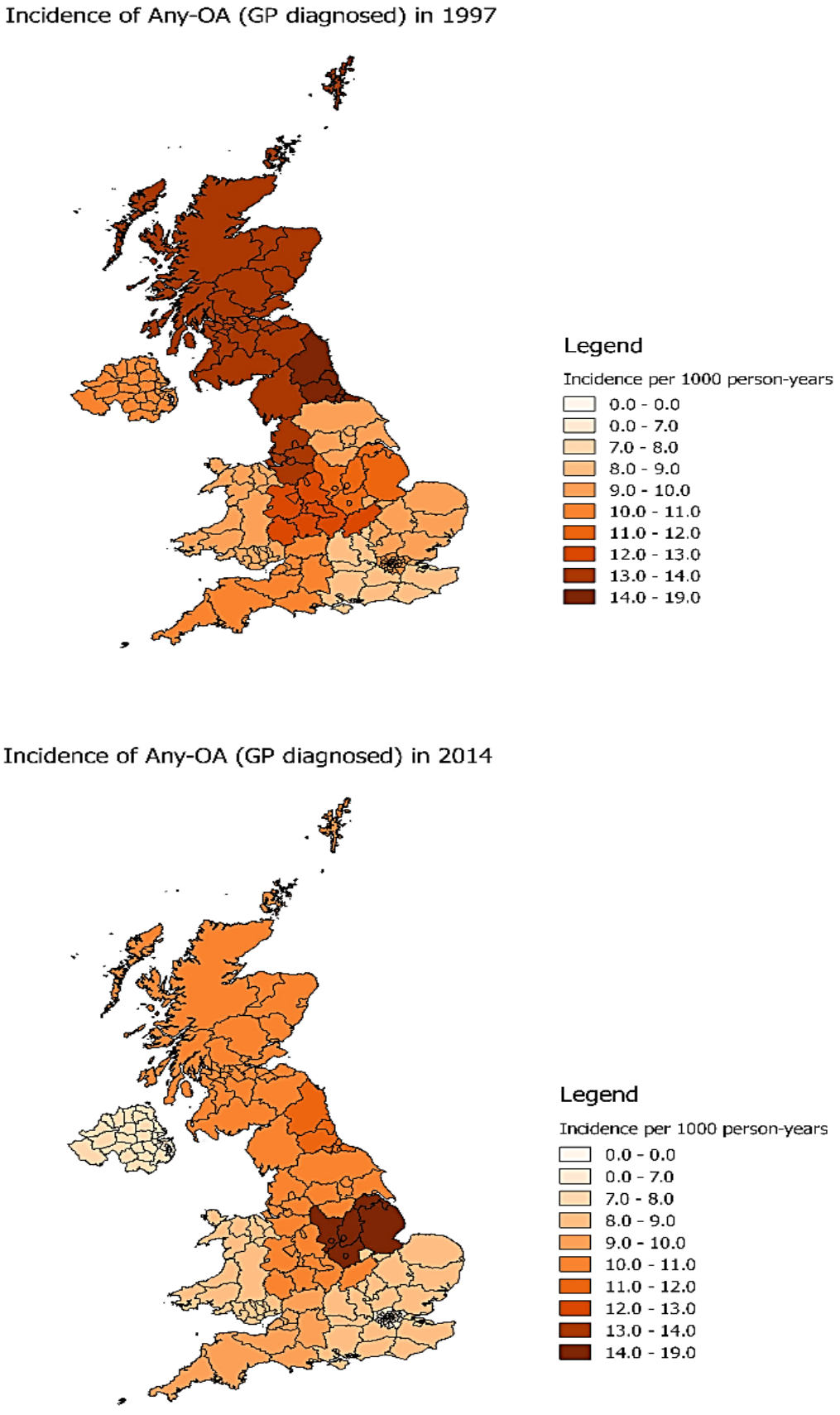
AAPC: Average annual percentage change; CI- Confidence interval; Age and length of data standardized using 2017 CPRD population; **P*-value <0.05

Table 3.3-6 Trend in incidence of GP diagnosed OA by joint in women

	Standardised Incidence (1000 person-years) [95% CI]				
Years	Hip	Knee	Wrist/Hand	Ankle/Foot	Unspecified
1998	1.20[1.12-1.27]	2.00 [1.89-2.11]	0.89 [0.80-0.98]	0.18 [0.02-0.99]	10.76 [6.91-14.62]
1999	1.08 [0.79-1.38]	2.12 [1.70-2.53]	0.85 [0.77-0.93]	0.15 [0.041-0.26]	12.12 [11.11-13.12]
2000	1.15 [1.01-1.30]	2.19 [1.98-2.40]	0.73 [0.67-0.79]	0.22 [0.16-0.29]	9.93 [9.49-10.37]
2001	1.24 [1.13-1.34]	2.77 [2.61-2.92]	0.61 [0.55-0.67]	0.23 [0.19-0.28]	9.02 [8.73-9.31]
2002	1.38 [1.28-1.48]	2.70 [2.55-2.84]	0.78 [0.71-0.85]	0.24 [0.19-0.28]	9.22 [8.95-9.49]
2003	1.32 [1.23-1.41]	2.95 [2.81-3.09]	0.74 [0.68-0.80]	0.21 [0.17-0.24]	9.72 [9.46-9.98]
2004	1.47 [1.37-1.57]	3.04 [2.90-3.19]	0.86 [0.80-0.92]	0.24 [0.20-0.28]	10.33 [10.07-10.59]
2005	1.44 [1.34-1.53]	3.00 [2.86-3.14]	0.94 [0.89-0.99]	0.25 [0.21-0.29]	10.42 [10.15-10.69]
2006	1.32 [1.24-1.41]	2.99 [2.85-3.12]	0.89 [0.85-0.94]	0.22 [0.18-0.26]	9.31 [9.07-9.55]
2007	1.41 [1.33-1.49]	3.01 [2.89-3.14]	0.92 [0.87-0.97]	0.24 [0.21-0.28]	8.96 [8.74-9.18]
2008	1.39 [1.31-1.47]	3.01 [2.90-3.13]	0.94 [0.88-0.97]	0.27 [0.23-0.30]	9.14 [8.93-9.35]
2009	1.43 [1.36-1.51]	3.14 [3.03-3.25]	1.06 [1.01-1.11]	0.28 [0.24-0.31]	8.68 [8.49-8.87]
2010	1.33 [1.27-1.40]	3.02 [2.92-3.12]	1.08 [1.02-1.13]	0.22 [0.19-0.25]	8.11 [7.94-8.28]
2011	1.39 [1.32-1.45]	2.98 [2.88-3.07]	1.09 [1.04-1.15]	0.23 [0.21-0.26]	7.71 [7.55-7.86]
2012	1.31 [1.25-1.38]	2.90 [2.81-2.99]	1.07 [1.01-1.12]	0.22 [0.20-0.25]	7.44 [7.29-7.6]
2013	1.30 [1.25-1.36]	2.80 [2.71-2.88]	1.07 [1.01-1.12]	0.20 [0.18-0.22]	7.26 [7.12-7.40]
2014	1.37 [1.31-1.43]	2.87 [2.78-2.96]	1.04 [0.97-1.10]	0.23 [0.20-0.25]	6.83 [6.7-6.97]
2015	1.31 [1.24-1.37]	2.75 [2.66-2.84]	1.09 [1.02-1.16]	0.22 [0.19-0.24]	6.65 [6.51-6.80]
2016	1.37 [1.30-1.44]	2.51 [2.42-2.61]	1.04 [0.97-1.11]	0.21 [0.19-0.24]	6.44 [6.27-6.60]
2017	1.37 [1.29-1.45]	2.63 [2.52-2.74]	1.04 [0.96-1.12]	0.21 [0.18-0.25]	6.02 [5.84-6.19]
AAPC	1.0 [0.2 to 1.7]*	1.6 [0.8 to 2.3]*	0.8 [-0.5 to 2.0]	3.6 [1.0 to 6.2]*	-3.3 [-5.4 to -1.2]*

AAPC: Average annual percentage change; CI- Confidence interval; Age and length of data standardized using 2017 CPRD population; **P*-value <0.05

Figure 3.3-14. Incidence of GP-diagnosed OA in different regions of the UK in 1997 and 2014



Standardised incidence rates decreased in all regions of the UK since 1997. In 2014 Yorkshire and Humber and the North East had the highest incidence rates of 14-15 per 1000 person-years and 12-13 per 1000 person-years, respectively. Lowest incidence rates were seen in Northern Ireland and South East England. (Figure 3.3-14)

3.3.1.6 Age period cohort analysis

Cohort effects

The incidence was found to decline according to birth cohorts. For people with the same age, those born later were less likely to have OA than those born earlier. (Figure 3.3-15) The reduction speeded up gradually after 1960, particularly for people aged 20-40 years, suggesting a potential aetiological change after 1960 that has made people less likely to develop OA. In contrast, prevalence increased and speeded up gradually by age, but remained almost unchanged for people born after 1960. The plot of distribution of incidence and prevalence across the age group for different periods is provided in Figure 3.3-16.

Figure 3.3-15. Age-period-cohort analysis of trend of OA (1997-2017) incidence (A) and prevalence (B) in the UK.

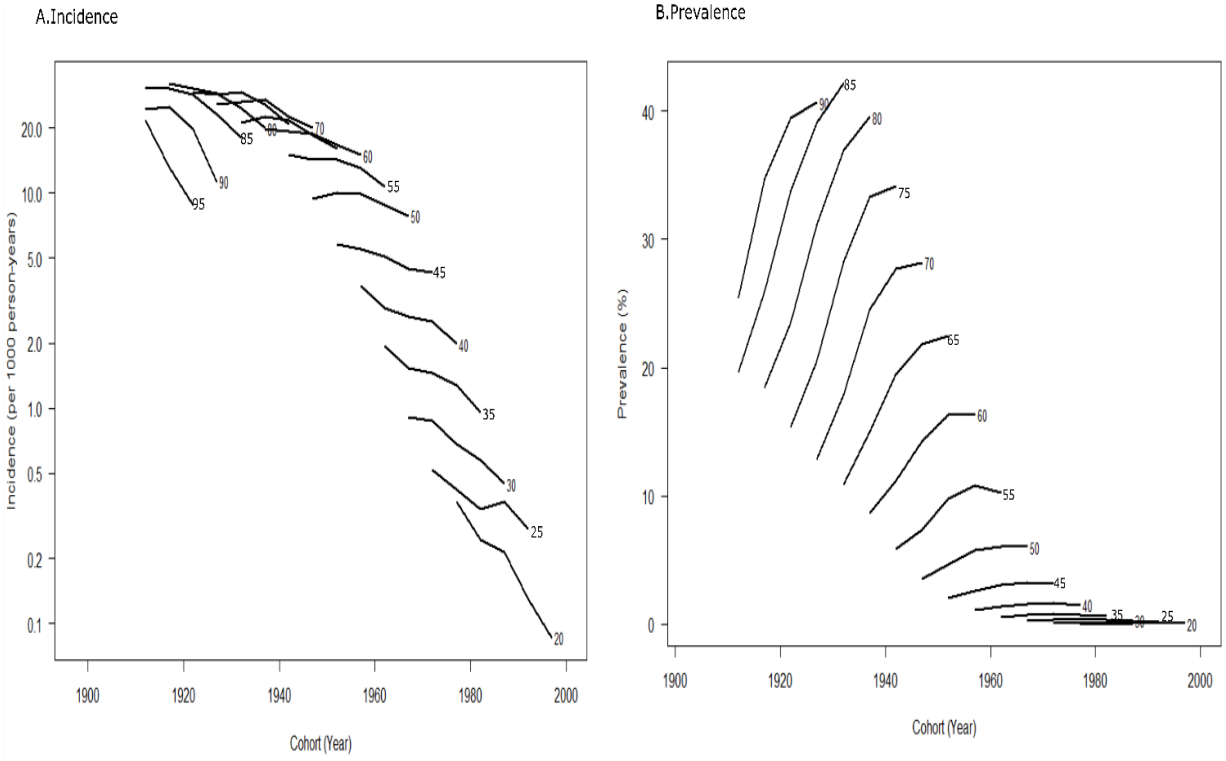
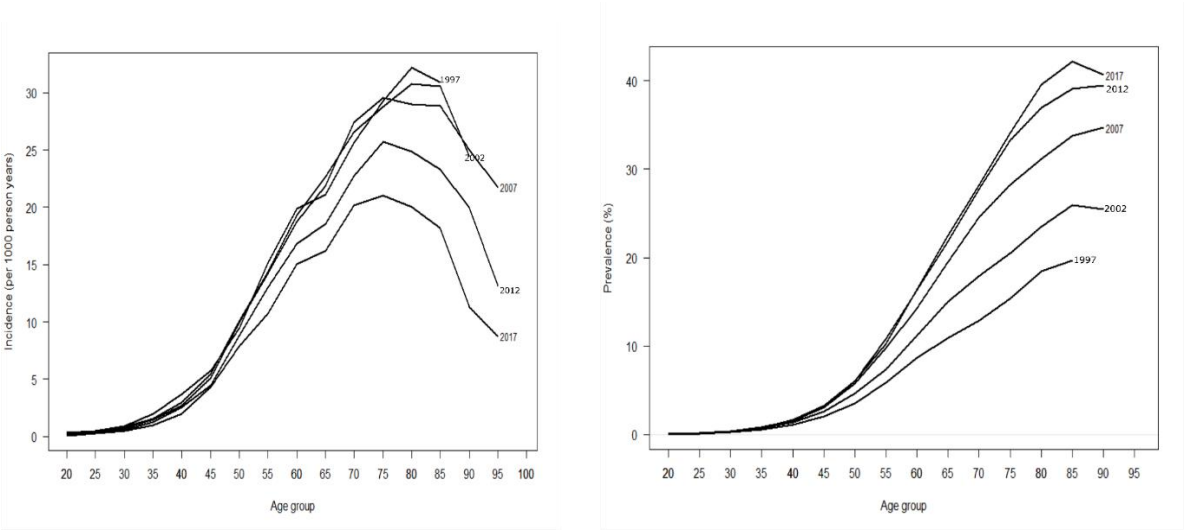


Figure 3.3-16. Incidence and prevalence by age and calendar years

A) Incidence B) Prevalence



Each line represents the distribution of OA across age in the period as stated. The lower incidence in the year 2017 compared to other years in same age group explains possible cohort effects.

Prevalence of any GP diagnosed OA varied between geographic regions. In 2017 the age-sex standardised prevalence in Scotland, West Midlands and Northern Ireland was between 7%-9%. The regions in the South of England had prevalence ranging from 3%-5%. Spatial-temporal maps of the prevalence show a decline in prevalence since 1997 in all UK regions except for Scotland and the West Midlands. The prevalence for East Midlands could not be estimated because of lack of information in the year 2017. (Figure 3.3-8)

3.3.2 Joint pain definition of OA

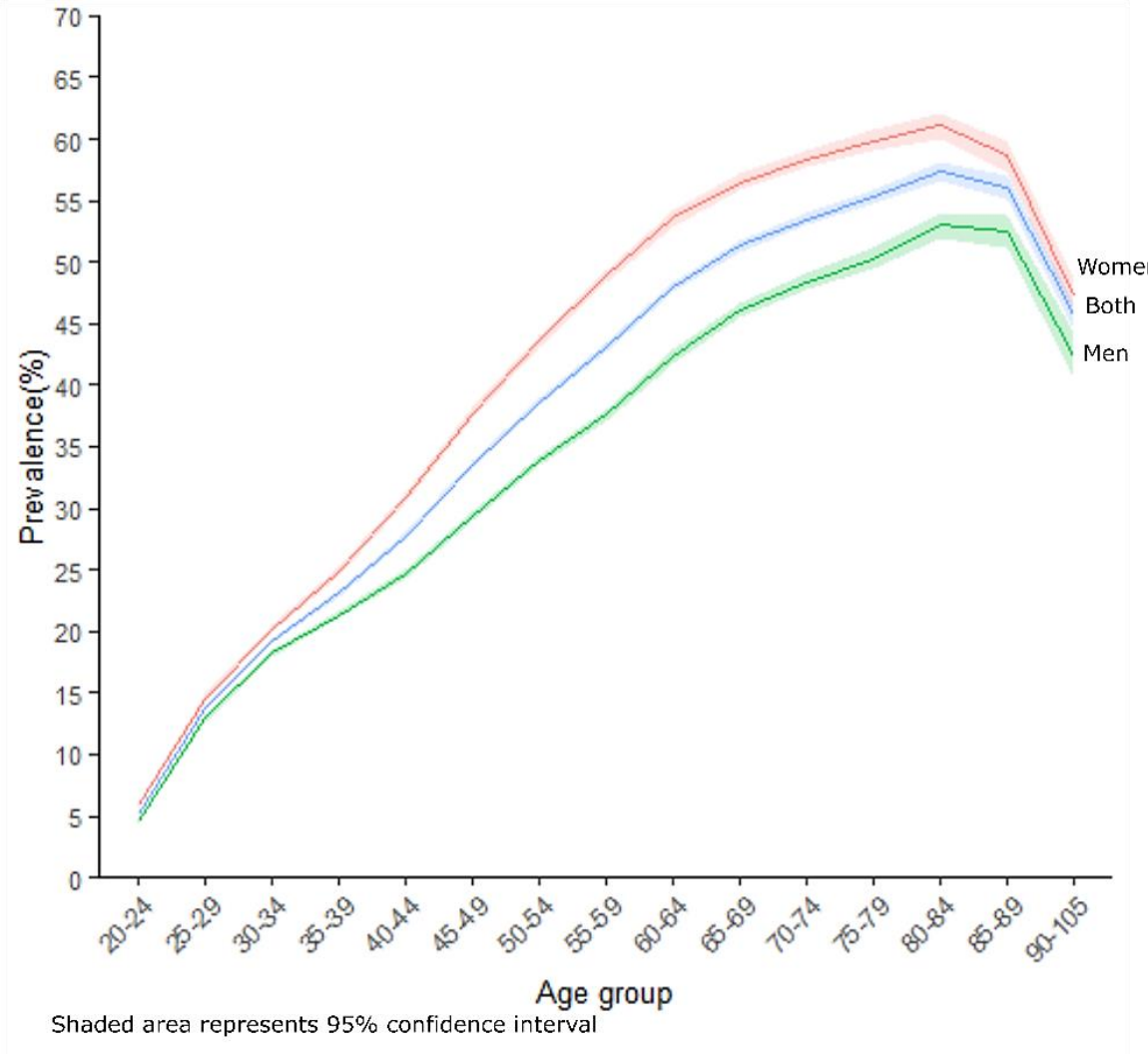
3.3.2.1 Prevalence of joint pain defined OA in 2017

In 2017 out of 1,600,036 eligible patients 544,763 cases of joint pain were recorded. The age-sex-LOD standardised prevalence in 2017 was 34.04% [95% CI 33.98%-34.11%]. (Table 3.3-7) Women (37.7%; 95% CI 37.5%-37.9%) had a greater prevalence of joint pain than men (29.9%; 95% CI 29.8%-30.0%).

Age specific distribution of prevalence for 2017 is given in Figure 3.3-17. In the youngest age group (20 to 24 years), the prevalence of joint pain was about 4% and was the same in men and women. The prevalence increased with age until the age group of 80 to 84 years, where it attained the peak of 57.4% [95% CI 56.6%-58.1%]. The gap between men and women widened after the age of 35-39 years which again became narrower in those over 85 years.

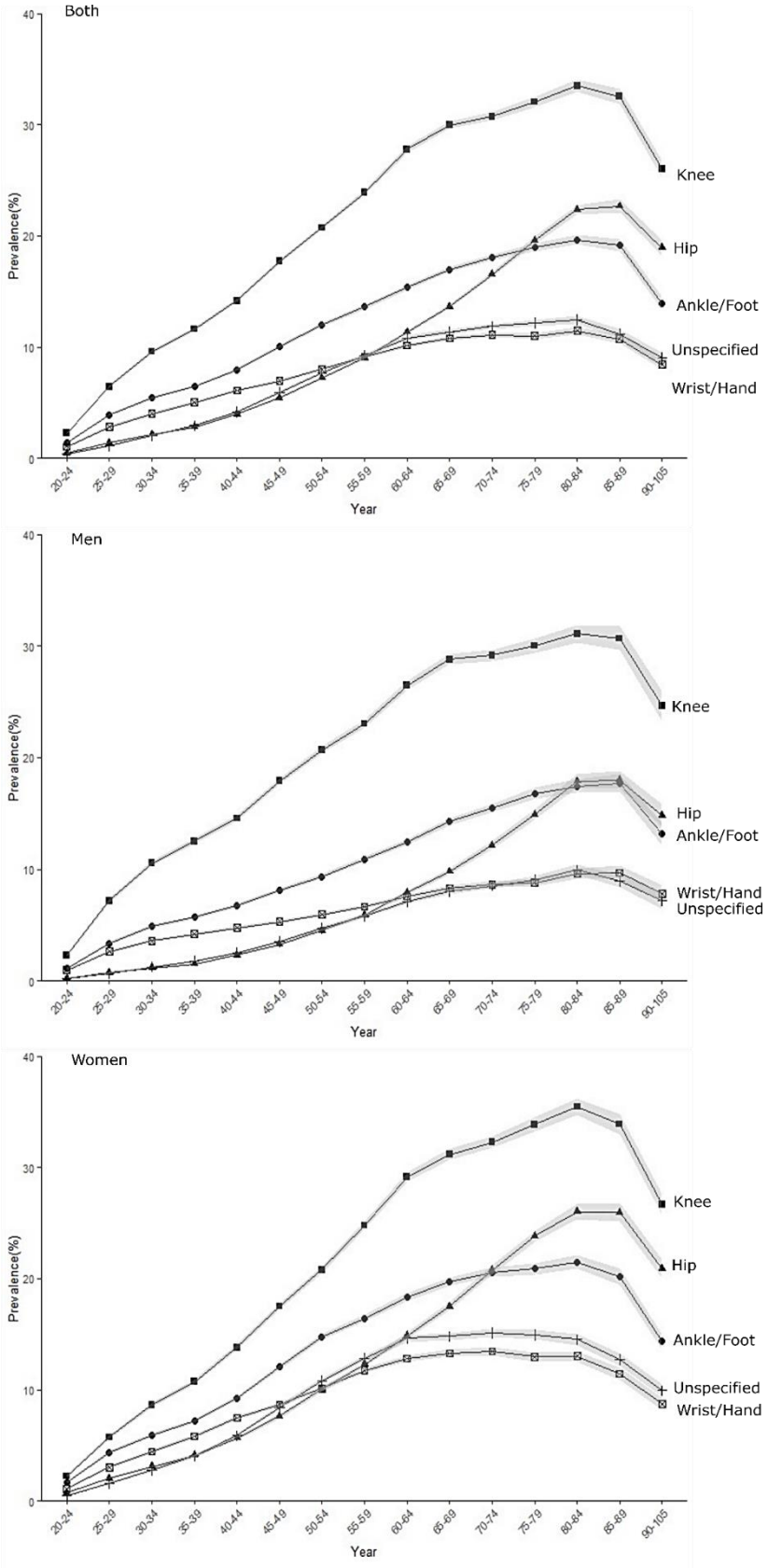
Site specific joint pain prevalence according to age group is shown in Figure 3.3-18. The prevalence increased with age at all joints. In 2017, knee joint pain had the highest prevalence (18.6%; 95%CI 18.5%-18.7%) followed by ankle and foot (10.7%; 95%CI 10.6%-10.7%), hip, (7.9%; 95% CI 7.8%-7.9%) wrist and hand (7.1%; 95% CI 7.0%-7.1%) and unspecified (6.5%; 95% CI 6.4%-6.5%).

Figure 3.3-17. Prevalence of joint pain defined OA by age in 2017



In the younger age group of 25-29 years the leading sites for prevalence were knee (6.4%; 95% CI 6.3%-6.6%), ankle and foot (3.9%; 95% CI 3.7%-3.9%), and wrist and hand (2.8%; 95%CI 2.7%-2.9%). In the older age group of 80-84 years, the order was knee (33.5%; 95%CI 32.9%-34.1%), hip (22.3%, 95%CI 21.9%-22.8%), ankle and foot (19.7%; 19.2%-20.1%), and unspecified (12.4%; 95% CI 12.1%-12.8%). A similar pattern occurred in both sexes. (Figure 3.3-18)

Figure 3.3-18. Prevalence of joint pain defined OA according to age and sex in 2017



3.3.2.2 Trends of prevalence of joint pain defined OA

Table 3.3-7 Prevalence of joint pain defined OA in the UK

			Prevalence (%)		
Years	N (Total eligible population)	Cases	Crude	Age-sex standardised	Age-sex and LOD standardised
1997	5255368	278466	5.29[5.22-5.35]	6.79[6.71-6.86]	8.99[8.91-9.06]
1998	5329284	313405	5.88[5.82-5.94]	7.46[7.32-7.52]	9.66[9.52-9.72]
1999	5400350	347664	6.43[6.40-6.46]	8.10[8.00-8.12]	10.3[10.2-10.32]
2000	5454259	383134	7.02[6.98-7.06]	8.93[8.91-8.96]	11.43[11.41-11.46]
2001	5465735	421949	7.71[7.65-7.77]	9.74[9.71-9.77]	12.04[12.01-12.07]
2002	5437891	467109	8.58[8.50-8.66]	10.72[10.69-10.75]	13.02[12.99-13.05]
2003	5374918	528317	9.82[9.78-9.86]	12.12[12.09-12.15]	14.32[14.29-14.35]
2004	5303535	599496	11.30[11.22-11.38]	13.76[13.73-13.80]	15.86[15.83-15.9]
2005	5225335	677041	12.95[12.90-13.00]	15.58[15.55-15.62]	17.58[17.55-17.62]
2006	5093687	745604	14.63[14.59-14.67]	17.36[17.32-17.39]	19.16[19.12-19.19]
2007	4938346	809153	16.38[16.32-16.44]	19.11[19.12-19.19]	20.61[20.62-20.69]
2008	4774617	875077	18.32[18.27-18.37]	21.09[21.06-21.13]	22.49[22.46-22.53]
2009	4604954	938701	20.38[20.32-20.44]	23.06[23.02-23.10]	24.36[24.32-24.4]
2010	4391276	980286	22.32[22.28-22.36]	24.87[24.83-24.91]	25.77[25.73-25.81]
2011	4145387	1010886	24.38[24.31-24.45]	26.69[26.64-26.73]	27.49[27.44-27.53]
2012	3909979	1029322	26.32[26.26-26.38]	28.33[28.29-28.38]	29.03[28.99-29.08]
2013	3582506	1011137	28.22[28.15-28.29]	29.85[29.80-29.89]	30.25[30.2-30.29]
2014	3118698	938341	30.08[30.00-30.16]	31.28[31.23-31.33]	31.48[31.43-31.53]
2015	2600729	823024	31.64[31.56-31.72]	32.41[32.35-32.46]	32.51[32.45-32.56]
2016	1980645	646372	32.63[32.57-32.69]	32.99[32.92-33.05]	33.04[32.97-33.1]
2017	1600036	544763	34.04[33.97-34.11]	34.04[33.98-34.11]	34.04[33.98-34.11]

Age-sex standardized rates are standardized with mid-2017 UK population as standard population.

The temporal crude and standardised prevalence values have increased since 1997. The standardised prevalence increased nearly 3.8 times from 9.0% in 1997 to 34% in 2017. A similar trend was seen in both men (from 8.1% in 1997 to 30.4% in 2017) and women (from 9.8% in 1997 to 37.7% in 2017). The difference in prevalence between men and women in 1997 was 1.7%, which increased to 7.3% in 2017. (Figure 3.3-19)

The joinpoint model shows three change points in prevalence trend at years 2000, 2007 and 2012. The rate of change during 2002-2007 was highest followed by that in 1997-2002. The APC was 8.5% [95% CI 8.2-8.7]. Both men and women showed a similar pattern. (Figure 3.3-20)

Figure 3.3-19. Trends of standardised prevalence of joint pain defined OA in the UK (1997-2017)

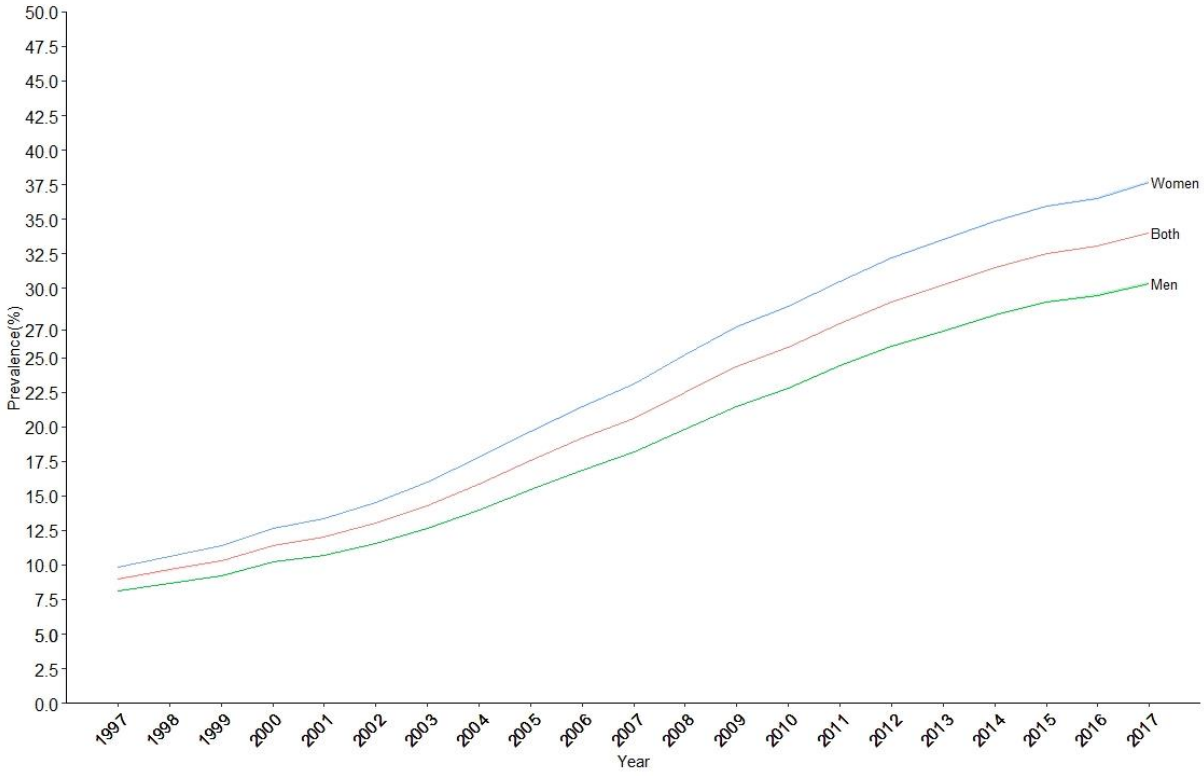
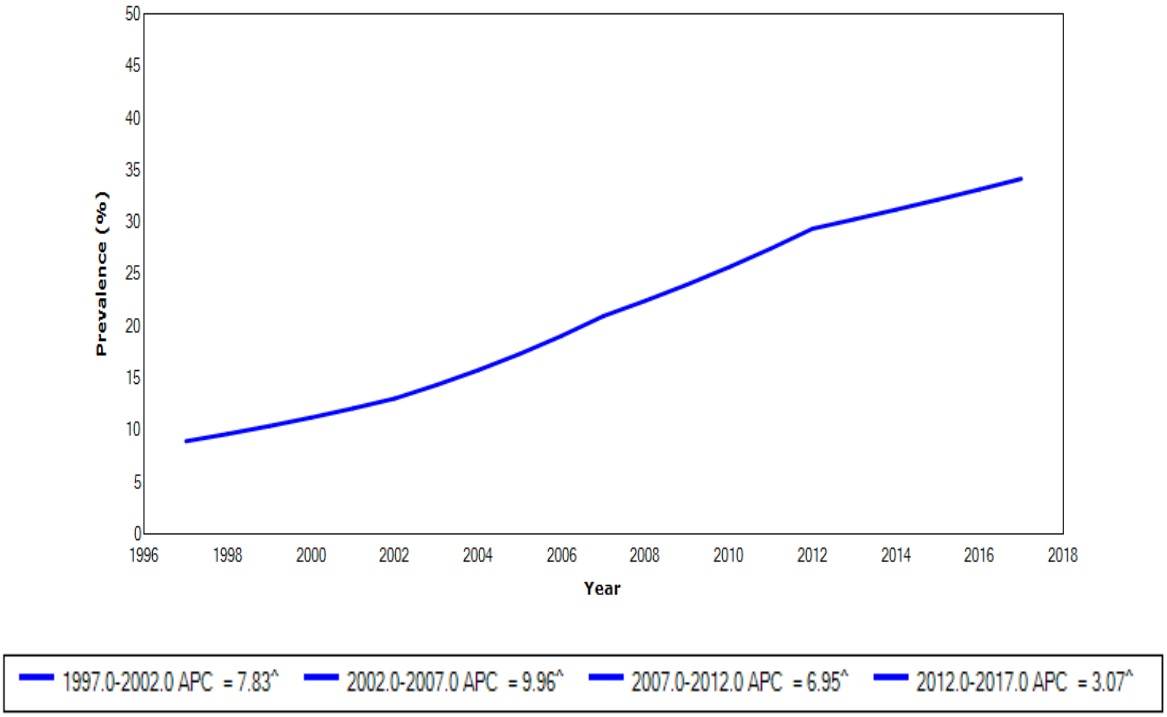


Figure 3.3-20. Joinpoint analysis of trend of Joint pain OA for any site



Joint specific analysis (Figure 3.3-21) shows an increasing trend for all sites, except for unspecified. The trend for knee pain grew most rapidly followed by ankle/foot, hip and wrist and hand pain. The prevalence of knee pain increased by 16% from 2.8% [95% CI 2.7-2.8%] in 1997 to 18.6% [95% CI 18.5%-18.7%] in 2017. Similarly, hip pain prevalence changed from 1.5% in 1997 to 7.9% in 2017 and ankle/foot pain increased from 1.3% to 10.7% in 2017. The trend shows an increase in APC for all regions. The highest increase was observed for wrist/hand pain (11.4%; 95% CI 11.0-11.9) and the lowest was reported for unspecified pain (7.9%; 95% CI 7.5-8.4). Details for region specific pain in men and women are given in Table 3.3-8 and Table 3.3-9.

Figure 3.3-21. Trends of joint pain defined OA prevalence for different sites (1997-2017)

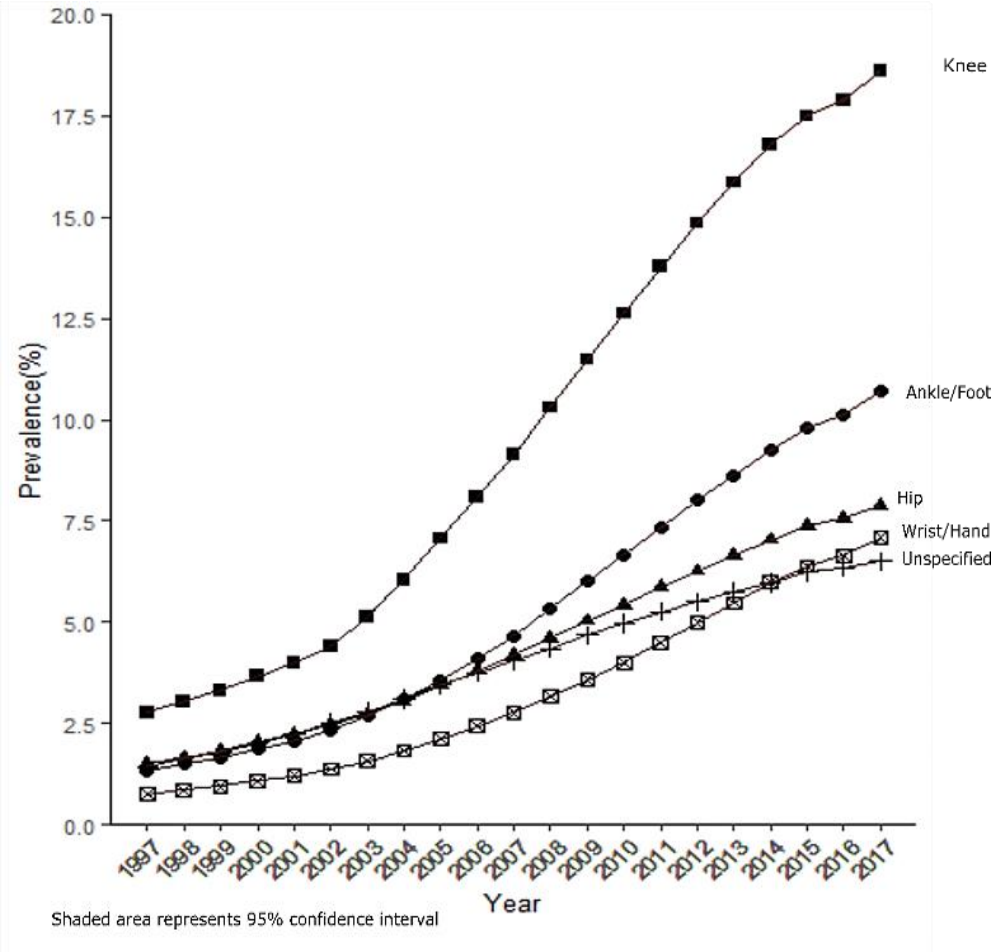


Table 3.3-8 Trend in prevalence of joint-pain defined OA for different sites in men

	Standardised Prevalence (%) [95% CI]				
Year	Hip	Knee	Wrist/Hand	Ankle/Foot	Unspecified
1997	1.06[1.04-1.08]	2.87[2.85-2.89]	0.56[0.54-0.58]	1.09[1.06-1.12]	0.95[0.93-0.97]
1998	1.17[1.15-1.19]	3.15[3.12-3.18]	0.63[0.60-0.66]	1.24[1.22-1.26]	1.08[1.06-1.10]
1999	1.27[1.25-1.29]	3.42[3.40-3.44]	0.71[0.68-0.74]	1.36[1.34-1.38]	1.19[1.17-1.21]
2000	1.40[1.38-1.42]	3.72[3.70-3.75]	0.80[0.79-0.81]	1.52[1.50-1.53]	1.33[1.31-1.35]
2001	1.52[1.51-1.54]	4.04[4.02-4.07]	0.90[0.89-0.92]	1.68[1.66-1.70]	1.49[1.47-1.51]
2002	1.66[1.65-1.68]	4.43[4.40-4.46]	1.02[1.01-1.04]	1.88[1.86-1.90]	1.70[1.68-1.71]
2003	1.84[1.82-1.86]	5.13[5.10-5.16]	1.18[1.16-1.19]	2.16[2.14-2.18]	1.90[1.88-1.92]
2004	2.05[2.03-2.08]	5.98[5.94-6.01]	1.37[1.35-1.38]	2.49[2.46-2.51]	2.10[2.08-2.12]
2005	2.30[2.28-2.32]	6.95[6.91-6.98]	1.60[1.58-1.62]	2.88[2.85-2.90]	2.33[2.31-2.35]
2006	2.55[2.53-2.58]	7.91[7.88-7.95]	1.84[1.82-1.86]	3.30[3.27-3.32]	2.52[2.50-2.54]
2007	2.80[2.78-2.83]	8.91[8.87-8.95]	2.10[2.08-2.12]	3.76[3.74-3.79]	2.72[2.69-2.74]
2008	3.08[3.05-3.10]	10.01[9.97-10.05]	2.39[2.37-2.42]	4.29[4.27-4.32]	2.91[2.88-2.93]
2009	3.38[3.35-3.40]	11.16[11.12-11.20]	2.71[2.69-2.74]	4.86[4.83-4.89]	3.13[3.10-3.15]
2010	3.64[3.62-3.67]	12.28[12.24-12.33]	3.06[3.03-3.08]	5.41[5.38-5.44]	3.32[3.29-3.34]
2011	3.94[3.91-3.97]	13.41[13.37-13.46]	3.46[3.44-3.49]	5.99[5.95-6.02]	3.51[3.49-3.54]
2012	4.19[4.16-4.22]	14.43[14.38-14.48]	3.85[3.82-3.88]	6.53[6.50-6.57]	3.66[3.64-3.69]
2013	4.44[4.40-4.47]	15.39[15.34-15.44]	4.22[4.19-4.25]	7.05[7.01-7.09]	3.81[3.78-3.84]
2014	4.68[4.64-4.71]	16.29[16.24-16.35]	4.63[4.60-4.67]	7.57[7.53-7.61]	3.95[3.92-3.98]
2015	4.89[4.85-4.93]	16.98[16.92-17.05]	4.96[4.92-5.00]	8.03[7.98-8.07]	4.13[4.09-4.16]
2016	5.02[4.97-5.06]	17.37[17.29-17.44]	5.14[5.10-5.19]	8.33[8.27-8.38]	4.19[4.15-4.23]
2017	5.21[5.17-5.26]	18.00[17.91-18.08]	5.47[5.42-5.52]	8.78[8.72-8.85]	4.21[4.17-4.26]
AAPC	8.3[8.0-8.6]*	9.7[9.3-10.0]*	12.1[11.9-12.3]*	11.0[10.7-11.3]*	7.8[7.5-8.0] *

AAPC: Annual average percentage change; *P value significant at 0.05

Table 3.3-9 Trend in prevalence of joint-pain defined OA for different sites in women

	Standardised Prevalence (%) [95% CI]				
Year	Hip	Knee	Wrist/Hand	Ankle/Foot	Unspecified
1997	1.94[1.92-1.96]	2.68[2.65-2.71]	0.95[0.91-0.99]	1.56[1.52-1.60]	1.91[1.88-1.94]
1998	2.14[2.11-2.17]	2.96[2.94-2.98]	1.07[1.05-1.09]	1.77[1.74-1.80]	2.15[2.12-2.18]
1999	2.35[2.32-2.38]	3.24[3.21-3.26]	1.20[1.00-1.40]	1.96[1.92-2.00]	2.38[2.35-2.41]
2000	2.67[2.64-2.69]	3.61[3.59-3.64]	1.35[1.34-1.37]	2.23[2.21-2.25]	2.65[2.62-2.67]
2001	2.92[2.90-2.94]	3.96[3.93-3.99]	1.51[1.49-1.53]	2.47[2.44-2.49]	2.93[2.90-2.95]
2002	3.22[3.20-3.25]	4.38[4.35-4.40]	1.70[1.68-1.72]	2.76[2.74-2.79]	3.28[3.25-3.30]
2003	3.58[3.55-3.61]	5.15[5.12-5.18]	1.95[1.93-1.97]	3.18[3.15-3.20]	3.67[3.64-3.69]
2004	4.01[3.99-4.04]	6.13[6.10-6.16]	2.26[2.24-2.28]	3.68[3.66-3.71]	4.08[4.05-4.10]
2005	4.50[4.47-4.53]	7.19[7.16-7.23]	2.63[2.61-2.65]	4.24[4.21-4.26]	4.54[4.51-4.57]
2006	4.99[4.96-5.02]	8.28[8.24-8.32]	3.01[2.99-3.04]	4.85[4.82-4.88]	4.94[4.91-4.97]
2007	5.49[5.46-5.52]	9.35[9.31-9.39]	3.42[3.40-3.45]	5.53[5.50-5.56]	5.32[5.29-5.35]
2008	6.03[6.00-6.06]	10.57[10.53-10.61]	3.89[3.86-3.92]	6.31[6.27-6.34]	5.74[5.70-5.77]
2009	6.60[6.56-6.63]	11.80[11.75-11.84]	4.40[4.37-4.43]	7.10[7.06-7.14]	6.17[6.14-6.21]
2010	7.13[7.09-7.17]	12.95[12.91-13.00]	4.91[4.88-4.94]	7.86[7.82-7.90]	6.55[6.52-6.59]
2011	7.71[7.67-7.74]	14.12[14.07-14.16]	5.50[5.47-5.54]	8.62[8.58-8.66]	6.92[6.88-6.96]
2012	8.23[8.19-8.27]	15.24[15.19-15.29]	6.07[6.04-6.11]	9.40[9.35-9.44]	7.26[7.22-7.30]
2013	8.74[8.70-8.78]	16.28[16.23-16.34]	6.66[6.62-6.70]	10.12[10.07-10.16]	7.60[7.56-7.64]
2014	9.26[9.21-9.30]	17.26[17.20-17.31]	7.28[7.24-7.32]	10.85[10.80-10.90]	7.91[7.86-7.95]
2015	9.71[9.66-9.76]	17.95[17.89-18.02]	7.75[7.71-7.80]	11.49[11.43-11.54]	8.23[8.18-8.28]
2016	9.96[9.9-10.02]	18.37[18.29-18.44]	8.06[8.01-8.11]	11.86[11.79-11.92]	8.40[8.34-8.45]
2017	10.41[10.35-10.48]	19.17[19.09-19.26]	8.60[8.54-8.66]	12.50[12.43-12.57]	8.67[8.61-8.73]
AAPC	8.8[8.3-9.3]*	10.4[9.9-10.8]*	11.7[11.4-11.9]*	10.9[10.6-11.2]*	7.9[7.6-8.1]*

AAPC: Annual average percentage change; P value <0.05

3.3.2.3 Prevalence of multiple joint-pain defined OA in 2017

According to the joint pain definition, in the year 2017 nearly 34% of the reported cases had at least two joints involved. The prevalence of joint pain at two or more sites in 2017 had increased from the 12% reported in 1997. There was an increasing trend of multiple joints involvement from 1997-2017 in the UK. (Figure 3.3-22)

Figure 3.3-22. Proportion of multiple joint pain between 1997 and 2017

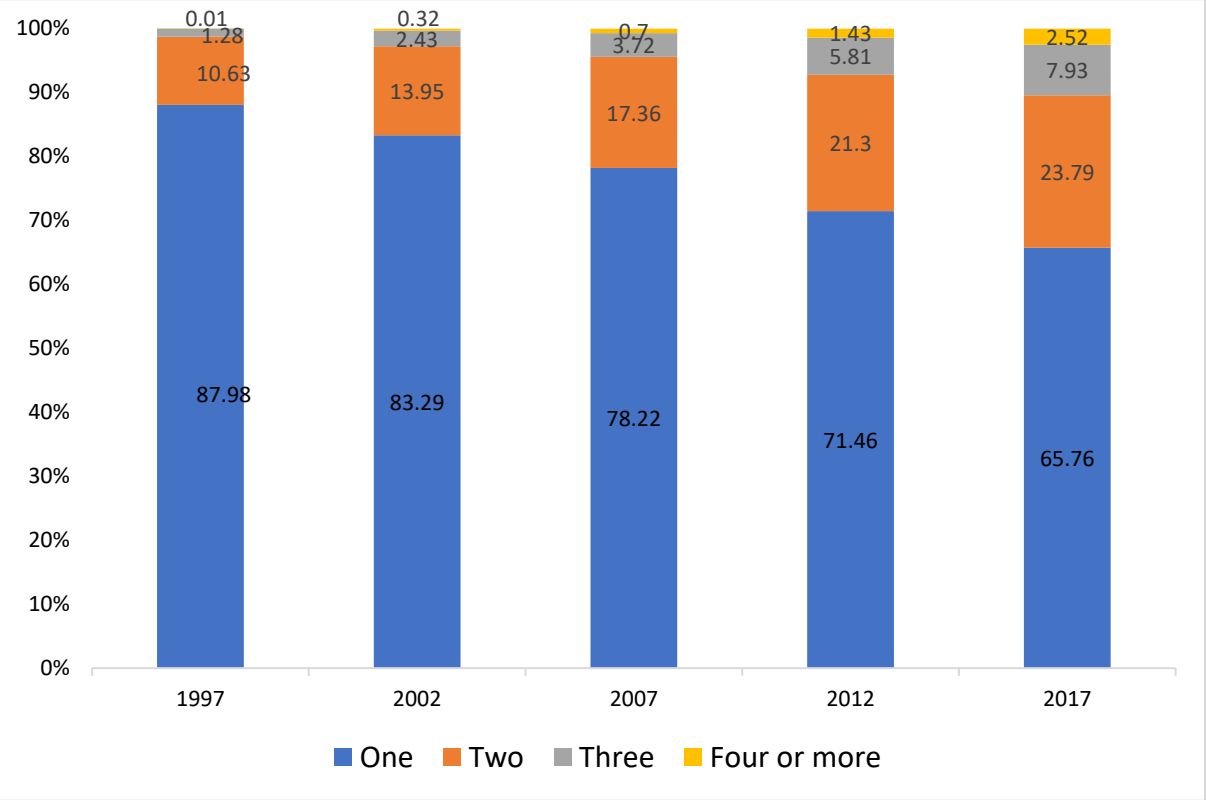
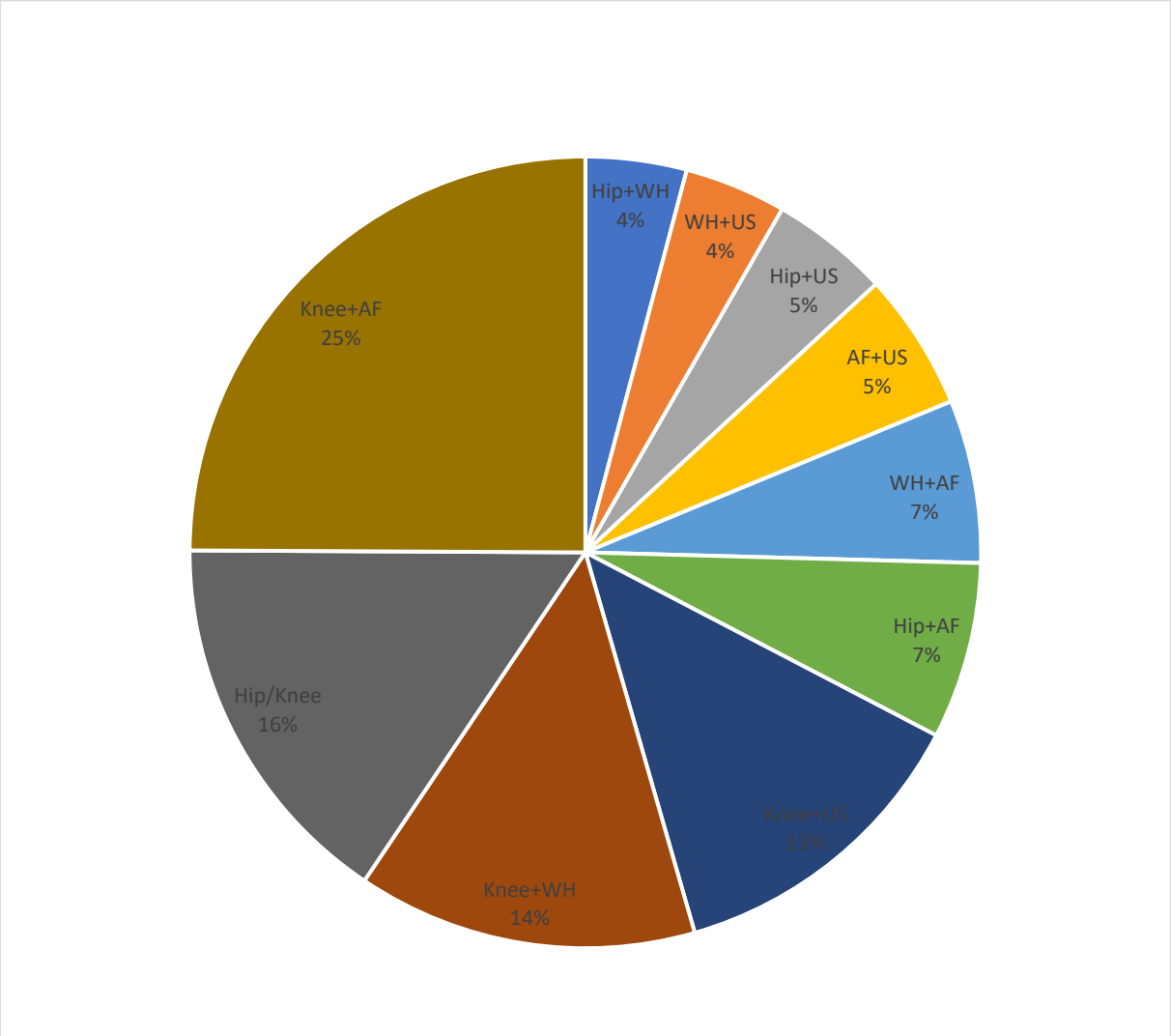


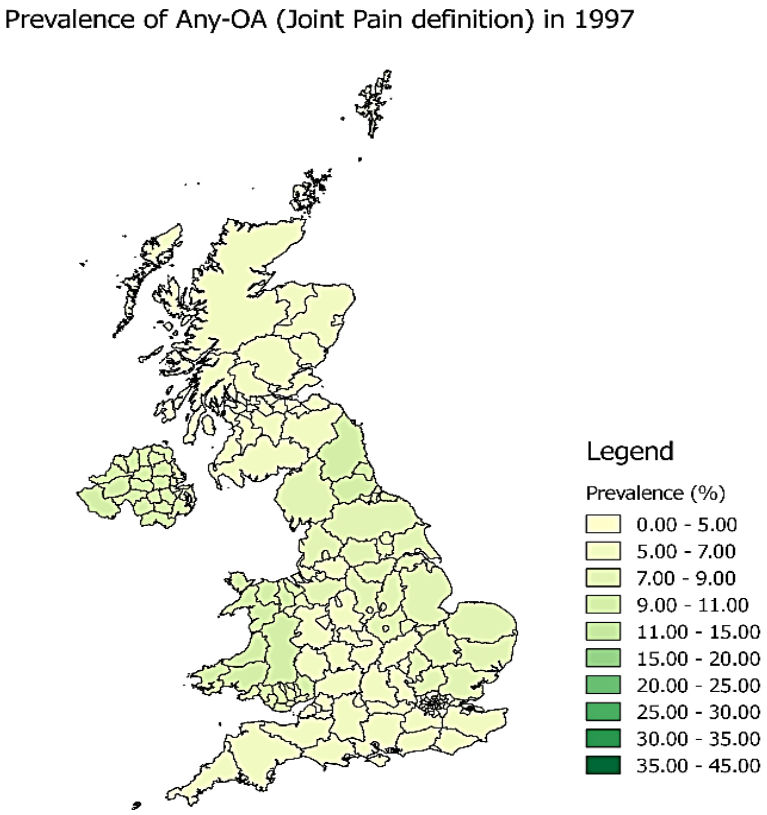
Figure 3.3-23. Pattern of sites involved in any two joint pain defined OA in 2017

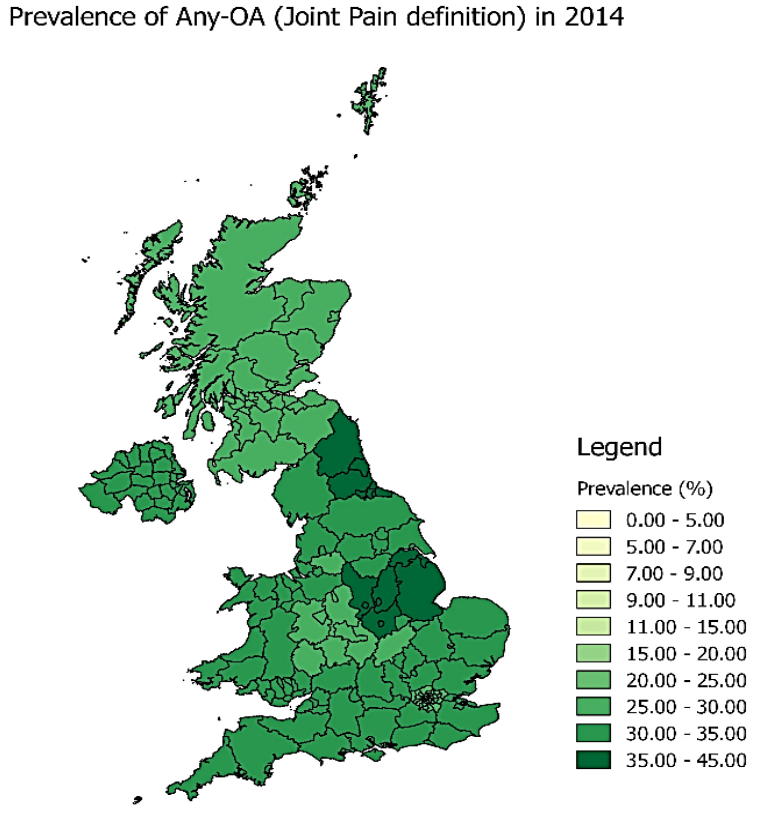


AF: Ankle/Foot; WH: Wrist/Hand; US: Unspecified

Exploring the pattern among people having two or more sites involved showed the leading pattern of joint pain was knee pain in combination with ankle/foot, hip, wrist/hand in decreasing order. (Figure 3.3-23)

Figure 3.3-24. Prevalence of joint pain defined OA in the UK during 1997 and 2014





There has been an increase in joint pain prevalence in all regions of the UK. In 2014 the highest prevalence was seen in the East Midlands and North-East regions of England. The prevalence is seen to be more uniform across the geographic regions within the range of 30%-45% (Figure 3.3-24).

3.3.2.4 Incidence of joint pain defined OA in 2017

During the 814,595 person-years of follow-up in 2017, 25,130 patients were newly diagnosed with joint pain. The age-sex and LOD standardised incidence was 30.85 per 1000 person-years (95% CI 30.47-31.22 per 1000 person-years). (Table 3.3-10) The incidence was higher in women (39.6 per 1000 person-years; 95% CI 38.9-40.2) than in men (31.3 per 1000 person-years; 95% CI 30.8-31.8).

Table 3.3-10 Incidence of joint-pain defined OA

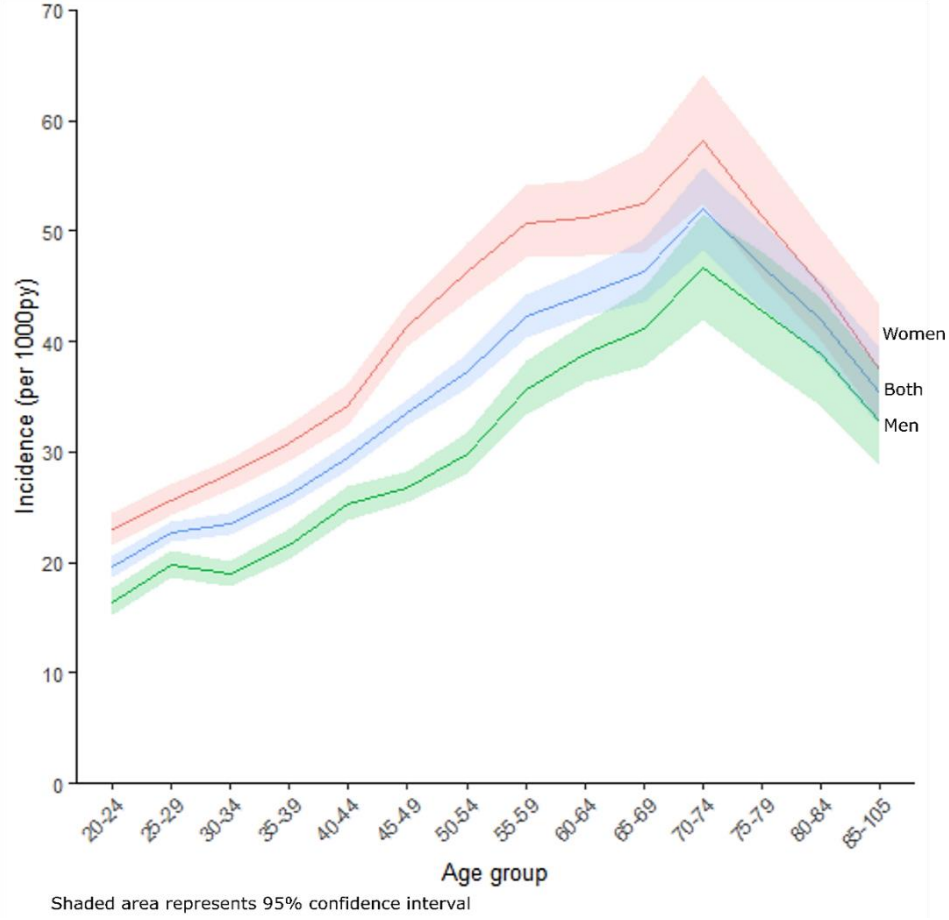
			Incidence (per 1000 person-years) 95%CI		
Year	Person-Years	Cases	Crude	Age-Sex standardised	Age-sex-LOD standardized
1997	1076611	25153	23.36[23.01-23.61]	21.37[21.10-21.64]	22.27[22.04-22.54]
1998	1233658	27944	22.65[22.35-22.95]	20.80[20.55-21.05]	21.59[21.35-21.85]
1999	1502017	31445	20.93[20.71-21.15]	19.20[18.99-19.42]	19.90[19.69-20.12]
2000	1842864	36857	20.00[19.72-20.28]	18.33[18.14-18.52]	19.03[18.84-19.22]
2001	1999954	41891	20.94[20.74-21.14]	19.28[19.09-19.47]	19.98[19.79-20.17]
2002	2289680	55593	24.28[24.06-24.40]	22.27[22.08-22.46]	22.97[22.78-23.16]
2003	2426284	69062	28.46[28.23-28.69]	26.18[25.98-26.38]	26.68[26.48-26.88]
2004	2520108	78154	31.01[30.71-31.29]	28.71[28.51-28.91]	29.21[29.01-29.43]
2005	2438636	78963	32.38[32.12-32.64]	31.16[30.76-31.57]	31.67[31.26-32.07]
2006	2447866	82525	33.71[33.41-33.91]	32.21[31.89-32.52]	32.71[32.39-33.02]
2007	2428912	87288	35.93[35.63-36.13]	34.26[33.98-34.54]	34.66[34.39-34.94]
2008	2342811	86841	37.06[36.82-37.30]	35.89[35.60-36.18]	36.29[36.00-36.58]
2009	2189674	83468	38.11[37.88-38.34]	37.23[36.93-37.53]	37.63[37.33-37.93]
2010	2146039	80878	37.68[37.45-37.91]	36.82[36.52-37.12]	37.02[36.72-37.32]
2011	2058909	78329	38.04[37.72-38.36]	37.15[36.86-37.43]	37.35[37.06-37.63]
2012	1968436	72291	36.72[36.42-36.92]	36.36[36.06-36.66]	36.56[36.26-36.86]
2013	1760340	65004	36.92[36.68-37.20]	36.63[36.32-36.94]	36.83[36.52-37.13]
2014	1603263	57952	36.14[35.84-36.44]	35.60[35.28-35.91]	35.70[35.38-36.01]
2015	1363172	46664	34.23[33.92-34.54]	33.69[33.37-34.02]	33.74[33.42-34.06]
2016	1079817	35304	32.69[32.33-33.05]	32.36[32.01-32.73]	32.41[32.06-32.78]
2017	814595	25130	30.85[30.40-31.20]	30.85[30.47-31.22]	30.85[30.47-31.22]

Age-sex standardized rates are standardized with CPRD 2017 UK population as standard population.

The incidence (per 1000 person-years) of joint pain defined OA in 2017, in decreasing order, was knee (14.96; 95% CI 14.70-15.22), ankle or foot (9.24; 95% CI 9.04-9.44), hip (7.29; 95%CI 7.10-7.49), wrist or hand (6.29, 95%CI 6.14-6.45) and unspecified (2.38, 95%CI 2.24-2.51).

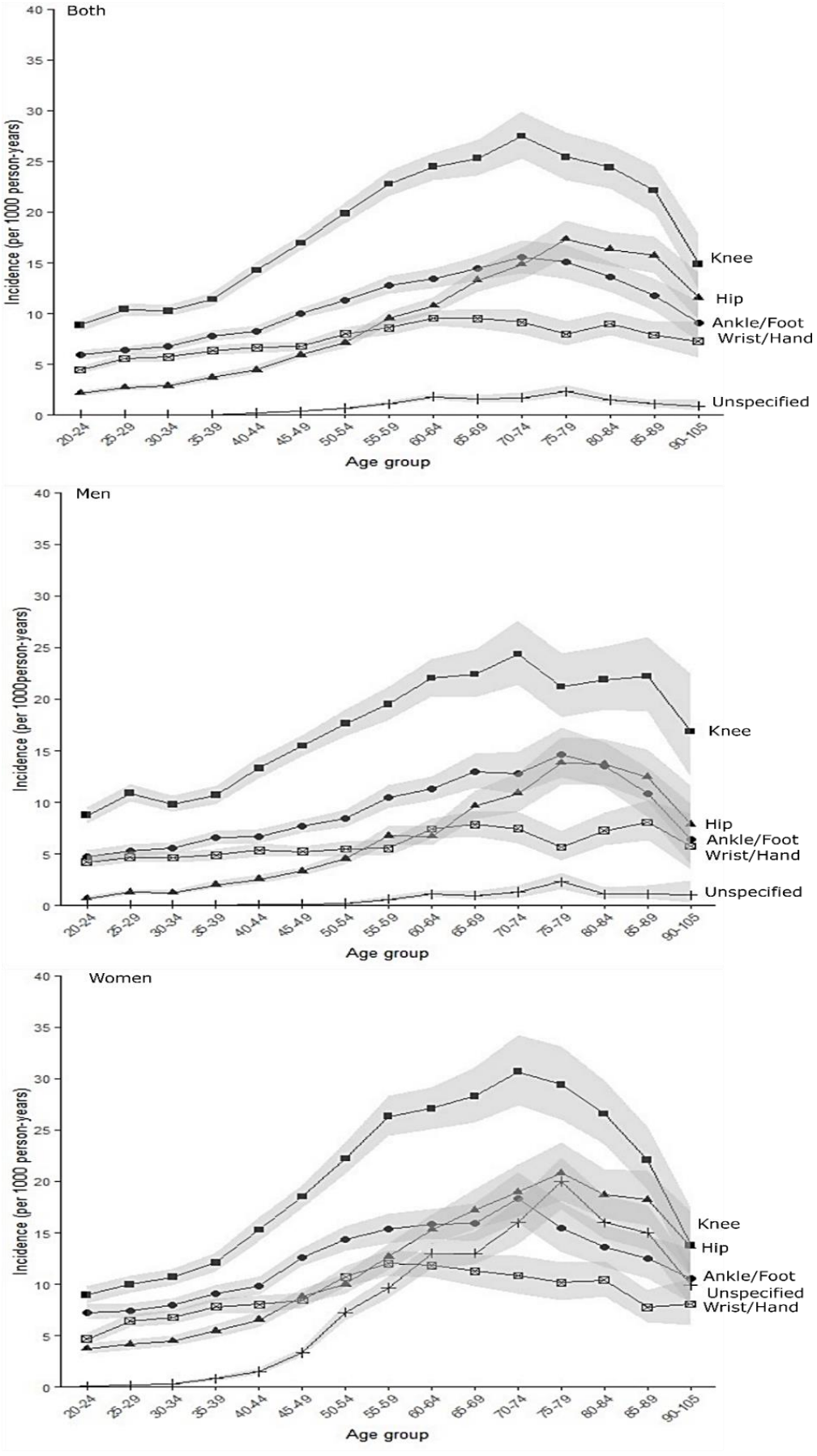
Age specific distribution for 2017 showed an incidence in the younger age group (20-24 years) of 19.6 per 1000 person-years [95% CI 18.7-20.6], which increased and peaked at 70-74 years of age (51.9 per 1000 person-years; 95% CI 48.3-55.7) followed by a decline with age. Women had higher incidence rates compared to men in all age groups. The difference in incidence rates between men and women was lowest (6 per 1000 person -years) in the younger age (25-29 years) but increased to 17 per 1000 person-years in the middle age group (50-54 years) and then narrowed in the elderly population (80-84 years). (Figure 3.3-25)

Figure 3.3-25. Incidence of joint pain defined OA across the age groups in 2017



The joint specific incidence of joint pain is depicted in Figure 3.3-26. The descending order of incidence rates are knee (14.9 per 1000 person-years; 95% CI 14.7-15.2), ankle/foot (9.2 per 1000 person-years; 95% CI 9.1-9.4), hip (7.3 per 1000 person-years; 95% CI 7.1-7.4), wrist/hand (6.3 per 1000 person-years; 6.1-6.4) and unspecified (3.8 per 1000 person-years; 95% CI 3.7-3.9). At each site, the incidence increased with age but then declined after age 70 years. A sharper rise in the incidence rate with age group was seen at the knee and hip compared to other sites. Up until age 50 years, the sites with the highest incidence rates were knee, ankle/foot, wrist/hand, and hip, but this order changed after 70 years of age to knee, hip, ankle/foot, wrist/hand and unspecified. A similar order was maintained with age in men, whereas in women older than 70 the leading sites in descending order of incidence were knee, hip, unspecified and ankle/foot.

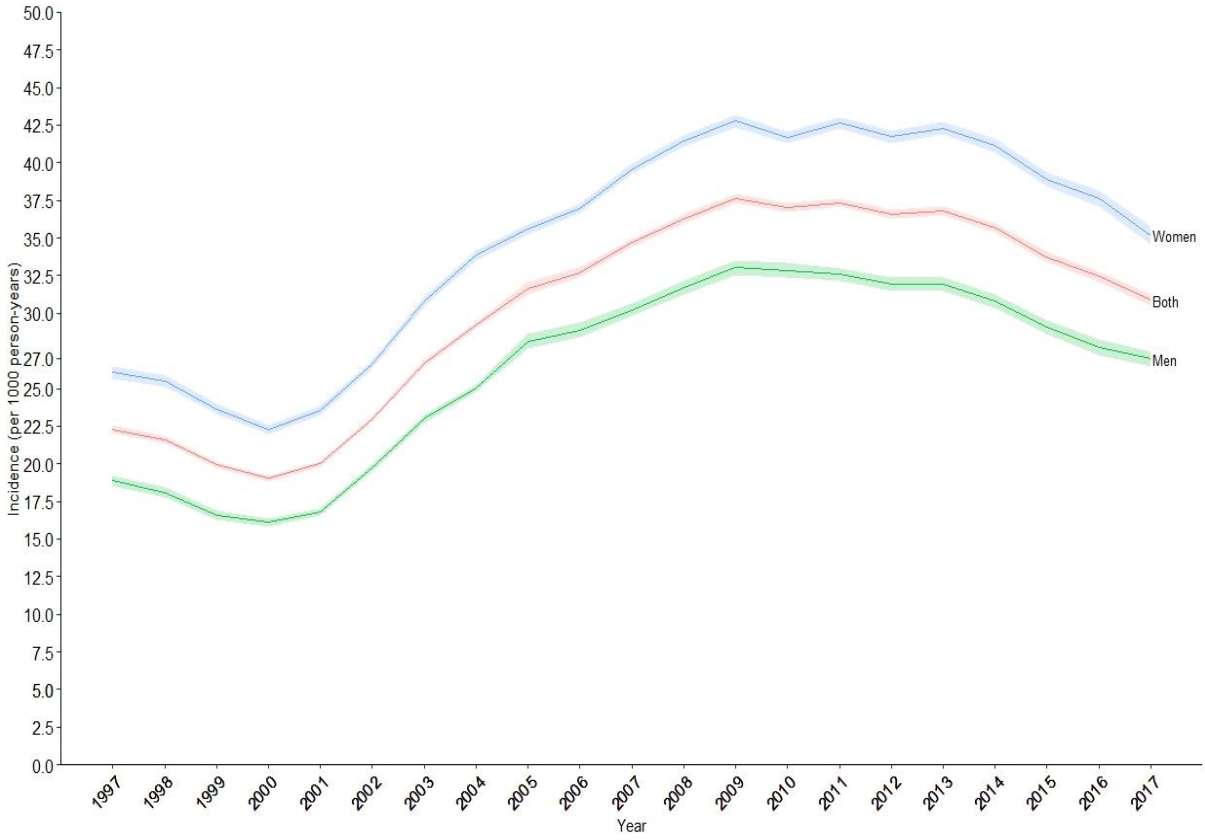
Figure 3.3-26. Incidence of joint pain defined OA across the age group by joint



3.3.2.5 Trends of joint pain defined OA incidence in the UK (1997-2017)

The trend of joint pain incidence showed a decline in earlier years which then increased after 2004. The standardised rate changed from 22.3 per 1000 person-years [95% CI 22.0-22.5 person-years] in 1997 to 30.8 per 1000 person-years [95% CI 30.5-31.22 person-years] in 2017. Both men and women showed a similar trend and remained in parallel throughout the study years. In men the incidence rate increased from 18.9 per 1000 person-years [95% CI 18.5-19.2] in 1997 to 26.9 per 1000 person-years [95% CI 26.5-27.4 person-years] in 2017. In women the incidence rate changed from 26.0 per 1000 person-years [95% CI 25.6-26.5 person-years] in 1997 to 35.2 per 1000 person-years [95% CI 34.6-35.7 person-years] in 2017. (Figure 3.3-27)

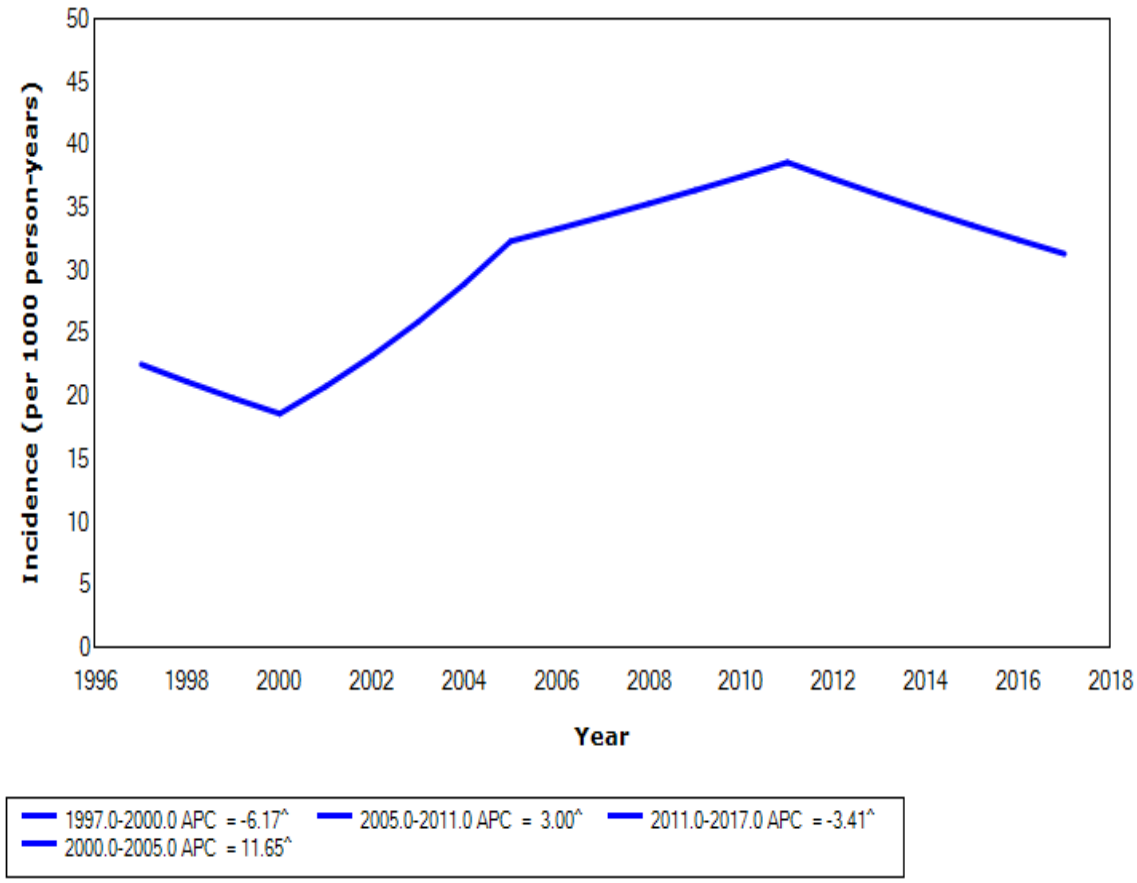
Figure 3.3-27. Trends of standardised incidence rates of joint pain defined OA in the UK (1997-2017)



The joinpoint analysis demonstrates the annual percentage change of the trend. It divides the trend line into four periods, specifically: declining (1997-2000); rapid increase

(2000-2005); slow increase (2005-2011); and declining (2011-2017). The APC between the years 2000-2005 was 11.65% and the average annual percentage change from 1997 to 2017 was 1.9% [95% CI 1.0-2.7]. (Figure 3.3-28) The AAPC in men (2.1%; 95% CI (-1.3% to 2.9%) was higher compared to women (1.7%; 95% CI 0.6%-2.8%), which was statistically significant. (Figure 3.3-28)

Figure 3.3-28. Joinpoint analysis of the trend of any joint pain defined OA incidence



The site-specific trends are shown in Figure 3.3-29. The increase in trend for joint pain was highest for knee pain followed by ankle/foot, hip, and wrist/hand pain, whereas the trend for unspecified joint pain showed a decline. The incidence rate of knee pain increased from 8.5 per 1000 person-years [95% CI 8.4 to 8.7] in 1997 to 14.9 per 1000 person-years [95% CI 14.7 to 15.2] in 2017. The wrist/hand region pain incidence increased from 2.7 per 1000 person-years [95% CI 2.6 to 2.8] in 1997 to 6.3 per 1000 person-years [95% CI 6.1 to 6.4] in 2017. Men and women showed similar trends over the study period. There was an increase in AAPC for all regions except for unspecified

(-1.1; 95% CI -2.6 to -0.4). The highest increase was seen for wrist/hand pain (4.4; 95% CI 3.2 to 5.7) followed by ankle/foot pain (3.1; 1.7 to 4.6). Details of the distribution and AAPC are given in Table 3.3-11 and Table 3.3-12.

Figure 3.3-29. Trends of joint pain defined OA for different joints in the UK (1997-2017)

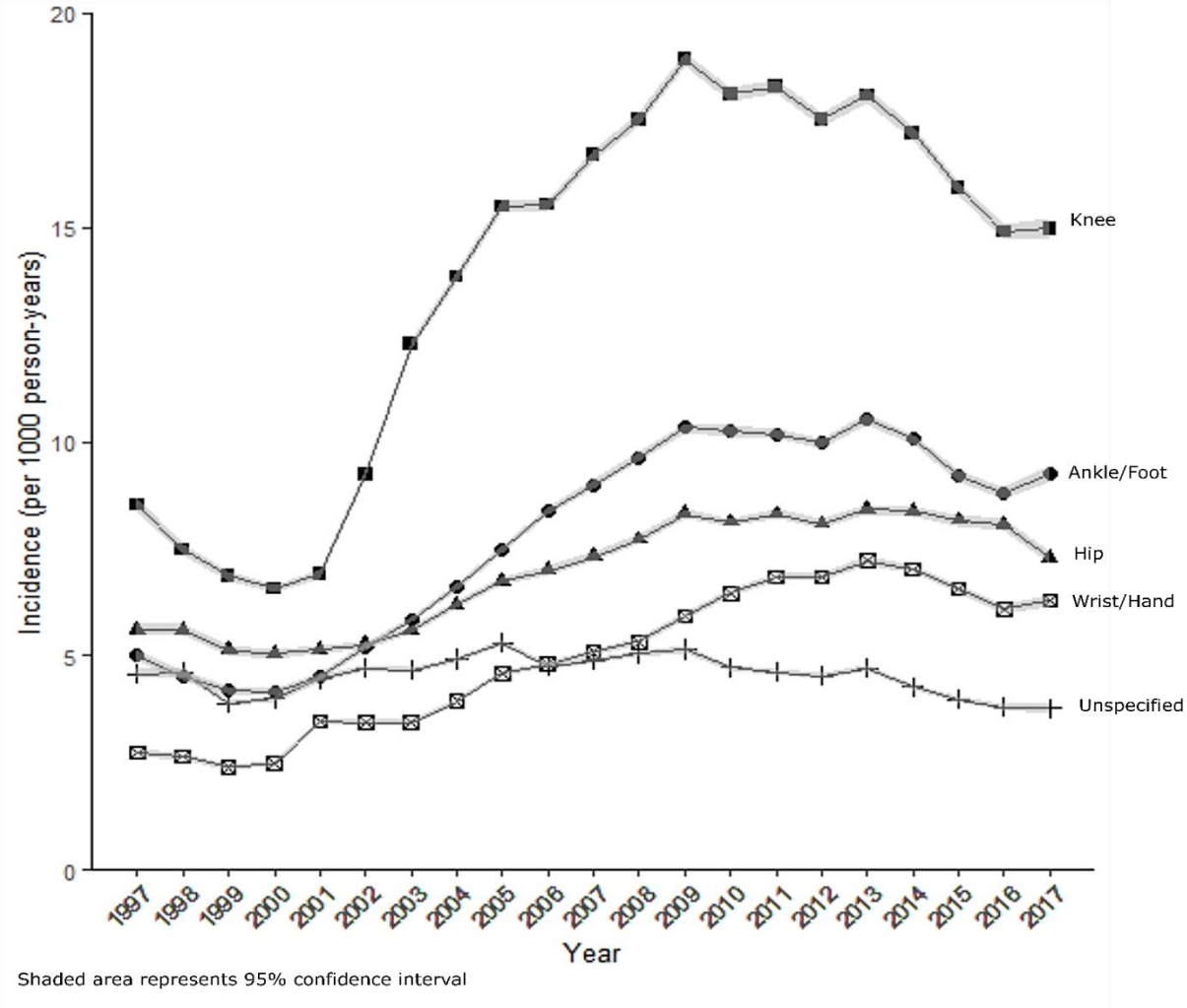


Table 3.3-11 Trends in incidence of joint-pain defined OA by joint in men

	Standardised Incidence (per 1000 person years) [95% CI]				
	Hip	Knee	Wrist/Hand	Ankle/Foot	Unspecified
1997	3.65[3.60-3.70]	8.25[8.20-8.30]	1.93[1.86-2.00]	4.24[4.18-4.30]	2.97[2.91-3.06]
1998	3.71[3.65-3.77]	7.21[7.14-7.28]	1.98[1.91-2.05]	3.73[3.67-3.79]	3.00[2.91-3.09]
1999	3.33[3.28-3.38]	6.54[6.44-6.64]	1.78[1.72-1.84]	3.32[3.22-3.42]	2.64[2.57-2.71]
2000	3.32[3.25-3.39]	6.30[6.21-6.39]	1.94[1.88-2.00]	3.34[3.24-3.14]	2.85[2.75-2.95]
2001	3.29[3.21-3.37]	6.52[6.42-6.62]	2.72[2.64-2.80]	3.71[3.63-3.79]	3.16[3.06-3.26]
2002	3.50[3.41-3.59]	8.76[8.71-8.81]	2.63[2.53-2.73]	4.29[4.21-4.37]	3.22[3.12-3.32]
2003	3.69[3.62-3.76]	11.62[11.58-11.66]	2.65[2.60-2.70]	4.72[4.65-4.77]	3.18[3.12-3.24]
2004	4.18[4.10-4.26]	12.96[12.90-13.02]	3.00[2.92-3.08]	5.38[5.32-5.44]	3.19[3.12-3.26]
2005	4.66[4.48-4.85]	14.57[14.33-14.80]	3.50[3.39-3.61]	6.16[6.00-6.33]	3.56[3.45-3.68]
2006	4.71[4.57-4.86]	14.71[14.48-14.93]	3.73[3.62-3.84]	6.94[6.79-7.09]	3.21[3.12-3.31]
2007	4.87[4.73-5.01]	15.71[15.49-15.93]	3.94[3.84-4.05]	7.35[7.20-7.50]	3.29[3.20-3.39]
2008	5.26[5.11-5.40]	16.41[16.18-16.64]	3.95[3.85-4.06]	7.96[7.80-8.13]	3.34[3.24-3.44]
2009	5.61[5.45-5.76]	18.07[17.82-18.32]	4.64[4.52-4.76]	8.60[8.43-8.78]	3.53[3.42-3.64]
2010	5.48[5.32-5.63]	17.28[17.03-17.52]	5.04[4.91-5.16]	8.55[8.39-8.72]	3.21[3.11-3.31]
2011	5.39[5.23-5.54]	16.94[16.70-17.19]	5.40[5.27-5.53]	8.44[8.27-8.60]	3.06[2.96-3.16]
2012	5.32[5.17-5.47]	16.56[16.31-16.81]	5.41[5.28-5.54]	8.30[8.12-8.47]	2.99[2.89-3.09]
2013	5.59[5.42-5.76]	16.88[16.62-17.15]	5.60[5.46-5.74]	8.67[8.49-8.86]	3.11[3.00-3.22]
2014	5.44[5.27-5.61]	15.99[15.73-16.26]	5.53[5.39-5.68]	8.23[8.05-8.41]	2.84[2.74-2.94]
2015	5.43[5.25-5.62]	14.73[14.46-15.00]	5.18[5.03-5.32]	7.53[7.35-7.72]	2.55[2.45-2.66]
2016	5.20[5.00-5.39]	13.56[13.27-13.85]	4.73[4.57-4.88]	7.29[7.08-7.49]	2.39[2.28-2.51]
2017	4.28[4.07-4.49]	13.48[13.14-13.83]	4.89[4.70-5.08]	7.90[7.63-8.18]	2.38[2.25-2.51]
AAPC	1.1[-0.4 to 2.5]	2.6[1.3 to 3.8]*	5.2[3.8 to 6.8]*	3.1[1.6 to 4.7]*	-0.9[-1.7 to -0.1]*

AAPC: Annual average percentage change; *P value <0.05

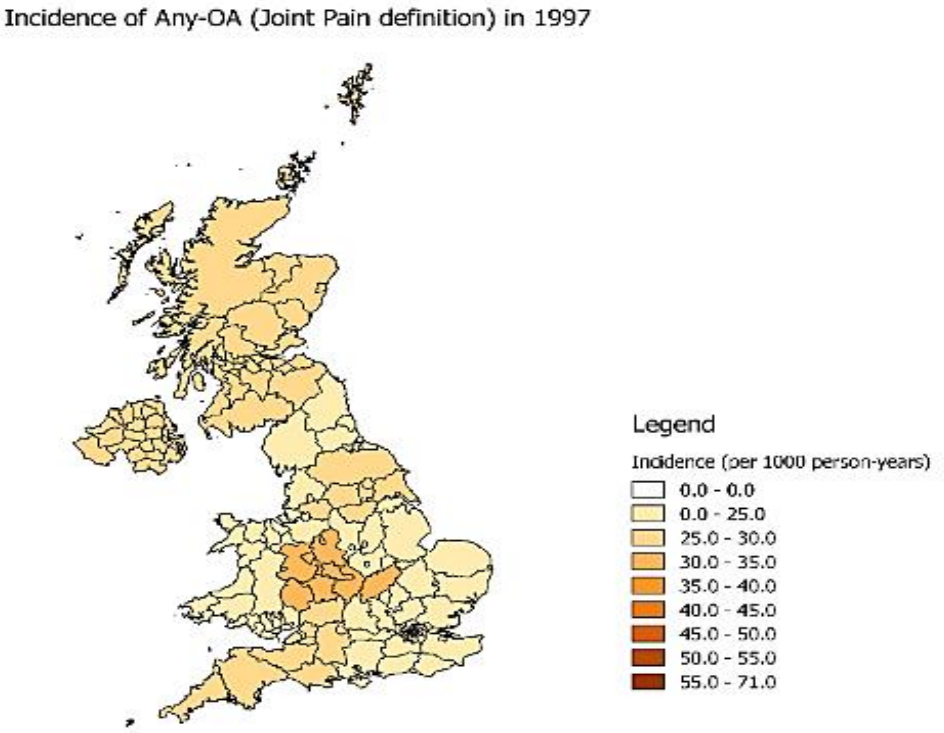
Table 3.3-12. Trends in incidence of joint pain defined OA by joint in women

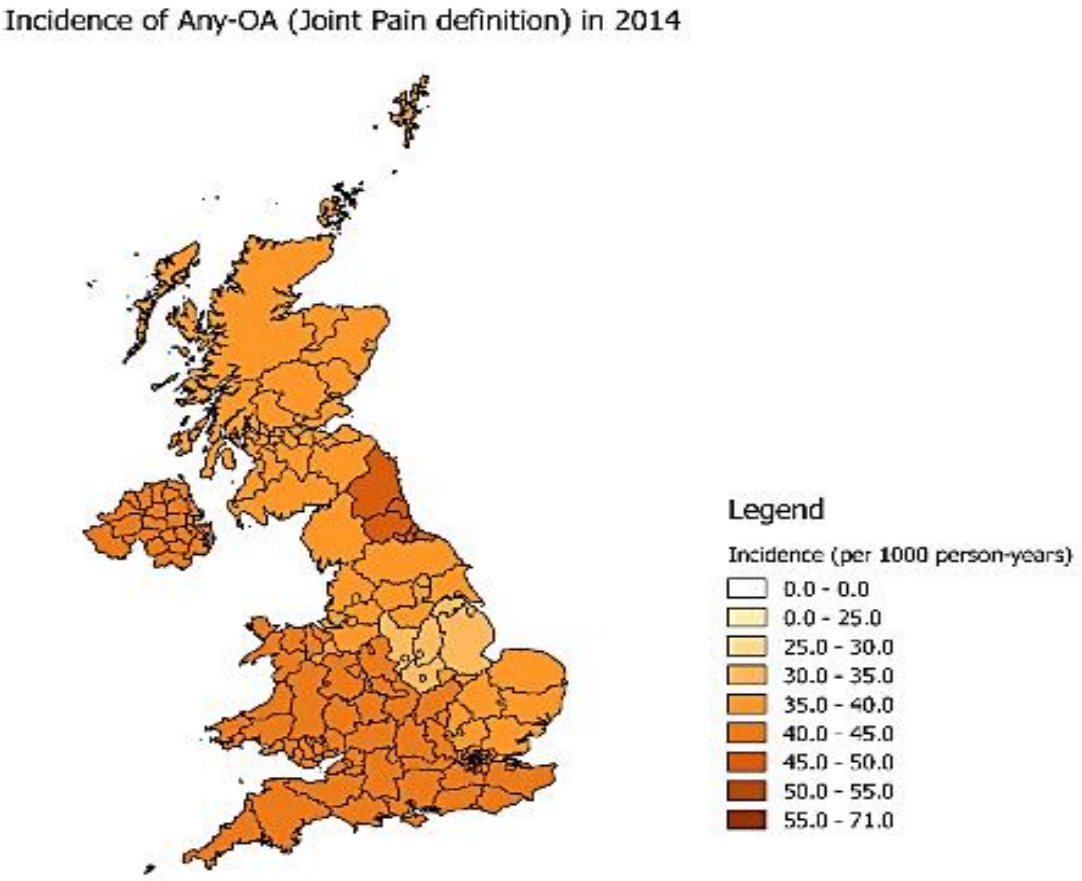
	Standardised Incidence (per 1000 person years) [95% CI]				
Year	Hip	Knee	Wrist/Hand	Ankle/Foot	Unspecified
1997	7.49[7.41-7.57]	8.80[8.60-9.00]	3.52[2.92-4.12]	5.75[4.75-6.75]	6.10[5.10-7.10]
1998	7.43[7.38-7.48]	7.74[7.54-7.94]	3.30[2.50-4.10]	5.30[4.30-6.30]	6.16[5.76-6.56]
1999	6.84[6.77-6.91]	7.19[6.90-7.39]	3.01[2.00-4.00]	5.05[4.05-6.05]	5.10[4.10-6.10]
2000	6.70[6.61-6.79]	6.80[6.65-6.95]	3.01[2.01-4.01]	4.91[4.00-5.82]	5.10[4.10-6.10]
2001	6.92[6.82-7.02]	7.28[7.16-7.40]	4.21[4.04-4.36]	5.29[4.59-5.99]	5.70[5.00-6.40]
2002	6.95[6.90-7.00]	9.67[9.56-9.78]	4.21[4.05-4.37]	6.06[5.06-7.06]	6.15[5.25-6.85]
2003	7.42[7.32-7.52]	12.88[12.78-12.98]	4.19[4.00-4.38]	6.94[6.76-7.12]	6.08[5.08-7.08]
2004	8.08[8.00-8.16]	14.69[14.59-14.79]	4.84[4.71-4.97]	7.75[7.55-7.95]	6.57[6.38-6.77]
2005	8.71[8.53-8.88]	16.36[16.14-16.59]	5.63[5.51-5.76]	8.72[8.56-8.89]	6.98[6.84-7.12]
2006	9.17[9.00-9.35]	16.33[16.11-16.55]	5.84[5.72-5.97]	9.74[9.58-9.91]	6.24[6.12-6.37]
2007	9.65[9.47-9.83]	17.60[17.37-17.84]	6.18[6.05-6.31]	10.55[10.38-10.73]	6.44[6.31-6.57]
2008	10.06[9.87-10.25]	18.55[18.30-18.79]	6.63[6.50-6.77]	11.21[11.03-11.40]	6.68[6.55-6.82]
2009	10.9[10.69-11.10]	19.72[19.45-19.98]	7.14[7.00-7.29]	11.97[11.78-12.17]	6.72[6.57-6.86]
2010	10.64[10.44-10.84]	18.91[18.65-19.16]	7.82[7.67-7.97]	11.87[11.68-12.07]	6.17[6.03-6.30]
2011	11.08[10.87-11.29]	19.54[19.27-19.80]	8.21[8.06-8.37]	11.83[11.63-12.02]	6.12[5.98-6.25]
2012	10.70[10.49-10.91]	18.44[18.17-18.71]	8.21[8.05-8.38]	11.54[11.34-11.74]	6.00[5.86-6.14]
2013	11.12[10.89-11.34]	19.21[18.92-19.50]	8.76[8.58-8.94]	12.28[12.07-12.50]	6.25[6.11-6.40]
2014	11.16[10.93-11.40]	18.35[18.06-18.64]	8.43[8.25-8.60]	11.79[11.58-12.01]	5.64[5.50-5.79]
2015	10.78[10.53-11.03]	17.06[16.76-17.35]	7.88[7.69-8.06]	10.80[10.58-11.03]	5.36[5.21-5.50]
2016	10.81[10.52-11.09]	16.17[15.84-16.50]	7.41[7.21-7.61]	10.22[9.98-10.46]	5.17[5.00-5.33]
2017	10.17[9.85-10.49]	16.38[15.99-16.77]	7.63[7.40-7.87]	10.52[10.23-10.81]	5.11[4.92-5.31]
AAPC	1.6[0.8 to 2.4]*	3.2[1.9 to 4.6]*	4.2[2.9 to 5.4]*	2.9 [1.9 to 3.9]*	-0.8[-2.3 to 0.7]

AAPC: Annual average percentage change; *P value <0.05

Geographic distribution of joint pain incidence showed an increase since 1997 in all regions except Yorkshire and The Humber. In 2017 the highest incidence was seen in Northern Ireland and the West Midlands, followed by the Southern England region. (Figure 3.3-30)

Figure 3.3-30. Incidence of joint pain defined OA in the UK during 1997 and 2014





3.4 Discussion

This is the first study to describe trends in both incidence and prevalence of GP-diagnosed OA and the alternative definition of OA (joint pain defined OA) in the UK. To date the available literature has not been able to provide clear epidemiological information. This study describes the higher burden of OA in the UK with a prevalence of 10.7% and incidence 6.8 per 1000 person-years in people aged 20 and over in 2017. The prevalence of site-specific OA in descending order is unspecified (7.6%), knee (2.9%), hip (1.5%), wrist/hand (0.5%) and ankle/foot (0.3%). Furthermore, the OA prevalence is increasing at a rate of 1.4% per year since 1998, whereas its incidence is declining at a rate of -1.6% per year. Geographically, the prevalence of OA is not uniformly distributed. Scotland, Northern Ireland, and the West Midlands had higher prevalence compared to the rest of the country. The incidence was highest in the East Midlands and North-East regions.

3.4.1 Prevalence and incidence of GP-diagnosed OA

Earlier studies from different countries have reported the overall prevalence among those aged 45 years and over to vary between 20% to 35% (Dillon *et al.*, 2006; Brennan-Olsen *et al.*, 2017). The prevalence in the USA was 28% (J. M. Jordan *et al.*, 2007) and in Australia, the prevalence of any OA in adults aged 18 years or more was 20.4%. In this study, the estimated prevalence among people aged 20 years or more using a primary care database and the findings is lower than others. The estimated prevalence from this study among people aged 45 years or more using the entire CPRD database was nearly 23%. Jordan et al, compared the prevalence of OA determined in four different databases and reported a prevalence of OA of 1.6% in those aged 15 years or more in the GPRD database, currently known as CPRD (K. Jordan *et al.*, 2007). A similar

finding was reported later in 2014 in the CPRD database of North Staffordshire (Jordan *et al.*, 2014).

Very few studies have examined the incidence of OA, especially in the UK. Using CPRD data, one study reported the standardised incidence of OA in 2014 to be 6.3 per 1000 person-years (Yu *et al.*, 2017) compared to 6.8 per 1000 person-years seen in this study in 2017. The estimates are lower than that reported in the Canadian study (Rahman *et al.*, 2014) and higher than those reported in Spain (Prieto-Alhambra *et al.*, 2014). However, the patterns in age and gender stratified rates are similar irrespective of the definition of OA. Use of different definitions makes comparison between studies very difficult. However, it should not affect the comparison within the study for different groups such as incidence by age and gender. Although 'physician-diagnosed' definition of OA was used, the findings are quite comparable to other administrative database studies. Another published report from GPRD found the annual incidence of OA was nearly 15 per 1000 person-years among those aged 40 years or more (Parsons *et al.*, 2011). An extra analysis was performed to calculate the incidence in 2017 among people aged 40 years or more. The standardized incidence of any-OA in 2017 among people aged 40 years or more to be 16.2 per 1000 person-years in this study. Yu et al reported the standardised incidence rate of any joint OA to be 8.6 per 1000 person-years among persons aged 15 years or more in a regional administrative database (Yu *et al.*, 2015) and suggesting a similar burden found in this study.

The increasing prevalence and incidence with age and in women in this study supports the existing epidemiological findings for OA. The sudden rise of both prevalence and incidence at the age group of 40 years has been biologically explained and reported uniformly in previous studies (Zhang and Jordan, 2010). Similarly, the decline of prevalence and incidence that was observed in later old age (85 years or more) is a common phenomenon for almost all chronic conditions. This may be because of two reasons: [1] the smaller sample size available for those aged more than 85 years; and

[2] people with OA are likely to have cardiovascular and other comorbidities hence die sooner that those without OA, resulting in relatively healthier people surviving into this age band ("healthy survivor" bias). The incidence pattern with age group accords with previous findings from the UK and other countries (Parsons *et al.*, 2011; Yu *et al.*, 2017). Also, there is enough scientific consensus regarding the biological relationship of sex and age with OA (Doherty, 2001). The association of OA with gender, along with other particular risk factors (e.g. increased weight and bone mass) and the increasing prevalence of OA in women following the menopause, has signalled the role of oestrogen as an influence (Spector and Campion, 1989; Spector *et al.*, 1996).

In both sexes, the prevalence and incidence of 'unspecified' site OA was high compared to other joint OA. A similar finding was reported by Yu et al (Yu *et al.*, 2015). The 'unspecified' reporting of the OA site explains the recording pattern in primary care. However, whether the 'unspecified' term used in the database is a substitute to record multiple joint involvement, remains unclear. Without proper radiographic evidence, physicians might record it as unspecified, which needs further exploration. This suggests that the GP coding for OA needs to be improved. In addition, care must be taken when interpreting the prevalence of OA by different body sites.

Many people with OA may never consult their GP about their OA (Yu, Jordan and Peat, 2018), so the prevalence from CPRD is generally lower than in studies that examine samples of the general population specifically for OA. For example, in this study, the prevalence of knee OA was nearly 3% whereas, a community-based study of adults in Canada reported the prevalence to be 8.5% (Plotnikoff *et al.*, 2015). Similarly, the prevalence of hip OA in this study was nearly 1.2%, whereas in the same Canadian study the prevalence was reported to be 10.5% (Plotnikoff *et al.*, 2015). However, studies have reported the prevalence of knee OA to be higher followed by hip and wrist/hand, similar to findings from this study (Kingsbury *et al.*, 2014).

The incidence rates of hip and knee OA are quite comparable to the incidence calculated using the Consultations in Primary Care Archive (CiPCA) database from the UK (Yu *et al.*, 2015). The hand/wrist OA incidence (0.65 per 1000) is slightly lower than the published rate (1.1 per 1000 person-years)(Yu *et al.*, 2015). The incidence rates for hand/wrist OA reported in those aged 40 years or more are 1.2 per 1000 person-years, close to the findings in the same age group in the CiPCA database. This could be because of the quality of the database, as CPRD GOLD database used for this study represents the whole country and wider heterogeneity in diagnosis and recording, while the published literature had a better supervised uniform recording of OA diagnosis in primary care database. Other factors could be because of the different health seeking behaviour of people in different parts of the country, for example a preference to see a physician for large joint problems compared to small joints.

Both the incidence and prevalence of ankle/foot OA in this study was much less compared to the previously reported research (Menz *et al.*, 2010; Roddy and Menz, 2018). This could be due to the difference in study population, diagnosis and recording of the ankle/foot OA by the GPs. McCarthy et al reported the inconsistency and incomplete recording of the musculoskeletal recording in primary care, especially the examination of the joints (McCarthy, Sheane and Cunnane, 2009). Other possible reasons could be the reporting of the symptoms by the people in primary care and diagnostic facilities available to identify specific causes.

Other factors for differences in results could be because of the broader age band used (20 years and more). This study included those aged 20 years or more, so lower estimates compared to other studies are justified in the context of a larger proportion of younger people in the denominator (eligible population). However, the standardised incidence is higher.

3.4.2 Trends of prevalence and incidence

In the year 2017 an increase in standardized prevalence from 1997 was observed. The annual percentage increase was 3.9%. So far, no literature is available on trends of OA prevalence. OA is a chronic condition without a specific cure, thus, once a person has consulted for OA the record remains in the database until the person dies or leaves the practice. This suggests increased healthcare utilization for OA in primary care. OA is a common complex disorder with variation in phenotypic expression and sometimes OA may be difficult to differentiate from other joint and regional pain conditions. No significant change in rate for joint-specific prevalence was found, but the 'unspecified' OA rate was declining, indicating possible improvement in clinical coding. Perhaps the increase in trend of 'joint-pain' after the year 2005 partially explains the gap if physicians became more prone to report symptoms rather than a specific diagnosis. The third reason for this increase may be related to a change in some risk factors for development of OA, such as the increase in obesity in recent decades, which is nearly parallel to the prevalence rise of OA.

Surprisingly, an overall slow decline in incidence rates for any-OA since 1997 was seen. The annual percentage change is on a downward trend at the rate of -1.7% per 1000 person-years. However, detailed joinpoint analysis reveals a slight rise in incidence from 2000-2004 followed by a slow decline. The decline might be because of the introduction of the Quality of Framework (QoF) introduced by the NHS for quality recording of cases which included OA, which could have increased the accuracy in recoding the OA in the data base. A similar trend was seen in a 'GP diagnosed' diabetes analysis in the CPRD (Tate *et al.*, 2017), which explains the change in pattern of coding after the year 2004. However, the insignificant change in rate for site specific incidence except for 'unspecified' which is declining at an annual rate of 2.3 per 1000 person-years, indicates improvement in coding. It may be that the increase in incidence of 'joint-pain' after the

year 2005 partially explains the gap, which is discussed later. (Appendix Fig 5, page 306) Physicians might be increasingly reporting the symptoms rather than the diagnosis, which can be a problem in large databases.

Age-period-cohort effects, length of data contribution and the participation of practices in the CPRD database influence the incidence estimates (K. Jordan *et al.*, 2007; Kuo *et al.*, 2014b). Age-period-cohort analysis shows a strong cohort effect in incidence among people born after the 1960s. It suggests that people born after this period may be exposed less to physically very demanding occupations such as coal-mining, farming and certain heavy industrial work because of changing patterns of occupation in the UK since the 1960s, including the decline in mining activities (*Long-term trends in UK employment: 1861 to 2018 - Office for National Statistics*, 2019). The maximum available length of data run-period was used to eliminate the problem of prevalent cases for OA for robust incidence estimates. In contrast, prevalence remained almost unchanged in people born after the 1960s.

3.4.3 Geographical distribution of the prevalence and incidence

There is clear evidence for regional variations in OA. OA is more prevalent in Scotland, the West Midlands and Northern Ireland compared to the rest of the UK. The prevalence for all the regions was estimated until 2014, but because of lack of data from the East Midland regions after 2014 comprehensive comparison could no longer made. However, the geographical distribution of the prevalence up until that time may represent consultation behaviour changes rather than a change in OA per se. Other reasons for a higher prevalence in certain regions could be because of different socio-economic conditions, lifestyles, and variations in health seeking behaviours in the population. Interestingly findings from this study largely match the obesity distribution mapping of UK undertaken by the Health Survey for England, Scotland and Northern Ireland (‘Statistics on Obesity, Physical Activity and Diet, England 2018’, 2018) in that the

prevalence of obesity is higher in the Northern region of the UK compared to the South, and obesity is recognised as an important risk factor for OA (Appendix-3). However, further research is required to explain the underlying factors for OA. The spatial distribution of the disease prevalence needs careful interpretation.

As seen for prevalence, geographical non-uniformity in incidence is clearly observed in this study. The middle regions of the UK had higher incidence rates in the year 2014 compared to the rest of the country. The given map aligns with the mapping done by Yu et al for the year 2013 (Yu *et al.*, 2017). This variation might be because the rates are sensitive to the practice areas involved, definitions, coding, incentives, and length of run-in period for practices. This supports the missing information from the East Midlands after the year 2014.

3.4.4 Prevalence and Incidence of joint-pain defined OA

Even though the information for OA related joint-pain was analysed, the primary purpose was to explore the comprehensive inclusion of the definition of OA i.e., both GP-diagnosed and OA related joint pain. In most general practices recording OA as joint pain may be common because of the absence of more definite (e.g., radiographic) confirmation of the diagnosis. The standardized prevalence of joint pain in this study was 34% and was higher among women compared to men. Jordan et al reported a similar prevalence using UK and Swedish primary care databases. Nearly 30% of people aged 18 years or more had any joint pain. Major injuries and other possible causes of joint pain that happened within 1 year before the index date were excluded from the analysis. However, the increased prevalence of OA related joint pain needs further exploration. Although, there are few studies available on joint pain in the UK, the findings are quite comparable to previous literature (Finney *et al.*, 2017).

The age distribution of joint pain follows a similar pattern to GP diagnosed OA, i.e., a linear increase with age until 80-85 years followed by a decline. However, nearly 10% of

people in the age group 20-29 years in 2017 consulted for joint pain. This suggests that the problem could be related to injuries, since diagnosis of OA at younger ages is rare or not diagnosed properly or could be an issue with coding. Possibly the absence of recording of injury history (unavailable) for the joint pain might have escaped the exclusion criteria. Jordan et al, documented the prevalence of arthralgia in the age group of 15-24 years to be 19.2%, which is much higher than findings from this study (Jordan *et al.*, 2010).

Interestingly, in younger ages the prevalence for knee pain was highest followed by ankle/foot and wrist/hand pain, whereas in elderly people, this was highest for knee pain, followed by hip and ankle/foot. This supports the above-mentioned explanation about the recording of the conditions, especially the missing injury records. Joint injuries in younger age especially knee and ankle/foot are quite common because of sport or physical activities. However, in the elderly population, joint pain of knee and hip are more likely to be linked with structural changes because of OA. Another study done by Finney et al also reported knee pain as the commonest site in the UK (Finney *et al.*, 2017). The age stratified findings indicate that careful use of the joint-pain definition for OA is needed in younger people in large primary care databases.

The definition of OA used in this study based on recording of joint pain found the standardized incidence rate to be 31 per 1000 person-years in 2017, which is less than the 40.5 per 1000 person-years for 'clinical-OA' reported by Yu et al for the year 2013 (Yu *et al.*, 2015). This could be because of the different age criteria used in these studies as explained before. Thus, the incidence of joint pain is less in the younger population lowering the overall estimates compared to that calculated for 45 years or more. The pattern of estimates across age groups is consistent with other findings, i.e. an increase with age and subsequent decline after the age of 70 years (Yu *et al.*, 2015). However, high consultations in the younger age group needs further exploration since identifying the reasons for joint pain in younger people is difficult in administrative database. The

finding of higher incidence among women from this study agrees with other studies. By the age of 70 years, nearly 60% of the population have consulted for joint pain, which supports the current evidence (Parsons *et al.*, 2011). This indicates the consultation burden of 'joint-pain' which might guide future planning of the care process.

The significant increase in the prevalence of joint-pain over twenty years highlight the rising burden. There has been a steep rise in the trend after the year 2004 with incidence trend. However, the rate of change subsequently slowed down after 2013. The possible reasons discussed for increasing prevalence of GP-diagnosed OA may best explain this. The rate of change was higher for knee, ankle/foot, and hip. This could be because of the clear recording of the body site based on symptoms. The lowest trend seen was for 'unspecified' joint pain which complements the 'GP-diagnosed' OA trend, where an exact diagnosis could not be ascertained. (Appendix Figure 5, page no 306) The rate of change for the wrist/hand was highest followed by ankle/foot and knee. The higher ankle/foot pain trend echoes the recent findings from Murray et al (Murray *et al.*, 2018). It is possible that joint pain may be attributable to other pathologies occurring at the joint site or in surrounding structures, for example ligamentous or tendon injury, ankle sprain, or referred pain from other areas. The cause of the increasing prevalence of joint pain needs further study. Since 1997, the trend of joint pain defined OA incidence has not been consistent. The rise in the incidence rate from 2003 could be because of the introduction of the QOF into the NHS (NHS, 2016), which might have reduced misclassification bias. This trend complements the trends of 'GP diagnosed' OA incidence, which is seen to fall after 2003. However, there is a downward trend seen after 2009. Similar trends were observed by Yu et al, who studied up until 2013 (Yu *et al.*, 2015).

An increase in trend of knee pain was observed compared to other studies. In one of the population-based studies of changes in the prevalence of knee OA symptoms and radiographic changes in the USA, Nguyen et al. found substantial increases in self-

reported knee pain but not radiographic OA between 1974 and 1994 after adjusting for the changing distribution of BMI (Nguyen *et al.*, 2011). However, the annual percentage change for wrist/hand was highest compared to other sites (van Saase *et al.*, 1989). These joint pain estimates can be a proxy indicator of future OA severity and burden in the country. The burden of joint pain is more common in Northern Ireland and Southern parts of the UK. This could be because of the more representation of GP practices from those areas to the CPRD GOLD database. However, such unequal distribution needs further research.

3.4.5 Study limitations

There are several limitations to this study. Firstly, this study based the OA case definition on diagnosis by the general practitioners rather than presence of structural OA on imaging. However, concordance between symptoms and radiographic OA (the usual way to assess structural OA) is variable and often poor, depending on the joint site being assessed (Hunter *et al.*, 2013). Furthermore, patient-centred outcomes rather than imaging changes are key determinants of disability and burden of disease, and NICE recommend that a purely clinical diagnosis is sufficient and that imaging should be reserved for specific situations such as atypical clinical features or rapid progression of symptoms (NICE, 2014). Site specific estimates could be biased because of coding used by the practitioners (e.g., OA or 'Unspecified'). A stand-alone primary care database was used for this study and the use of alternative 'joint pain' definition reflects the primary care burden through the broader selection criteria.

Secondly, the analysis the temporal trend of BMI using CPRD GOLD data was difficult owing to its incompleteness. Therefor the national health survey data was used to explain the findings. This comparison has helped to explain high BMI trend as one of the risk factors for the prevalence and incidence.

Thirdly, the index date reflects the date of allocation of Read codes for OA and does not reflect actual disease onset. However, the date of allocation of a Read code for OA would be expected to be within a few months of the date of diagnosis. Sensitivity analysis was not performed to examine the effects of changes in Read codes on the results. However, the codes were verified in this study as explained in the methods and do not suggest a change in the Read code list will alter the findings on temporal trends in the epidemiology of OA. Exclusion criteria used in this study might have led to underestimation of the burden.

Another limitation is the geographical presentation of the estimates, which needs cautious interpretation because of the non-uniform distribution of the practices involved in the database.

3.4.6 Conclusion

The standardised incidence of GP-diagnosed OA has been declining from 2010, however the prevalence is rising gradually. An increase in the trends of both standardised incidence and prevalence of joint pain related to OA was observed. Nearly one in every 10 adults aged 20 years or more has OA and the knee is the leading site reported to be involved in both sexes. The increasing rate of 'joint pain' incidence and prevalence is a matter for concern. The results from this study suggest that the changing burden of OA and joint-pain in primary care necessitates an appropriate policy and intervention for prevention and care.

Summary of Chapter 3

Chapter 3 the burden of OA in the primary care settings in the UK. Key findings are

- The prevalence of OA in the primary care is nearly 10% and annual incidence was 7 per 1000 person years
- Women contributed to the burden more compared to men
- There was an increase in burden of OA after the age of 40 years in both men and women
- Knee OA is the most reported joint specific OA followed by hip, wrist/hand, and ankle/foot
- The trend of prevalence is increasing while the incidence is declining.

As the burden of the OA in the primary care is evident, it would be interesting to understand the burden of the other different chronic conditions present with OA. Thus, in the next chapter (chapter 4) the association of OA with various comorbidities will be explored.

4 Chapter 4

Comorbidities in OA

4.1 Introduction

Presence of two or more chronic conditions in an individual has become a norm rather than an exception (Fortin *et al.*, 2010). Multiple chronic conditions in an individual could be explained through shared pathogenesis, shared aetiology or ageing in which the risk of developing other chronic conditions becomes high (Piette and Kerr, 2006). There has been a growing interest in researching the comorbidities that may associate with OA. However, to date the list of comorbidities studied are primarily limited to CVD, diabetes, depression and chronic obstructive pulmonary diseases (Nieves-Plaza *et al.*, 2013; Stubbs *et al.*, 2016b; H. Wang *et al.*, 2016; Parkinson, Waters and Franck, 2017). According to Versus Arthritis, three in ten people with OA have more than one long term condition (Loftis, Ellis and Margham, 2014).

The systematic review shows that on average 67% of people with OA had one or more other chronic conditions, and the risk of having comorbidity was 20% greater than in those without OA (Swain *et al.*, 2019). Other systematic reviews on comorbidities in OA report significant associations with cardio-vascular conditions and diabetes (Nieves-Plaza *et al.*, 2013; H. Wang *et al.*, 2016). The presence of additional comorbidities escalates the disease severity and healthcare utilization, and demands complex management guidelines (Bähler *et al.*, 2015). The temporal and causal association between these conditions has yet to be established as most studies are cross-sectional. Current evidence is restricted to a few comorbidities and the 'time-to-event' i.e. occurrence of comorbidities after diagnosis of OA has not been studied.

Except for shared risk factors such as ageing and obesity, little is known about biological plausibility to explain concurrence of OA and associated comorbidities.

Furthermore, whether comorbidities in OA occur because of OA itself, or because of the medication used to treat OA remains unclear. Lack of evidence on causality, pattern and distribution of comorbidities in OA makes the management of OA with comorbidities less clearly defined (de Rooij *et al.*, 2014). To date, no large database studies are available on the reported associations, and many possible associations have not been investigated. Therefore, using data representative of the general population of UK from the CPRD, this study aimed to examine the burden of comorbidity at the time of first diagnosis of OA (i.e., newly diagnosed, or incident OA) compared with matched controls. I further followed patients with incident OA and their matched controls after diagnosis to compare their subsequent accumulation of comorbidities.

4.2 Methods

4.2.1 Source of data

CPRD GOLD data was used for this study.

4.2.2 Study population

For this analysis, data available for registered people since inception of CPRD (1st Jan 1985) to 31st December 2017 was used.

General criteria for inclusion were that participants should:

- be aged 20 years or more at study entry
- have had active registration for at least 36 months with the UTS practice prior to the study,
- be flagged as acceptable data (determined by CPRD database standards)

4.2.3 Case definition of OA

For this study only GP-diagnosed OA was used, because of the ambiguity in using a 'joint-pain' definition to define the OA population (please see Chapter 3 discussion section for OA and joint pain).

GP-diagnosed OA was defined as:

- at least one recorded physician diagnosis of OA for hip, knee, ankle/foot, wrist/hand, or recorded as ‘unspecified’
- any recording of knee or hip replacement in the absence of recording of GP-diagnosed OA during the study period

4.2.4 Eligible study population for retrospective and prospective analysis

For the analysis, controls were participants registered with the UTS practices who had no record of diagnosed OA, OA related joint pain or total joint replacement. One control was selected per OA case and was matched in a 1:1 ratio by age (±2 years), gender, year of first registration and practice and having at least one consultation recording in the database. The same index date as their matched case was used (date of first OA diagnosis). The matched controls were selected using the ‘sttocc’ command in Stata. Details of the method of patient selection are given in Appendix Figure 6 (page 307).

4.2.5 Comorbidity definition and extraction

Details of the comorbidity selection is provided in chapter 2, section 2.1.7.1. Forty nine chronic conditions excluding OA were extracted for the study. The comorbidities in this study were further categorised into eight groups namely, musculoskeletal, respiratory, genitourinary, neuropsychiatric, cancer, circulatory, metabolic/endocrine, and digestive. In addition, a list of six conditions were grouped as ‘other’ category. The definition of all these conditions was based on physician diagnoses recorded as Read codes. A summary of the disease list with primary codes is given in the Appendix Table 4 (page 309).

In the CPRD, the code list for important comorbidities was obtained using the medical browser provided by the CPRD interface. Wherever required, this was further refined after comparing with codes used by other researchers in the department. A final list of the codes was shared with the general practitioner collaborator (CM) for input and verification. Finally, the corrected codes were read and agreed by all the research

team. Most of the code lists for comorbidities listed are externally validated (Deyo, 1992; Nada F. Khan, Harrison and Rose, 2010).

4.2.6 Covariates

Because of the longitudinal nature of the data, the health behaviour of participants was subject to vary over time. For example, BMI status, alcohol use and smoking habits can change multiple times during 20 years of follow-up. Studies have confirmed that such health risk behaviours largely influence the incidence of comorbidities (Bhaskaran *et al.*, 2013). The whole study period was divided into five follow-up intervals (0-1 year, 0-5 years, 0-10 years, 0-15 years, and 0-20 years) after and before the index date. The status of each covariate (BMI, alcohol use and smoking status) at the end of each follow-up interval was extracted from an additional file provided by the CPRD. The purpose of restricting to five follow-up intervals was to prevent the large expansion of the database to save the data memory and time for running time varying covariate (TVC) analysis. Such an approach is suggested in the Stata manual for TVC analysis.

In case of missing information for one follow-up time, it was imputed using last observation carried forward, assuming the status remained unchanged. BMI was categorised into four groups based on the values (Kg/m^2) such as underweight (<18.50), normal (18.50-24.99), overweight (25.00-29.99) and obese (30.00 and above) (NHS, 2018). Smoking status was divided into ex-smokers, current smokers, and non-smokers. Alcohol use was grouped into non-user, ex-user, current user 1-9 units/week, current users >10 units/week and current users (unknown quantity). As missing data information was less than 10%, the whole dataset was used for analysis for which complete information on covariates was available.

4.2.7 Statistical methods

Both case-control and cohort designs were used to explore the temporal association with comorbidities. The case-control design assessed comorbidities that were present

on or before the first diagnosis of OA (up to a maximum of 20 years before the index date), whereas the cohort design assessed the occurrence of comorbidities after the diagnosis of OA.

For the case-control analysis, the prevalence of a specific comorbidity in cases and controls was estimated by calculating the proportions of people diagnosed with any comorbidities during the past 1, 5, 10, 15 and 20 years before the index date. This method was used primarily to examine whether the longer observational period would give greater prevalence - to assess the observational bias (Kuo *et al.*, 2014a). The denominator for prevalence calculation was the total number of cases or matched controls during each study period. The ORs of having two or more comorbidities (other than OA) during the retrospective time points was calculated. During the retrospective analysis, the association of multiple chronic conditions with OA was examined. The study population was divided into five groups (none, single, two, three and four or more comorbidities) based on their total count of comorbidities. ORs and 95% CIs were used to estimate the association between OA and each coexisting medical condition. Conditional logistic regression was used to adjust for age, gender, BMI, smoking, alcohol use and index date. Even though controls were matched for age and gender, for the retrospective study the outcome of interest was OA. So, these factors were adjusted in the model.

In the prospective cohort study, the incident comorbidity was assessed as the earliest date of diagnosis after the index date. The study period was until the incident date of comorbidity, death date, transfer out or end of the study (31st Dec 2017), whichever came first. Only people at risk for a given comorbidity (not having such comorbidity at index date) were considered to estimate HRs of a specific comorbidity. For calculation of cumulative probabilities, at baseline the percentage of the specific comorbidity was zero and the subsequent risk was calculated among the at-risk population. The Nelson-Aalen method was used to display the cumulative probability of each

comorbidity in people with incident OA and matched controls. Along with that, HRs with 95% CIs were calculated for each comorbidity separately using Cox proportional hazards model adjusting for age, smoking, alcohol use, BMI, and index date. Proportionality assumption for each comorbidity was tested looking at Kaplan-Meir curves and use of the Schoenfeld residual test. The Schoenfeld test was done for both global and individual covariates and OA. The Cox model incorporated a time varying covariate analysis which accounted for the change in age, BMI, alcohol use and smoking status over time, and time invariable factors such as sex and the index year.

The association and risk of 49 comorbidities with OA was tested. This simultaneous testing of several hypothesis creates the risk of higher false discovery rate which is known as 'multiple testing' (Greenland, 2008). To explain further, even though the significance level was fixed at 0.05, not considering the multiplicity of tests would increase the probability that some of the true null hypotheses were being rejected by chance alone. To address the problem of multiple testing p values were adjusted to identify significant associations. The false discovery rate method proposed by Benjamini and Hochberg was used to calculate adjusted p values for both retrospective and prospective analyses (Benjamini and Yekutieli, 2001). Details of the method are given in Appendix Method 1 (page 352). R software was used to calculate adjusted p values using 'fisheries stock assessment' (FSA) package using the 'False discovery method' (Ogle, & dunnTest, 2019). For sensitivity analysis various other methods were proposed for calculating adjusted p values. (Appendix Method1, page 344)

Comorbidity association with joint specific OA was also explored using the above-mentioned methods. The analysis was restricted to hip, knee, wrist/hand, and ankle/foot because of the higher incidence rate for having enough statistical power. The statistical analyses were performed using STATA statistical software V.15 and R software V3.5.

4.2.8 Sensitivity analysis

For sensitivity analysis, the matched cases and control were who had none of the studied comorbidities on or before the index date. (Appendix Fig 7, page 308) This analysis was performed to study the temporal association of OA only with specific comorbidities. This was because in the main analysis, HR was calculated among people 'at-risk' for that comorbidity only. This makes it difficult to explain the direct association of each comorbidity with OA, which could have been influenced by presence of other conditions. Whereas the sensitivity analysis looked at people without any comorbidities at the index date i.e., OA-only population may provide better interpretation. However, this population might not be true representative of the OA phenotypes.

4.3 Results

4.3.1 Retrospective case-control study

4.3.1.1 Characteristics of the study population

During the period 1997 to 2017, 221,807 incident OA patients were identified with a median age of 61 years at diagnosis (IQR: 52.18-70.43 years) and nearly 58% being women. Individuals with same sex, age ($\pm$2 years) and from the same practice but without OA were selected as matched controls for the 221,807 OA cases. Table 4.3-1 shows characteristics of cases and matched controls. Both unadjusted and adjusted associations of BMI, smoking and alcohol use with OA were significant. Being obese was associated with 2.15 times (95% CI 2.11-2.18) higher risk of developing OA compared to normal weight people. Ex-smokers and current smokers had 10-15% higher risk of developing OA compared to non-smokers. The odd ratio of having OA was 1.15 (95% CI 1.14-1.17) among ex-smokers compared to non-smokers. Details of the comparison of other covariates are given in Table 4.3-1.

Table 4.3-1 Characteristics of incident OA patients and controls from 1997- 2017 at index date

	Incident OA (N=221,807) n(%)	Controls (N=221,807) n(%)	Unadjusted Odds Ratio (95%CI)	Adjusted Odds Ratio[#] (95% CI)
Age (years)				
<40 years	12701(6.07)	13501(5.71)	NA	NA
40-49 years	30813(14.24)	31673(13.86)	NA	NA
50-59 years	60300(27.12)	59606(26.81)	NA	NA
60-69 years	60462(27.19)	59924(26.95)	NA	NA
70-79 years	40891(18.39)	40418(18.18)	NA	NA
80-89 years	15932(7.16)	15815(7.11)	NA	NA
≥90 years	1191(0.53)	1353(0.60)	NA	NA
Gender				
Men	94067(42.31)	94067(42.31)	NA	NA
Women	128223(57.69)	128223(57.69)	NA	NA
BMI (kg/m^2)				
<18.5 (Underweight)	4866(1.39)	3091(2.19)	0.85 (0.82-0.90)*	0.86 (0.82-0.89)*
18.5- 24.9 (Normal)	86872(28.69)	63674(30.09)	Reference	Reference
25.0-29.9 (Overweight)	83188(37.29)	82870(37.42)	1.38 (1.36-1.40)*	1.38 (1.36 -1.40)*
≥30 (Obese)	47373(32.65)	72556(21.31)	2.14 (2.11-2.18)*	2.15 (2.11- 2.18)*
Alcohol consumption (units/week)				
Never	41534(19.90)	44328(18.68)	Reference	Reference
Ex-drinker	5425(2.75)	6099(2.42)	1.04 (1.00-1.08)	1.05 (1.01-1.09)*
Current 1-9	80506(34.96)	77699(36.22)	0.89 (0.88-0.91)*	0.90 (0.88-0.91)*
Current >=10	43282(19.45)	43233(19.47)	0.92 (0.91-0.95)*	0.93 (0.92-0.95)*
Current Unknown	51560(22.95)	51004(23.20)	0.92 (0.91-0.94)*	0.92 (0.91-0.94)*
Smoking Status				
Never smoked	124190(55.87)	117839(53.01)	Reference	Reference
Ex-smoker	40366(18.16)	41812(18.81)	1.15 (1.14-1.17)*	1.15 (1.14-1.17)*
Current smoker	57723(15.97)	62679(28.18)	1.10 (1.08-1.12)*	1.10 (1.08-1.12)*
Mean age (SD)	61.14(13.03)	60.98(13.15)		
Mean BMI (SD)	28.28(5.62)	26.62(4.98)		

#Adjusted by index date and age; *P value < 0.05; NA- not applicable; BMI- body mass index
Mean age (Overall: 60.96 years, SD 13.24 years; Men-60.71 sd-12.78; Women-61.21years SD 13.31).

4.3.1.2 Association of comorbidities

Comorbidities diagnosed prior to the diagnosis of any OA and present before the index date in both case and control groups are shown in Table 4.3-2. Comorbidities diagnosed within 1 year, 5 years, 10 years, 15 years, and 20 years before the index date were analysed. Within the one-year observation period prior to the index date, the prevalence of two or more chronic conditions among cases was 2.69% compared to 1.59% in the control group. This increased to 53.05% and 41.78%, respectively, for

diagnoses recorded during the 20 years before the index date. The longer the observation period, the more prevalent cases were identified. In both groups, leading comorbidities recorded within the 1 year before the index date were back pain (OA 3.44%, non-OA 2.02%), hypertension (OA 2.16%, non-OA 1.77%), high cholesterol (OA 1.37%, non-OA 1.02%), depression (OA 1.34%, non-OA 0.82%) and hearing problems (OA 1.06%, non-OA 0.75%). Within the twenty years before the index date, leading comorbidities recorded were back pain (OA 40.12%, non-OA 29.61%), hypertension (OA 25.60%, non-OA 22.03%), depression (OA 18.32%, non-OA 13.10%), high cholesterol (OA 12.67%, non-OA 10.47%) and hearing problems (OA 9.21%, non-OA 7.46%). (Table 4.3-2)

Table 4.3-3 provides information on the association of comorbidities with incident OA over different time periods (1 year to 20 years). Out of 49 comorbidities studied, within the 1-year time, a significant association was seen with 33, which increased to 39 comorbidities in the 10 years period and 43 comorbidities in the 20 years period. The comorbidities reported within 1 year before the index date with the strongest associations with OA were rheumatoid arthritis (aOR: 3.69; 95% CI 2.90-4.68), fibromyalgia (aOR: 2.77; 95% CI 2.21-3.46), Sjogren's syndrome (aOR: 2.60; 95% CI 1.44-4.69), epilepsy (aOR: 1.89; 95% CI 1.40-2.54), psychosis (aOR: 1.89; 95% CI 1.06- 3.39), and Parkinson's disease (aOR: 1.75; 95% CI 1.33-2.31). Whereas, within 20 years before the index date, the strongest associations were seen with rheumatoid arthritis (aOR: 1.95; 95% CI 1.80-2.11), fibromyalgia (aOR: 1.89; 95% CI 1.75-2.04), polymyalgia (aOR: 1.74; 95% CI 1.62-1.87), back pain (aOR: 1.67; 95% CI 1.64-1.69), Sjogren's syndrome (aOR 1.67; 95% CI 1.39-2.00), systemic lupus erythematous (SLE) (aOR: 1.54; 95% CI 1.15-2.07), ankylosing spondylitis (aOR: 1.53; 95% CI 1.44-1.62), gout (aOR: 1.52; 95% CI 1.46-1.57) and heart failure (aOR: 1.52; 95% CI 1.43-1.62). (Table 4.3-3)

The association of OA with multimorbidity before the index date is depicted in Figure 4.3-1. One year before the index date, the adjusted odds ratio of OA among people with two or more chronic conditions was 1.52 (95% 1.45-1.59) compared to those who had less than 2 comorbidities. The odds ratio for OA in the same groups for the past 20 years prior to the index date was 1.71 (95% CI 1.69-1.74). (Figure 4.3-1)

Associations of comorbidities with joint specific incident OA are shown in Table 4.3-4. For hip joint OA, within 20 years before the index date leading comorbidities having a positive association with hip OA were back pain (aOR 1.66; 95% CI 1.59-1.73), ankylosing spondylitis (aOR 1.62; 95% CI 1.39-1.90), fibromyalgia (aOR 1.51; 95% CI 1.17-1.92), gastro-intestinal bleeding (aOR 1.49; 95% CI 1.23-1.80), polymyalgia (aOR 1.39; 95% CI 1.14-1.69) and depression (aOR 1.32; 95% CI 1.25-1.39).

Leading comorbidities associated with knee OA within 20 years before the index date were musculoskeletal conditions such as fibromyalgia (aOR 1.75; 95% CI 1.49-2.05), polymyalgia (aOR 1.56; 95% CI 1.32-1.77), ankylosing spondylitis (aOR 1.55; 95% CI 1.37-1.73), back pain (aOR 1.51; 95% CI 1.47-1.56), and gout (aOR 1.49; 95% CI 1.39-1.61), as well as depression (aOR 1.46; 95% CI 1.43-1.49) and sleep disorder (aOR 1.44; 95% CI 1.26-1.53).

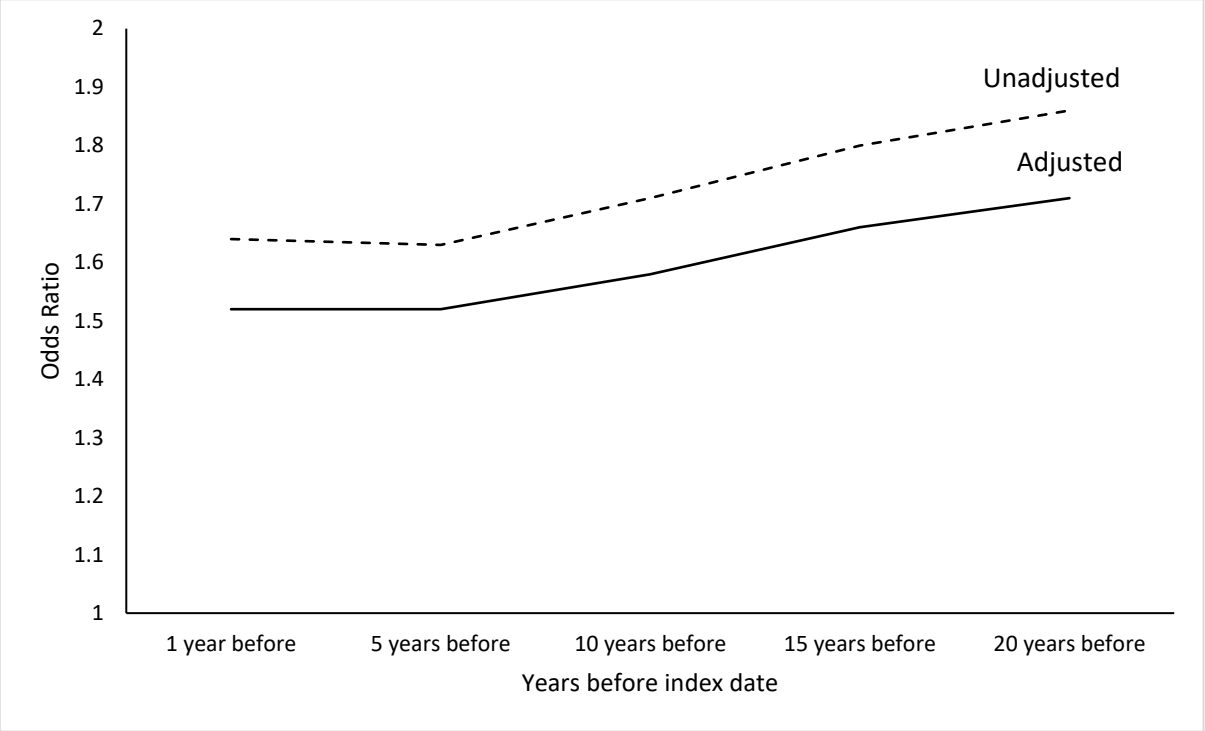
For wrist and hand OA, leading associations were seen with gout (aOR 1.70; 95% CI 1.39-2.08), back pain (aOR 1.58; 95% CI 1.49-1.69), ankylosing spondylitis (aOR 1.57; 95% CI 1.24-1.96), benign prostate hypertrophy (aOR 1.56; 95% CI 1.29-1.89), hypertension (aOR 1.50; 95% CI 1.22-1.86), depression (aOR 1.48; 95% CI 1.34-1.57) and migraine (aOR 1.47; 95% CI 1.30-1.67).

Comorbidities associated with ankle/foot OA within 20 years before the index date were gout (aOR 2.56; 95% CI 2.01-3.14), inflammatory bowel disease (aOR 1.63; 95% CI 1.29-2.06), back pain (aOR 1.59; 95% CI 1.45-1.23), gastritis (aOR 1.45; 95% CI 1.18-1.78), gall bladder stone (aOR 1.45; 95% CI 1.14-1.83), hearing problems (aOR

1.41; 95% CI 1.23-1.67) and benign prostate hypertrophy (aOR 1.40; 95% CI 1.10-1.22).

Details of associations with comorbidities at each 5 years observation period interval are provided in Appendix Table 6 (page 311).

Figure 4.3-1. Association with two or more comorbidities before the index date



Adjusted for age, gender, BMI, Smoking, Alcohol use and index year

Table 4.3-2. Comorbidities in the 1 and 20 years prior to the diagnosis of OA at any joint

	Within 1 year				Within 20 years			
	Controls		OA cases		Controls		OA cases	
	n	%	n	%	n	%	n	%
Musculoskeletal								
Ankylosing spondylitis	132	0.06	215	0.09	2158	1.03	3258	1.55
Back pain	4452	2.02	7632	3.44	61835	29.61	84092	40.12
Gout	493	0.22	749	0.34	4829	2.31	8013	3.82
Osteoporosis	632	0.28	1166	0.52	4896	2.34	6260	2.98
Polymyalgia	170	0.08	323	0.14	1243	0.59	2226	1.06
Rheumatoid arthritis	91	0.04	367	0.16	972	0.46	1956	0.93
Sjogren's syndrome	18	0.01	48	0.02	202	0.09	340	0.16
Systemic lupus erythematosus	5	0.00	11	0.00	81	0.04	122	0.05
Fibromyalgia	115	0.05	404	0.18	1073	0.51	2162	1.03
Fatigue	218	0.09	360	0.16	1739	0.83	2453	1.17
Respiratory								
Asthma	691	0.31	1081	0.48	12320	5.90	17029	8.12
COPD	602	0.27	927	0.42	9296	4.45	12642	6.05
Genito-Urinary								
Chronic kidney disease	1566	0.71	2002	0.90	7527	3.60	8965	4.27
Benign prostatic hypertrophy^	639	0.29	989	0.45	6365	3.05	8436	4.02
Renal stone	114	0.05	158	0.07	1567	0.75	1923	0.91
Neuro/Psychiatric								
Stroke	1354	0.61	1773	0.80	14200	6.80	16158	7.70
Dementia	235	0.11	355	0.16	990	0.47	1068	0.51
Epilepsy	72	0.03	144	0.06	1125	0.54	1376	0.65
Multiple sclerosis	22	0.01	35	0.01	433	0.20	348	0.17
Parkinson's disease	80	0.03	161	0.07	502	0.24	696	0.33
Migraine	500	0.23	745	0.33	8489	4.06	11359	5.41
Depression	1799	0.82	2978	1.34	27362	13.10	38417	18.32
Psychosis	18	0.001	41	0.01	419	0.20	398	0.19
Schizophrenia	51	0.02	84	0.04	1034	0.49	1073	0.51
Cancer	874	0.39	902	0.41	7984	3.80	8972	4.28
Circulatory								
Coronary heart disease	967	0.44	1257	0.56	14262	6.83	18302	8.73
Arterial/Venous	116	0.05	176	0.08	1062	0.51	1429	0.68
Heart failure	289	0.13	444	0.20	1847	0.88	3113	1.48
Hypertension	3906	1.77	4805	2.16	46012	22.03	53659	25.60
Peripheral vascular disease	413	0.18	753	0.34	3906	1.87	5539	2.64
Metabolic								
High cholesterol	2239	1.02	3053	1.37	21865	10.47	26558	12.67
Diabetes mellitus	1397	0.63	1948	0.88	12656	6.06	16147	7.70
Hyperthyroid	137	0.06	142	0.06	1843	0.88	2047	0.97
Hypothyroidism	895	0.40	1203	0.54	9793	4.69	12276	5.85
Digestive								
Gastritis	610	0.28	997	0.45	7551	3.61	10527	5.02
Gastrointestinal bleed	155	0.07	270	0.12	1570	0.75	2253	1.07
Gall stones	533	0.24	660	0.30	6461	3.09	9189	4.38
Inflammatory bowel disease	578	0.26	805	0.36	6409	3.06	8704	4.15
Liver disease	73	0.03	135	0.06	689	0.32	1029	0.49
Irritable bowel syndrome	986	0.44	1421	0.63	10015	4.79	14335	6.83
Others								
Hearing	1666	0.75	2357	1.06	15587	7.46	19315	9.21
Vision problem	130	0.06	136	0.06	1136	0.54	1313	0.62
Psoriasis	277	0.12	439	0.19	3655	1.75	4602	2.19
Scleroderma	2		12		54	0.02	55	0.02
Sleep disorder	481	0.22	724	0.32	3820	1.82	5148	2.45
Tuberculosis	16	0.01	32	0.01	342	0.16	417	0.19
Anaemia	588	0.26	920	0.41	5406	2.59	6732	3.21
Comorbidities (count)								
No comorbidity	195859	88.10	184311	82.91	77845	35.01	59752	26.88
Single comorbidity	22891	10.29	31971	14.38	51546	23.18	44541	20.03
Any two comorbidities	3058	1.37	5042	2.26	38897	17.49	41327	18.59
Any three comorbidities	415	0.19	787	0.35	25282	11.37	31429	14.14
Four or more	67	0.03	179	0.08	28720	12.92	45241	20.35

COPD- Chronic Obstructive Pulmonary Disease; ^only for men Expanded version of this table is available at Appendix Table 5 (page 304).

Table 4.3-3. Association between any OA and comorbidities in different time periods prior to the index date

	20 years	10 years	5 years	1 year
	Adjusted OR[#]	**Adjusted OR[#]**	**Adjusted OR[#]**	**Adjusted OR[#]**
>= 2 comorbidities	1.71 (1.69-1.74)*	1.58 (1.56-1.60)*	1.53 (1.49-1.55)*	1.52(1.45-1.59)*
Musculoskeletal				
Ankylosing spondylitis	1.53 (1.44-1.62)*	1.63 (1.52-1.75)*	1.63 (1.47-1.79)*	1.49 (1.19-1.86)*
Back pain	1.67 (1.64-1.69)*	1.51 (1.48-1.53)*	1.45 (1.43-1.48)*	1.60 (1.54-1.69)*
Gout	1.52 (1.46-1.57)*	1.52 (1.45-1.59)*	1.49 (1.41-1.58)*	1.26 (1.11-1.42)*
Osteoporosis	1.41 (1.35-1.47)*	1.42 (1.36-1.49)*	1.49 (1.42-1.58)*	1.74 (1.57-1.93)*
Polymyalgia	1.74 (1.62-1.87)*	1.86 (1.72-2.01)*	1.86 (1.69-2.05)*	1.71 (1.41-2.08)*
Rheumatoid arthritis	1.95 (1.80-2.11)*	2.17 (1.98-2.38)*	2.50 (2.21-2.82)*	3.69 (2.90-4.68)*
Sjogren's syndrome	1.67 (1.39-2.00)*	1.94 (1.56-2.40)*	2.47 (1.85-3.30)*	2.60 (1.44-4.69)
Systemic lupus erythematosus	1.54 (1.15-2.07)	1.59 (1.09-2.29)	1.72 (1.05-2.82)	2.31 (0.76-7.05)
Fibromyalgia	1.89 (1.75-2.04)*	2.07 (1.89-2.25)*	2.19 (1.96-2.45)*	2.77 (2.21-3.46)*
Fatigue	1.42 (1.32-1.51)*	1.46 (1.36-1.57)*	1.48 (1.36-1.62)*	1.56 (1.30-1.86)*
Respiratory				
Asthma	1.33 (1.30-1.37)*	1.35 (1.31-1.39)*	1.37 (1.31-1.43)*	1.36 (1.23-1.51)*
COPD	1.35 (1.31-1.39)*	1.36 (1.31-1.41)*	1.37 (1.30-1.43)*	1.42 (1.28-1.58)*
Genito-Urinary				
Chronic kidney disease	1.12 (1.08-1.16)*	1.12(1.08-1.16)*	1.15 (1.10-1.19)*	1.16 (1.08-1.24)*
Benign prostatic hypertrophy	1.38 (1.33-1.43)*	1.37 (1.32-1.43)*	1.37 (1.32-1.46)*	1.37 (1.24-1.53)*
Renal stone	1.16 (1.09-1.25)*	1.21 (1.11-1.32)*	1.21 (1.08-1.36)*	1.31 (1.02-1.68)
Neuro/Psychiatric				
Stroke	1.15 (1.11-1.19)*	1.15 (1.12-1.19)*	1.17 (1.13-1.22)*	1.24 (1.15-1.34)*
Dementia	1.09 (0.99-1.19)	1.13 (1.03-1.24)	1.23 (1.11-1.36)*	1.44 (1.21-1.71)*
Epilepsy	1.18 (1.08-1.29)*	1.24 (1.11-1.37)*	1.17 (1.03-1.35)	1.89 (1.40-2.54)*
Multiple sclerosis	0.80 (0.69-0.93)*	0.95 (0.78-1.14)	0.95 (0.72-1.20)	1.55 (0.89-2.67)
Parkinson's disease	1.39 (1.23-1.57)*	1.39 (1.22-1.57)*	1.47 (1.27-1.70)*	1.75 (1.33-2.31)*
Migraine	1.37 (1.33-1.41)*	1.42 (1.36-1.47)*	1.44 (1.37-1.53)*	1.40 (1.25-1.59)*
Depression	1.49 (1.46-1.52)*	1.49 (1.46-1.52)*	1.49 (1.45-1.54)*	1.51 (1.42-1.61)*
Psychosis	0.86 (0.75-1.00)	0.83 (0.69-0.98)	0.95 (0.75-1.19)	1.89 (1.06-3.39)*
Schizophrenia	0.95 (0.87-1.04)	0.97 (0.87-1.08)	1.08 (0.92-1.26)	1.36 (0.95-1.96)
Cancer	1.12 (1.09-1.16)*	1.12 (1.08-1.17)*	1.12 (1.08-1.18)*	0.96 (0.87-1.05)
Circulatory				
Coronary heart disease	1.24 (1.21-1.27)*	1.22 (1.18-1.25)*	1.17 (1.12-1.21)*	1.12 (1.03-1.23)
Arterial/Venous	1.29 (1.19-1.41)*	1.30 (1.19-1.43)*	1.35 (1.20-1.52)*	1.41 (1.10-1.81)
Heart failure	1.52 (1.43-1.62)*	1.52 (1.43-1.63)*	1.53 (1.41-1.65)*	1.30 (1.11-1.52)*
Hypertension	1.08 (1.06-1.10)*	1.06 (1.04-1.07)*	1.04 (1.02-1.06)*	1.03 (0.98-1.08)
Peripheral vascular disease	1.45 (1.39-1.51)*	1.51 (1.44-1.59)*	1.54 (1.45-1.64)*	1.62 (1.43-1.84)*
Metabolic/Endocrine				
High cholesterol	1.18 (1.16-1.20)*	1.18 (1.15-1.21)*	1.20 (1.16-1.23)*	1.20 (1.13-1.28)*
Diabetes mellitus	1.06 (1.03-1.09)*	1.06 (1.02-1.10)*	1.06 (1.02-1.09)*	1.12 (1.04-1.20)*
Hyperthyroid	1.09 (1.02-1.16)*	1.05 (0.97-1.14)	1.04 (0.93-1.15)	0.92 (0.71-1.17)
Hypothyroidism	1.18 (1.15-1.22)*	1.17 (1.12-1.20)*	1.16 (1.11-1.21)*	1.19 (1.08-1.30)*
Digestive				
Gastritis	1.42 (1.36-1.45)*	1.45 (1.39-1.50)*	1.45 (1.38-1.52)*	1.55 (1.39-1.72)*
Gastrointestinal bleed	1.42 (1.33-1.52)*	1.44 (1.33-1.56)*	1.49 (1.34-1.64)*	1.66 (1.36-2.03)*
Gall bladder stone	1.27 (1.22-1.31)*	1.26 (1.21-1.31)*	1.23 (1.17-1.30)*	1.05 (0.93-1.18)
Inflammatory bowel disease	1.36 (1.32-1.41)*	1.42 (1.36-1.47)*	1.44 (1.36-1.52)*	1.33 (1.19-1.48)*
Liver disease	1.42 (1.29-1.57)*	1.48 (1.32-1.67)*	1.45 (1.26-1.68)*	1.47 (1.09-1.99)
Irritable bowel syndrome	1.52(1.47-1.56)*	1.55(1.49-1.60)*	1.58(1.51-1.66)*	1.59(1.23-1.95)*
Others				
HIV infection/AIDS	2.08 (0.76-5.75)	1.49 (0.54-4.16)	3.17 (0.84-12.03)	-
Hearing	1.26 (1.23-1.29)*	1.26 (1.22-1.29)*	1.26 (1.22-1.30)*	1.30 (1.22-1.39)*
Psoriasis	1.20 (1.14-1.25)*	1.24 (1.17-1.31)*	1.26 (1.17-1.36)*	1.32 (1.12-1.55)*
Scleroderma	0.97 (0.65-1.44)	1.02 (0.64-1.64)	1.76 (0.94-3.30)	5.75 (1.22-22.09)
Sleep disorder	1.35 (1.28-1.41)*	1.37 (1.30-1.44)*	1.37 (1.28-1.46)*	1.41 (1.24-1.59)*
Tuberculosis	1.25 (1.08-1.45)	1.24 (1.04-1.50)	1.23 (0.95-1.59)	1.71 (0.91-3.18)
Anaemia	1.25 (1.21-1.30)*	1.31(1.26-1.37)*	1.40(1.32-1.48)*	1.42(1.28-1.59)*
Vision problem	1.11 (1.02-1.21)	1.13 (1.03-1.24)	1.17 (1.03-1.32)	0.96 (0.74-1.23)

*P value <0.01 adjusted for multiple testing using 'False discovery rate'.

[#]Adjusted for age, gender, BMI, Smoking, Alcohol, multimorbidity and index year ^Only for men.

COPD- Chronic Obstructive Pulmonary Disease

For expanded version of this table, please refer to Appendix Table 6. (page 311)

Table 4.3-4. Association between joint specific OA and comorbidities diagnosed in the 20 years prior to the index date

	Hip	Knee	Wrist/Hand	Ankle/Foot
	20 years	20 years	20 years	20 years
Musculoskeletal				
Ankylosing Spondylitis	1.62(1.39-1.90)*	1.55(1.37-1.73)*	1.57(1.24-1.96)*	1.40(0.96-2.00)
Back pain	1.66(1.59-1.73)*	1.51(1.47-1.56)*	1.58(1.49-1.69)*	1.59(1.45-1.73)*
Gout	1.21(1.09-1.35)*	1.49(1.39-1.61)*	1.70(1.39-2.08)*	2.56(2.01-3.14)*
Osteoporosis	1.30(1.16-1.46)*	1.25(1.13-1.34)*	1.26(1.05-1.53)	1.34(1.04-1.85)
Polymyalgia	1.39(1.14-1.69)*	1.56(1.32-1.77)*	1.58(1.07-2.35)	1.38(0.81-2.37)
Rheumatoid Arthritis	1.25(0.99-1.63)	1.43(1.21-1.70)*	1.57(0.99-1.99)	1.30(0.62-1.72)
Sjogren's syndrome	1.93(1.08-3.47)	1.47(1.04-2.09)	1.32(0.63-2.74)	1.30(0.32-5.22)
SLE	-	1.19(0.62-2.29)	0.38(0.09-1.38)	-
Fibromyalgia	1.51(1.17-1.92)	1.75(1.49-2.05)*	1.53(1.14-2.07)	1.29(0.81-2.03)
Fatigue	1.32(1.09-1.60)	1.38(1.21-1.59)*	1.42(1.09-1.84)	1.10(0.66-1.53)
Respiratory				
Asthma	1.19(1.11-1.28)*	1.38(1.31-1.46)*	1.31(1.18-1.47)*	1.38(1.18-1.62)*
COPD	1.20(1.11-1.31)*	1.33(1.25-1.41)*	1.23(1.07-1.41)	1.25(1.02-1.52)
Genito-Urinary				
CKD	1.10(0.99-1.21)	1.04(0.97-1.13)	0.92(0.77-1.10)	1.05(0.82-1.34)
Benign prostatic hypertrophy^	1.40(1.27-1.55)*	1.32(1.25-1.43)*	1.56(1.29-1.89)*	1.40(1.10-1.72)*
Renal stone	1.05(0.87-1.28)	1.41(1.15-1.51)*	1.22(0.86-1.73)	1.60(1.07-2.39)
Neuro/Psychiatric				
Stroke	1.09(1.00-1.16)	1.20(1.10-1.22)*	1.24(1.09-1.40)*	1.17(0.90-1.28)
Dementia	1.11(0.86-1.44)	0.93(0.77-1.12)	0.72(0.44-1.17)	0.96(0.45-2.01)
Epilepsy	1.27(0.99-1.61)	1.29(1.09-1.51)	1.12(0.76-1.66)	0.81(0.45-1.44)
Multiple sclerosis	0.73(0.48-1.10)	0.96(0.70-1.29)	0.67(0.34-1.32)	0.63(0.25-1.61)
Parkinson's Disease	0.87(0.62-1.23)	1.20(0.96-1.52)	1.11(0.55-2.24)	1.49(0.66-3.35)
Migraine	1.16(1.06-1.28)*	1.35(1.27-1.44)*	1.47(1.30-1.67)*	1.38(1.14-1.67)*
Depression	1.32(1.25-1.39)*	1.46(1.43-1.49)*	1.48(1.34-1.57)*	1.40(1.27-1.60)*
Psychosis	1.09(0.70-1.71)	1.05(0.82-1.44)	0.58(0.30-1.12)	
Schizophrenia	0.99(0.77-1.29)	1.15(0.96-1.38)	0.83(0.55-1.25)	0.56(0.38-1.13)
Cancer	1.24(1.13-1.35)*	1.11(1.03-1.17)	0.92(0.79-1.07)	1.16(0.93-1.45)
Circulatory				
CHD	1.18(1.10-1.26)*	1.15(1.09-1.21)*	1.02(0.90-1.16)	1.38(1.17-1.63)*
Arterial/Venous	1.34(1.09-1.65)	1.21(1.03-1.42)	0.96(0.63-1.52)	0.84(0.47-1.51)
Heart failure	1.38(1.16-1.63)*	1.34(1.20-1.53)*	1.17(0.89-1.83)	1.57(1.03-2.38)
Hypertension	1.12(1.07-1.17)*	1.10(1.07-1.15)*	1.02(0.95-1.10)	1.08(0.97-1.20)
PVD	1.37(1.21-1.55)*	1.29(1.16-1.38)*	1.50(1.22-1.86)*	1.44(1.09-1.89)
Metabolic/Endocrine				
High Cholesterol	1.15(1.09-1.22)*	1.14(1.09-1.19)*	1.22(1.11-1.35)*	1.11(1.01-1.33)
Diabetes Mellitus	1.06(0.98-1.13)	1.02(0.98-1.08)	0.97(0.85-1.10)	0.95(0.79-1.13)
Hyperthyroid	1.13(0.94-1.38)	1.13(0.99-1.30)	1.05(0.79-1.39)	1.15(0.72-1.85)
Hypothyroidism	1.23(1.13-1.34)*	1.17(1.10-1.24)*	1.21(1.07-1.38)*	1.11(0.91-1.37)
Digestive				
Gastritis	1.22(1.11-1.34)*	1.39(1.30-1.47)*	1.26(1.09-1.45)*	1.45(1.18-1.78)*
Gastrointestinal bleed	1.49(1.23-1.80)*	1.37(1.21-1.56)*	1.35(0.99-1.83)	1.48(0.96-2.31)
Gall bladder stone	1.22(1.11-1.35)*	1.33(1.25-1.43)*	1.31(1.13-1.52)*	1.45(1.14-1.83)*
IBD	1.23(1.11-1.37)*	1.35(1.26-1.44)*	1.22(1.04-1.40)	1.63(1.29-2.06)*
Liver Disease	1.14(0.85-1.55)	1.32(1.08-1.62)	1.18(0.71-1.96)	1.51(0.74-3.06)
Irritable bowel syndrome	1.33(1.22-1.46)*	1.38(1.29-1.48)*	1.25(1.08-1.42)	1.61(1.24-2.02)*
Others				
HIV infection/AIDS	0.86(0.13-5.86)	2.05(0.20-20.69)	-	-
Hearing	1.14(1.06-1.21)*	1.24(1.18-1.29)*	1.31(1.18-1.46)*	1.41(1.23-1.67)*
Psoriasis	1.07(0.93-1.22)	1.13(1.04-1.25)	1.07(0.86-1.29)	1.15(1.01-1.81)
Scleroderma	1.29(0.33-5.06)	072(0.26-1.52)	-	
Sleep Disorder	1.25(1.-8-1.43)	1.44(1.26-1.53)*	1.44(1.15-1.78)*	1.48(1.09-2.02)
Tuberculosis	0.86(0.56-1.32)	1.35(1.00-1.84)	3.44(1.23-9.58)	2.56(0.93-7.07)
Anaemia	1.21(1.08-1.36)*	1.26(1.16-1.35)*	1.31(1.10-1.53)	1.14(0.87-1.49)
Vision problem	1.05(0.83-1.33)	1.11(0.95-1.31)	1.27(0.85-1.90)	0.77(0.39-1.51)

*P value <0.05 adjusted for multiple testing using 'False discovery rate'.

Adjusted for age, gender, body mass index (BMI), Smoking, Alcohol use and index year; ^for men only

SLE – Systemic Lupus Erythematous; COPD- Chronic Obstructive Pulmonary Disease; CHD- Coronary Heart Disease; PVD- Peripheral vascular disease; IBD- Inflammatory Bowel Disease

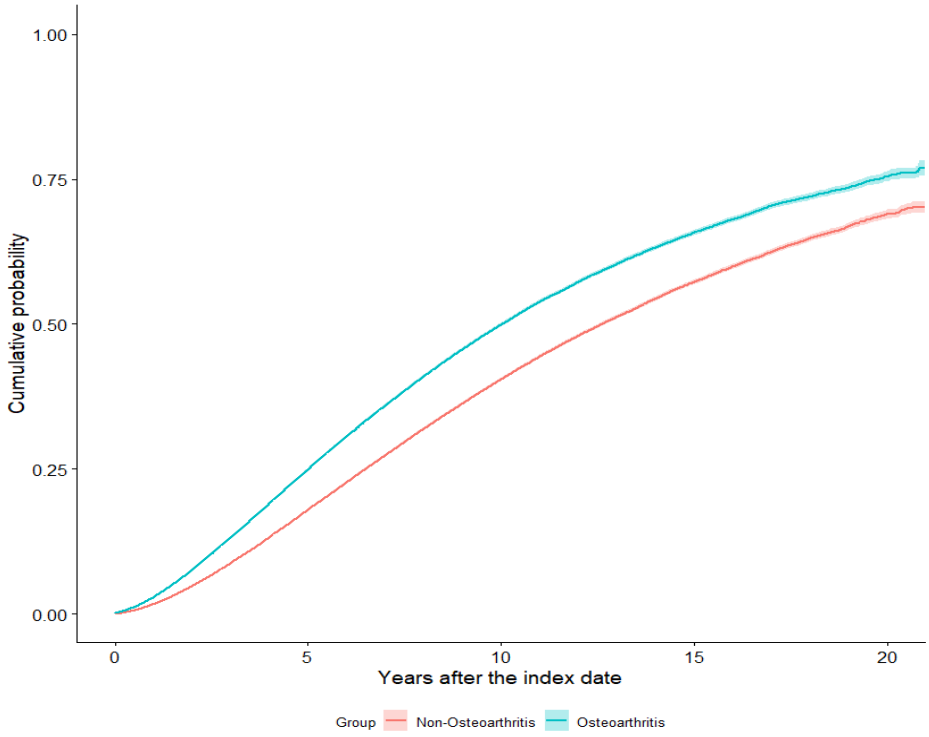
For expanded version of this table, please refer to Appendix Table 7.(page 312)

4.3.2 Comorbidities diagnosed after incident OA (prospective analysis)

The adjusted cumulative probabilities of having multimorbidity at 5, 15 and 20 years following the index date were 27.3%, 68.4% and 77.4% in people with incident OA and 19.5%, 42.9% and 70.7% in controls, respectively (Figure 4.3-2). The adjusted HR (aHR) for incident multimorbidity was 1.29 (95% CI 1.28-1.31) in OA cases compared with controls (Table 4.3-6) (log-rank test, p<0.001). The median time to develop any two comorbidities among patients with OA and matched controls was 7.15 years (IQR 3.60-11.36) and 8.90 years (IQR 4.84-12.92) respectively.

The cumulative probabilities of all comorbidities were higher in the OA group than the control group in each year of follow-up. (Table 4.3-5) Table 4.3-5 shows the cumulative probabilities of specific comorbidities in incident OA cases and controls diagnosed within 1 year, 5 years, 10 years, 15 years, and 20 years after the index date. The cumulative risk of all comorbidities was higher in incident OA cases than matched controls (p<0.05).

Figure 4.3-2 Cumulative probabilities of having additional multimorbidity after the index date



OA: Osteoarthritis; Non-OA: Non-Osteoarthritis

Table 4.3-5. Cumulative probabilities (%) of incident comorbidities after index date

	OA (years)					Non-OA (Years)				
	1	**5**	**10**	**15**	**20**	**1**	**5**	**10**	**15**	**20**
Additional multimorbidity	18.5	27.3	53.00	68.37	77.40	10.3	19.5	42.91	59.56	70.73
Musculoskeletal										
Ankylosing spondylitis	0.07	0.43	0.81	1.14	1.41	0.06	0.25	0.50	0.72	0.98
Back pain	5.12	20.5	34.30	44.18	50.89	3.07	14.0	25.35	34.30	41.12
Gout	0.36	2.12	4.42	6.73	8.63	0.25	1.27	2.67	4.24	5.91
Osteoporosis	0.57	2.41	4.98	7.86	10.52	0.31	1.84	4.24	7.15	10.15
Polymyalgia	0.22	0.71	1.39	2.07	2.94	0.08	0.42	0.94	1.52	2.42
Rheumatoid arthritis	0.24	0.78	1.40	1.96	2.51	0.04	0.18	0.37	0.54	0.73
Sjogren's syndrome	0.02	0.09	0.16	0.26	0.26	0.01	0.04	0.09	0.12	0.23
Systemic lupus erythematosus	0.01	0.03	0.06	0.08	0.10	0.00	0.01	0.03	0.04	0.05
Fibromyalgia	0.20	0.74	1.24	1.66	2.03	0.04	0.20	0.38	0.53	0.69
Fatigue	0.15	0.78	1.53	2.27	3.00	0.10	0.53	1.97	1.57	2.08
Respiratory										
Asthma	0.41	1.89	3.37	4.61	5.94	0.33	1.45	2.53	3.43	4.29
COPD	0.42	1.99	4.05	6.16	8.64	0.31	1.52	3.33	5.34	7.07
Genito-Urinary										
Chronic kidney disease	1.08	6.33	14.87	20.27	24.43	0.83	5.31	12.80	17.89	22.40
Benign prostatic hypertrophy^	0.45	2.09	3.92	5.56	7.18	0.30	1.59	3.12	4.54	5.75
Renal stone	0.07	0.36	0.76	1.12	1.59	0.06	0.28	0.59	0.96	1.34
Neuro/Psychiatric										
Stroke	0.81	4.05	8.29	12.75	16.98	0.65	3.25	6.92	10.79	14.93
Dementia	0.23	1.50	3.94	7.30	11.00	0.12	0.98	2.85	5.71	9.22
Epilepsy	0.05	0.25	0.50	0.75	1.17	0.04	0.18	0.38	0.57	0.70
Multiple sclerosis	0.01	0.05	0.10	0.12	0.16	0.01	0.05	0.08	0.11	0.12
Parkinson's disease	0.08	0.39	0.81	1.19	1.58	0.04	0.24	0.55	0.94	1.50
Migraine	0.32	1.36	2.37	3.19	4.11	0.22	1.00	1.78	2.38	2.85
Depression	1.67	6.95	11.70	15.73	19.43	0.99	4.29	7.64	10.50	13.33
Psychosis	0.01	0.09	0.18	0.28	0.38	0.02	0.09	0.17	0.26	0.35
Schizophrenia	0.04	0.19	0.36	0.51	0.66	0.03	0.16	0.30	0.43	0.53
Cancer	0.82	4.37	9.54	14.81	20.07	0.45	2.73	6.37	10.48	14.90
Circulatory										
Coronary heart disease	0.69	3.23	6.12	8.70	11.29	0.47	2.37	4.52	6.68	8.48
Arterial/Venous	0.10	0.54	1.15	1.89	2.55	0.06	0.34	0.84	1.38	1.95
Heart failure	0.30	1.47	2.89	4.43	5.92	0.12	0.73	1.64	2.73	3.91
Hypertension	2.83	11.8	21.01	27.99	33.75	2.17	10.2	18.90	25.59	31.54
PVD	0.34	1.47	2.88	4.26	5.52	0.19	0.98	2.00	3.07	3.91
Metabolic/Endocrine										
High cholesterol	1.52	7.00	12.59	16.95	19.36	1.17	5.76	10.82	14.71	17.77
Diabetes Mellitus	1.03	5.28	11.18	17.30	23.19	0.77	3.94	8.43	13.58	18.98
Hyperthyroid	0.09	0.37	0.71	1.00	1.29	0.07	0.31	0.58	0.84	1.10
Hypothyroidism	0.54	2.41	4.46	6.26	7.45	0.41	2.02	3.92	5.53	6.79
Digestive										
Gastritis	0.53	2.45	4.77	7.01	9.23	0.27	1.41	3.00	4.60	6.24
Gastrointestinal bleed	0.13	0.70	1.42	2.09	2.65	0.07	0.40	0.83	1.35	1.94
Gall bladder stone	0.36	1.96	3.91	5.96	7.66	0.27	1.33	2.72	4.10	5.49
Inflammatory bowel disease	0.41	1.97	3.87	5.44	6.85	0.27	1.31	2.58	3.84	4.60
Liver disease	0.06	0.31	0.64	1.05	1.40	0.03	0.17	0.39	0.57	0.79
Irritable bowel syndrome	0.90	2.00	3.48	4.63	5.66	0.40	1.33	2.31	3.13	3.85
Others										
HIV infection/AIDS	0.00	0.00	0.0001	0.0001	0.0001	0.00	0.00	0.00	0.00	0.00
Hearing	1.06	5.73	11.74	17.82	24.18	0.89	4.74	10.17	16.06	21.59
Psoriasis	0.18	0.70	1..28	1.81	2.28	0.12	0.57	1.05	1.47	1.77
Scleroderma	0.01	0.02	0.04	0.06	0.11	0.00	0.02	0.03	0.05	0.06
Sleep Disorder	0.36	1.59	3.06	4.38	5.58	0.24	1.11	2.04	2.95	3.78
Tuberculosis	0.02	0.06	0.12	0.17	0.25	0.01	0.05	0.09	0.11	0.14
Anaemia	0.57	2.70	5.45	8.22	11.04	0.29	1.56	3.53	5.65	7.53
Vision problem	0.07	0.37	0.79	1.15	1.57	0.05	0.28	0.68	1.11	1.53
Cataract	1.45	4.74	9.87	15.30	20.45	1.20	4.18	9.10	14.35	19.20

COPD- Chronic Obstructive Pulmonary Disease; ^ only for men; PVD -Peripheral vascular diseases

4.3.2.1 Relative risk of developing incident comorbidities

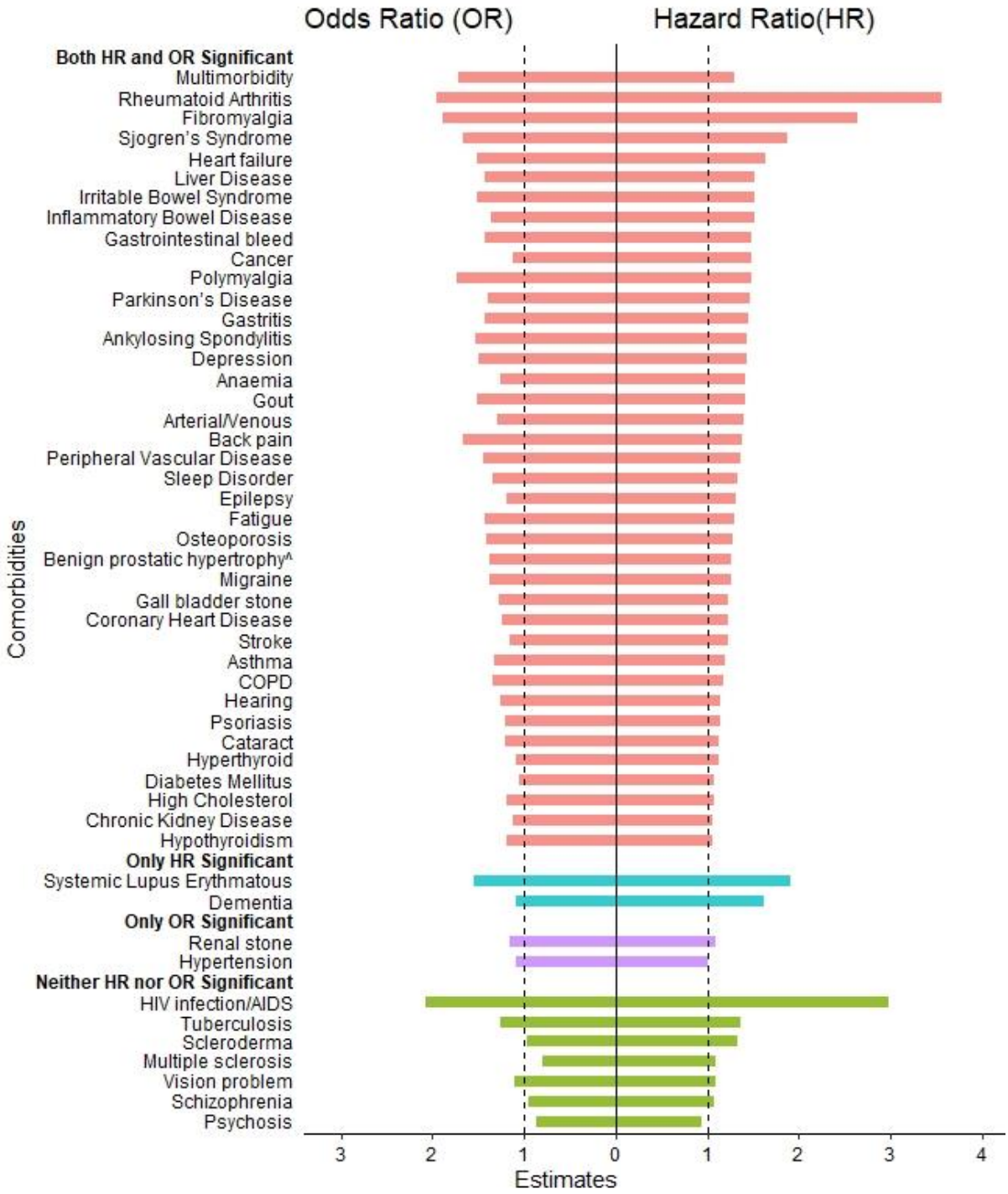
Except for HIV/AIDS, psychosis, multiple sclerosis, tuberculosis, scleroderma, vision problem, schizophrenia, hypertension, and renal stones, the risks of developing each of the other comorbidities were significantly higher in people with OA. (Table 4.3-6) Patients with OA were over three times more likely to develop rheumatoid arthritis (aHR 3.56; 95% CI 3.26-3.89) and 2.6 times more likely to develop fibromyalgia (aHR 2.64; 95% CI 2.41-2.89) than matched controls. Besides musculoskeletal conditions people with OA had significantly higher risk compared to matched controls of developing heart failure (aHR 1.63; 95% CI 1.56-1.71), dementia (aHR 1.62; 95% CI 1.56-1.68), liver diseases (aHR 1.51; 95% CI 1.37-1.67), irritable bowel syndrome (IBS) (aHR 1.51; 95% CI 1.45-1.58), gastro-intestinal bleeding (aHR 1.49; 95% CI 1.39-1.59), cancer (aHR 1.49; 95% CI 1.46-1.53), Parkinson's disease (aHR 1.46; 95% CI 1.34-1.59), gastritis (aHR 1.45; 95% CI 1.40-1.51), depression (aHR 1.43; 95% CI 1.39-1.47), anaemia (aHR 1.42; 95% CI 1.37-1.47), and peripheral vascular diseases (aHR 1.36; 95% CI 1.30-1.43).

Table 4.3-6. Hazard ratio and 95% confidence interval for each comorbidity comparing incident OA cases and controls (Time varying covariate cox regression)

	Controls at-risk (Incidence per 1000 p-ys)	**Cases at-risk (Incidence per 1000 p-ys)**	Unadjusted HR (95% CI)	Adjusted HR (95% CI)
Additional multimorbidity	77695(6.76)	74111(5.12)	1.37(1.36-1.39)	1.29(1.28-1.30)*
Musculoskeletal				
Ankylosing spondylitis	218496 (0.8)	217711 (0.48)	1.63(1.49-1.77)	1.44(1.32-1.58)*
Back pain	117392 (42.82)	144323 (28.99)	1.45(1.43-1.47)	1.38(1.36-1.41)*
Gout	213278 (4.46)	214843 (2.77)	1.63(1.57-1.69)	1.41(1.35-1.46)*
Osteoporosis	215723 (5.21)	215211 (4.47)	1.19(1.15-1.23)	1.28(1.24-1.32)*
Polymyalgia	219904 (1.43)	218863 (0.9)	1.49(1.40-1.59)	1.48(1.39-1.58)*
Rheumatoid arthritis	219874 (1.42)	219077 (0.36)	3.82(3.50-4.17)	3.56(3.26-3.89)*
Sjogren's disease	221805 (0.16)	219902 (0.08)	2.01(1.64-2.46)	1.87(1.52-2.29)*
Systemic lupus erythematous	222027 (0.06)	220031 (0.02)	2.14(1.52-3.01)	1.90(1.34-2.69)*
Fibromyalgia	219834 (1.28)	218978 (0.37)	3.32(3.04-3.63)	2.64(2.41-2.89)*
Fatigue	219556 (1.54)	218276 (1.06)	1.45(1.36-1.54)	1.30(1.22-1.38)*
Respiratory				
Asthma	197561 (3.5)	201834 (2.53)	1.35(1.29-1.40)	1.20(1.15-1.25)*
COPD	207583 (4.13)	209489 (3.42)	1.22(1.17-1.26)	1.18(1.14-1.22)*
Genito-Urinary				
Chronic Kidney Disease	212998 (1.46)	212652 (1.26)	1.17(1.15-1.19)	1.06(1.04-1.08)*
Benign prostatic hypertrophy^	213434 (4.01)	213577 (3.13)	1.27(1.22-1.32)	1.27(1.22-1.32)*
Renal stone	219574 (0.74)	217980 (0.6)	1.25(1.15-1.36)	1.10(1.01-1.19)
Neuro/Psychiatric				
Stroke	204629 (8.68)	204936 (7.26)	1.21(1.18-1.24)	1.22(1.19-1.26)*
Dementia	221101 (4.05)	219204 (3.18)	1.36(1.32-1.42)	1.62(1.56-1.68)*
Epilepsy	219002 (0.51)	217678 (0.37)	1.39(1.25-1.54)	1.31(1.18-1.46)*
Multiple sclerosis	221632 (0.09)	219473 (0.07)	1.18(0.93-1.49)	1.09(0.86-1.39)
Parkinson's Disease	221470 (0.79)	219635 (0.58)	1.41(1.29-1.53)	1.46(1.34-1.59)*
Migraine	205856 (2.44)	208048 (1.74)	1.36(1.29-1.43)	1.26(1.20-1.33)*
Depression	170180 (12.86)	182837 (7.92)	1.58(1.54-1.62)	1.43(1.39-1.47)*
Psychosis	221619 (0.19)	219562 (0.17)	1.10(0.93-1.29)	0.94(0.79-1.10)
Schizophrenia	220303 (0.36)	218301 (0.29)	1.21(1.07-1.36)	1.08(0.96-1.22)
Cancer	212110 (9.87)	211362 (6.72)	1.50(1.47-1.54)	1.49(1.46-1.53)*
Circulatory				
Coronary Heart Disease	201870 (6.32)	204490 (4.6)	1.35(1.31-1.39)	1.22(1.18-1.26)*
Arterial/Venous	220674 (1.17)	219035 (0.84)	1.43(1.33-1.53)	1.39(1.30-1.49)*
Heart failure	219010 (2.92)	218309 (1.69)	1.74(1.66-1.83)	1.63(1.56-1.71)*
Hypertension	161900 (23.68)	169134 (20.58)	1.13(1.11-1.15)	1.01(0.99-1.03)
Peripheral vascular disease	216126 (2.93)	215876 (2.02)	1.45(1.38-1.51)	1.36(1.30-1.43)*
Metabolic/Endocrine				
High Cholesterol	194351 (1.34)	197519 (1.11)	1.18(1.16-1.21)	1.08(1.05-1.10)*
Diabetes Mellitus	204495 (11.83)	206477 (9.05)	1.33(1.30-1.36)	1.08(1.06-1.11)*
Hyperthyroid	219061 (0.7)	217505 (0.57)	1.21(1.11-1.32)	1.12(1.03-1.22)*
Hypothyroidism	208088 (4.59)	209156 (3.9)	1.16(1.12-1.20)	1.06(1.02-1.09)*
Digestive				
Gastritis	207695 (4.94)	209676 (3.05)	1.62(1.57-1.68)	1.45(1.40-1.51)*
Gastrointestinal bleed	219414 (1.4)	218162 (0.85)	1.65(1.54-1.76)	1.49(1.39-1.59)*
Gall bladder stone	209651 (4.0)	211412 (2.76)	1.45(1.40-1.51)	1.23(1.18-1.28)*
Inflammatory bowel Disease	211501 (3.89)	212175 (2.59)	1.49(1.45-1.55)	1.31(1.26-1.37)*
Liver Disease	220977 (0.65)	219294 (0.38)	1.74(1.58-1.92)	1.51(1.37-1.67)*
Irritable bowel syndrome	222101 (3.49)	222145 (2.45)	1.50(1.44-1.56)	1.51(1.45-1.58)*
Others				
HIV infection/AIDS	222161 (0.001)	220123 (0.001)	3.79(1.23-11.65)	2.98(0.95-9.37)
Hearing	200102 (12.48)	202329 (10.92)	1.16(1.13-1.19)	1.14(1.11-1.16)*
Psoriasis	215401 (1.3)	214766 (1.03)	1.23(1.15-1.31)	1.14(1.06-1.21)*
Scleroderma	222097 (0.03)	220060 (0.02)	1.50(1.05-21.3)	1.33(0.93-1.92)
Sleep Disorder	216765 (3.11)	216231 (2.06)	1.49(1.43-1.56)	1.33(1.27-1.39)*
Tuberculosis	220697 (0.1)	218804 (0.08)	1.45(1.16-1.79)	1.36(1.09-1.69)
Anaemia	214130 (5.62)	213681 (3.62)	1.57(1.52-1.62)	1.42(1.37-1.47)*
Vision problem	220721 (7.62)	218929 (6.89)	1.12(1.03-1.21)	1.09(1.00-1.18)
Cataract	222200 (10.35)	222215 (9.63)	1.09(1.07-1.12)	1.13(1.10-1.16)*

Adjusted for age, sex, BMI, alcohol, smoking status, multimorbidity count and index date; p-y person years.
**p*-value <0.05 'False discovery rate' (FDR) adjusted; COPD- Chronic Obstructive Pulmonary Disease

Figure 4.3-3. Comparison of Odds Ratio and Hazard Ratio for comorbidities in OA for 20 years observation period



COPD- Chronic obstructive pulmonary diseases; *p-value <0.05; ^Only for men

Figure 4.3-3 depicts the comparison of adjusted ORs and HRs for comorbidities in OA. It shows that people who had musculoskeletal or other pain-related conditions before diagnosis of OA/index date are more likely to develop OA, and OA people are also more likely to develop other pain-related conditions after the diagnosis. Also, the HR of

diabetes, chronic kidney disease, schizophrenia, respiratory diseases, stroke, coronary heart disease, cancer, gastritis, dementia, gastro-intestinal bleeding, depression, sleep disorders and IBD were higher than the respective ORs. (Figure 4.3-3)

Table 4.3-7. Adjusted hazard ratio and 95% confidence interval for each comorbidity comparing incident OA cases (joint wise)

	Hip Adjusted HR (95% CI)	**Knee** Adjusted HR (95% CI)	**Wrist/Hand** Adjusted HR (95% CI)	**Ankle/Foot** Adjusted HR (95% CI)
>=2 comorbidities	1.16(1.11-1.21)*	1.24(1.20-1.28)*	1.46(1.36-1.56)*	1.17(1.07-1.29)*
Musculoskeletal				
Ankylosing Spondylitis	1.92(1.47-2.51)*	1.59(1.31-1.93)*	1.82(1.29-2.56)*	1.72(0.98-2.99)
Back pain	1.36(1.29-1.43)*	1.41(1.36-1.46)*	1.30(1.21-1.39)*	1.38(1.24-1.53)*
Gout	1.35(1.21-1.51)*	1.42(1.32-1.53)*	1.59(1.34-1.89)*	1.71(1.37-2.13)*
Osteoporosis	1.28(1.17-1.40)*	1.37(1.28-1.46)*	1.45(1.27-1.66)*	1.22(0.98-1.52)
Polymyalgia	1.42(1.18-1.69)*	1.38(1.20-1.58)*	1.67(1.27-2.20)*	1.43(0.90-2.27)
Rheumatoid arthritis	3.20(2.40-4.27)*	2.64(2.20-3.17)*	2.27(1.76-2.91)*	2.22(1.28-3.87)*
Sjogren's Disease	0.95(0.49-1.83)	1.61(0.99-2.58)	1.72(0.86-3.45)	1.72(0.34-8.63)
Systemic lupus erythematosus	1.38(0.58-3.31)	1.60(0.78-3.27)	1.39(0.38-5.04)	-
Fibromyalgia	2.32(1.69-3.19)*	2.32(1.88-2.86)*	1.68(1.24-2.28)*	1.68(0.93-3.05)
Fatigue	1.42(1.18-1.72)*	1.32(1.15-1.50)*	1.17(0.92-1.50)	1.10(0.74-1.64)
Respiratory				
Asthma	1.05(0.91-1.20)	1.16(1.07-1.28)*	1.25(1.05-1.49)	1.30(0.99-1.71)
COPD	1.24(1.12-1.38)*	1.15(1.07-1.24)*	1.13(0.95-1.35)	0.99(0.77-1.25)
Genito-Urinary				
Chronic Kidney Disease	1.14(1.08-1.20)*	1.12(1.07-1.17)*	1.25(1.13-1.38)*	1.23(1.07-1.41)*
Benign prostatic hypertrophy^	1.27(1.14-1.42)*	1.42(1.32-1.53)*	1.22(1.01-1.47)	1.30(1.04-1.62)
Renal stone	1.29(1.01-1.65)	1.30(1.10-1.54)	0.99(0.69-1.41)	1.29(0.73-2.31)
Neuro/Psychiatric				
Stroke	1.21(1.13-1.31)*	1.24(1.18-1.31)*	1.15(1.02-1.30)	1.23(1.04-1.45)
Dementia	1.66(1.51-1.84)*	1.72(1.60-1.85)*	1.89(1.57-2.28)*	1.95(1.49-2.55)*
Epilepsy	1.58(1.17-2.12)	1.41(1.13-1.74)	1.34(0.81-2.19)	1.07(0.57-2.01)
Multiple sclerosis	2.18(1.08-4.36)	1.05(0.61-1.80)	0.82(0.25-2.74)	1.33(0.38-4.69)
Parkinson's Disease	1.68(1.34-2.12)*	1.69(1.43-1.99)*	1.25(0.81-1.94)	1.83(1.04-3.20)
Migraine	1.06(0.89-1.25)	1.23(1.09-1.37)*	1.27(2.05-2.54)	1.25(0.93-1.69)
Depression	1.43(1.33-1.54)*	1.44(1.36-1.51)*	1.36(1.22-1.51)*	1.57(1.34-1.85)*
Psychosis	0.94(0.57-1.55)	0.99(0.68-1.43)	1.23(0.53-2.83)	0.78(0.25-2.44)
Schizophrenia	1.26(0.87-1.84)	0.96(0.74-1.24)	0.77(0.42-1.42)	0.91(0.42-1.97)
Cancer	1.60(1.49-1.72)*	1.59(1.51-1.67)*	1.46(1.30-1.63)*	1.65(1.40-1.94)*
Circulatory				
Coronary Heart Disease	1.29(1.17-1.41)*	1.30(1.22-1.39)*	1.32(1.14-1.53)*	1.09(0.89-1.34)
Arterial/Venous	1.71(1.42-2.07)*	1.54(1.33-1.77)*	0.93(0.64-1.35)	1.64(1.01-2.67)
Heart failure	1.64(1.45-1.86)*	1.82(1.66-2.00)*	1.58(1.24-1.99)*	1.36(0.97-1.90)
Hypertension	1.05(0.99-1.11)	1.04(1.01-1.08)	1.08(0.99-1.17)	1.01(0.91-1.13)
Peripheral vascular disease	1.52(1.34-1.73)*	1.41(1.29-1.55)*	1.46(1.19-1.79)*	1.42(1.05-1.93)
Metabolic/Endocrine				
High Cholesterol	0.97(0.91-1.04)	1.08(1.03-1.12)*	1.09(0.99-1.19)	1.16(1.01-1.33)
Diabetes Mellitus	1.07(1.00-1.15)	1.19(1.14-1.25)*	1.24(1.11-1.38)*	1.12(0.97-1.30)
Hyperthyroid	1.02(0.79-1.34)	1.04(0.86-1.27)	1.52(1.04-2.22)	1.07(0.62-1.86)
Hypothyroidism	1.02(0.92-1.14)	0.96(0.89-1.04)	1.16(0.99-1.34)	1.14(0.91-1.42)
Digestive				
Gastritis	1.57(1.41-1.75)*	1.51(1.40-1.63)*	1.31(1.12-1.53)*	1.39(1.11-1.74)*
Gastrointestinal bleed	1.62(1.34-1.96)*	1.97(1.71-2.26)*	1.28(0.94-1.74)	1.52(1.00-2.30)
Gall bladder stone	1.33(1.19-1.50)*	1.31(1.20-1.42)*	1.45(1.23-1.70)*	1.13(0.88-1.46)
Inflammatory bowel disease	1.41(1.25-1.59)*	1.41(1.29-1.53)*	1.33(1.12-1.58)*	1.62(1.26-2.08)*
Liver Disease	1.48(1.09-2.02)	1.64(1.33-2.00)*	1.38(0.85-2.21)	1.49(0.82-2.72)
Others				
Hearing	1.17(1.10-1.25)*	1.19(1.15-1.25)*	1.23(1.11-1.35)*	1.37(1.19-1.57)*
Psoriasis	1.09(0.89-1.33)	1.05(0.91-1.20)	1.12(0.85-1.47)	0.97(0.64-1.48)
Scleroderma	1.23(0.47-3.24)	1.31(0.54-3.22)	0.96(0.24-3.82)	-
Sleep Disorder	1.35(1.19-1.54)*	1.39(1.27-1.52)*	1.66(1.35-2.03)*	1.39(1.05-1.86)
Tuberculosis	1.58(0.68-3.66)	1.36(0.85-2.19)	2.55(0.99-6.54)	0.87(0.24-3.12)
Anaemia	1.74(1.59-1.92)*	1.61(1.51-1.72)*	1.33(1.14-1.55)*	1.55(1.25-1.92)*
Vision problem	1.11(0.87-1.40)	1.09(0.93-1.29)	1.39(0.93-2.09)	1.37(0.76-2.48)

Adjusted for age, sex, BMI, alcohol use, smoking status, and index date; p-y person years; **P*-value <0.05 'False discovery rate' (FDR) adjusted; COPD- Chronic Obstructive Pulmonary Disease

The risks of developing comorbidities following a diagnosis of joint specific OA are given in Table 4.3-7. The risks of additional multimorbidity were higher for people with wrist/hand OA (aHR 1.46; 95% CI 1.36-1.56), knee OA (aHR 1.24; 95% CI 1.20-1.28), ankle/foot OA (aHR 1.17; 95% CI 1.07-1.29) and hip OA (aHR 1.16; 95% CI 1.11-1.21).

People with hip OA had higher risk of being subsequently diagnosed with rheumatoid arthritis (aHR 3.20; 95% CI 2.40-4.27), fibromyalgia (aHR 2.32; 95% CI 1.69-3.19), ankylosing spondylitis (aHR 1.92; 95% CI 1.47-2.51), anaemia (aHR 1.74; 95% CI 1.59-1.92), arterial/venous diseases (aHR 1.71; 95% CI 1.42-2.07), Parkinson's disease (aHR 1.68; 95% CI 1.34-2.12) and dementia (aHR 1.66; 95% CI 1.51-1.84).

Whereas, among people with knee OA the leading comorbidities diagnosed prospectively were rheumatoid arthritis (aHR 2.64; 95% CI 2.20-3.17), fibromyalgia (aHR 2.32; 95% CI 1.88-2.86), gastro-intestinal bleeding (aHR 1.97; 95% CI 1.71-2.26), heart failure (aHR 1.82; 95% CI 1.66-2.00), dementia (aHR 1.72; 95% CI 1.60-1.85) and Parkinson's disease (aHR 1.69; 95% CI 1.43-1.99).

After incident wrist and hand OA, the risks of being diagnosed with comorbidities were, rheumatoid arthritis (aHR 2.27; 95% CI 1.76-2.91) , dementia (aHR 1.89; 95% CI 1.57-2.28), ankylosing spondylitis (aHR 1.82; 95% CI 1.29-2.56), fibromyalgia (aHR 1.68; 95% CI 1.24-2.28), polymyalgia rheumatica (aHR 1.67; 95% CI 1.27-2.20), sleep disorders (aHR 1.66; 95% CI 1.35-2.03), gout (aHR 1.59; 95% CI 1.34-1.89) and heart failure (aHR 1.58; 95% CI 1.24-1.99).

In people with ankle/foot OA, prospectively there was an increased risk of diagnosis of rheumatoid arthritis (aHR 2.22; 95% CI 1.28-3.82), dementia (aHR 1.95; 95% CI 1.49-2.55), gout (aHR 1.71; 95% CI 1.37-2.13), cancer (aHR 1.65; 95% CI 1.40-1.94), inflammatory bowel disease (aHR 1.62; 95% CI 1.26-2.08) and depression (aHR 1.57; 95% CI 1.34-1.85). (Table 4.3-7)

4.3.3 Sensitivity analysis

From the eligible individuals, cases and matched controls were selected having no comorbidities diagnosed prior to or on the index date for the prospective analysis. Of 221,807 incident OA cases 22,333 (10.1%) were without any of the comorbidities of interest on the index date. An equal number of controls without comorbidities was selected matched by age (±2 years), sex and practice area. The mean age was 56.7 years (SD- 13.6) in OA cases and 56.5 years (SD- 13.6) in matched non-OA controls, 52.4% in both groups being women. The median length of follow up after the index date was 8.05 years (IQR 4.15-19.96 years) and the mean length of follow up was 11.54 years (SD 5.37 years). Details of the distribution of covariates are shown in Table 4.3-8.

Table 4.3-8. Characteristics of incident OA patients and controls at the risk for comorbidities (without any comorbidities at the baseline)

	Incident OA (n=22,333)	Controls (n=22,333)	Unadjusted Odds Ratio (95%CI)	Adjusted Odds Ratio[#] (95% CI)
Age (years)				
<40 years	2484(12.09)	2577(12.57)	NA	NA
40-49 years	4017(17.99)	4101(18.36)	NA	NA
50-59 years	6594(29.53)	6496(29.09)	NA	NA
60-69 years	5477(24.52)	5490(24.58)	NA	NA
70-79 years	2949(13.20)	2885(12.92)	NA	NA
80-89 years	776(3.47)	746(3.34)	NA	NA
≥90 years	36(0.16)	38(0.17)	NA	NA
Gender				
Men	10622(47.56)	10622(17.56)	NA	NA
Women	11711(52.43)	11711(52.43)	NA	NA
BMI (kg/m^2)				
<18.5	279(1.25)	452(2.02)	0.85(0.74-0.97)*	0.82(0.71-0.96)*
18.5- 24.9	6214(27.85)	8493(38.04)	Reference	Reference
25.0-29.9	8314(37.26)	8367(37.48)	1.40(1.34-1.46)*	1.38(1.32-1.44)*
≥30	7503(33.63)	5010(22.44)	2.09(2.01-2.19)*	2.09(1.99-2.20)*
Missing	16(0.05)	28(0.10)	NA	NA
Alcohol use (units/week)				
Never	4318(19.33)	4139(18.53)	Reference	Reference
Ex-drinker	536(2.40)	465(2.08)	1.04(0.92-1.17)	1.10(0.96-1.26)
Current 1-9	8052(36.05)	8245(36.92)	0.94(0.89-0.99)*	0.93(0.88-0.99)*
Current >=10	4147(18.57)	4237(18.97)	0.93(0.88-0.98)*	0.93(0.87-0.98)*
Current Unknown	5277(23.63)	5246(23.49)	0.94(0.89-0.99)*	0.96(0.91-1.02)
Missing	3(0.01)	3(0.01)	NA	NA
Smoking Status				
Never smoked	11715(52.45)	12160(54.44)	Reference	Reference
Ex-smoker	6101(27.31)	5774(25.85)	1.12(1.07-1.16)*	1.10(1.05-1.12)*
Current smoker	4516(20.22)	4399(19.69)	1.06(1.02-1.11)*	1.06(1.02-1.11)*
Missing	0	1(0.003)	NA	NA
Age in years (Mean, SD)	56.71(13.55)	56.53(13.58)		
BMI in Kg/M^2(Mean, SD)	28.44(5.68)	26.80(5.05)		

#Adjusted by age, index year and first year of registration; *P value <0.05; NA-not applicable; BMI- Body mass index; SD- Standard deviation
Mean age (Men 56.67 years, sd-13.42 years; Women- 58.97, SD 13.76 years)

The covariates adjusted cumulative probability of having multimorbidity was higher in incident OA cases compared to controls at all time-points after the index date (log-rank test, p<0.001). (Appendix Table 8, page 313) The cumulative probabilities of having multimorbidity at 5, 15 and 20 years following the index date were 0.64%, 22.14% and 52.93% in people with incident OA and 0.25%, 15.53% and 38.00% in controls, respectively. (Appendix table 8 (page 313) and Appendix Figure 8 (page 315))

The risk of having multimorbidity was 34% higher (aHR 1.34; 95% CI 1.28-1.41) in OA cases compared with controls after adjusting for other covariates such as age, BMI, smoking status, alcohol consumption and index date. The risk of developing incident comorbidity in musculoskeletal, neurological, cardio-vascular, cancer and digestive systems was higher in patients with OA (Appendix Table 5, page 310). For example, patients with OA were five times more likely to develop fibromyalgia (aHR 5.29; 95% CI 2.65-10.50), more than four times more likely to develop rheumatoid arthritis (aHR 4.31; 95% CI 2.68-6.92) and three times more likely to develop liver diseases (aHR 3.36; 95% CI 1.89-5.97) than matched controls. For dementia and ankylosing spondylitis, the risks were nearly two times higher in patients with OA compared to matched controls. Patients with OA were 1.5 to 2 times more likely to develop osteoporosis, benign prostatic hypertrophy, depression, peripheral vascular diseases, heart failure, gastrointestinal bleeding, sleep disorder, and anaemia compared to matched controls. Also, the risks of developing gastritis (aHR 1.41; 95% CI 1.15-1.74) and diabetes (aHR 1.26; 95% CI 1.11-1.43) were significantly higher in patients with OA compared to the matched controls. (Appendix Table 9, page 314)

Appendix Figure 9 (page 316) depicts the comparison of adjusted ORs and HRs for comorbidities in OA. It shows that people who had musculoskeletal or other pain-related conditions before diagnosis of OA/index date are more likely to develop OA, and OA people are also more likely to develop other pain-related conditions after the diagnosis. Appendix Figure 10 (page 317) compares the hazard ratio from the two samples.

4.4 Discussion

This study estimated the burden of comorbidities prior to the diagnosis of OA and the risk of developing comorbidities following the diagnosis of OA using a nationally representative large UK primary care database. The key findings are: (1) people diagnosed with OA were significantly more likely to have multimorbidity both prior and

following the diagnosis of OA; (2) while musculoskeletal (MSK), gastrointestinal (GI), cardiovascular (CV) and psychological conditions (MH) were associated with OA in both temporal directions, dementia and systemic lupus erythematous (SLE) were only associated with OA after its diagnosis; and (3) additionally, there was a bidirectional association both before and after the diagnosis of OA with anaemia, inflammatory bowel disease (IBD), benign prostatic hypertrophy (BPH), gall bladder stones, liver diseases, cancer and hearing impairment.

4.4.1 Associations in both retrospective and prospective analyses

Multimorbidity associations with OA before and after the diagnosis reveal the important role of MSK conditions. Both multimorbidity and OA have positive relationships with ageing, which was accounted for in the analysis. Multiple shared risk factors such as obesity, physical inactivity, medication use and the possible role of inflammation in multimorbidity might lead to OA and vice-versa (Friedman E.M., Christ S.L., and Mroczek D.K., 2015; Chudasama *et al.*, 2019). Especially in this work, the association with development of new multimorbidity after adjusting for comorbidity burden at the baseline was estimated. The adjusted HR of 1.29 indicates the higher burden of multimorbidity among people with OA after the diagnosis.

Associations of OA with some of the identified MSK comorbidities in this study accord with previous studies, though the causes remain speculative (Reeuwijk *et al.*, 2010). For example, systemic inflammatory disease, such as rheumatoid arthritis might damage joints and lead to “secondary” OA, and a lesser inflammatory component is increasingly recognised in OA pathogenesis (Berenbaum, 2013). Association of OA with some of the musculoskeletal comorbidities are well known. For example, the association between OA and rheumatoid arthritis (RA) is very consistently reported (Reeuwijk et al., 2010; Ruiz-Medrano, Espinosa-Ortega and Arce-Salinas, 2019). People presenting with RA are more at risk of developing OA in the future and the opposite is also reported. Although the latter

has not been studied in detail, Lu et al reported that the risk of RA diagnosis among people with OA was five times higher compared to controls (Lu et al., 2015), whereas in this study the risk was three fold. The exact reasons for the association are not well studied, but it can be hypothesized that OA triggers multiple factors for development of RA. For example, OA increases inflammatory chemicals such as pro-inflammatory cytokines, including interleukins, into the circulation which is also found in RA patients (Berenbaum, 2013). Also possession of HLA-DRB1 and citrulline proteins found in patients with OA in the presence of appropriate risk factors such as genetics and environmental factors, including diet, may influence the autoimmune system to develop RA (Wojdasiewicz, Poniatowski and Szukiewicz, 2014). Other common risk factors for both OA and RA could be obesity. Studies have found that, obesity is linked with OA and adipokines - a type of cytokine secreted by adipose tissue which may increase the risk of RA (Gómez et al., 2011). In the retrospective analysis, the association of the multiple musculoskeletal conditions with OA was higher as the time-period before the index date shortens. This indicates there was possibilities of misdiagnosis or problem in differential diagnosis to exclude the probable diseases. Care must be taken to interpret such findings and more research is needed to understand the accuracy of the reporting of the MSK conditions in primary care.

Similarly, the bidirectional associations with discrete chronic pain-related conditions such as fibromyalgia, back pain and IBS could result from shared non-restorative sleep and central pain sensitization, which result in reduced pain threshold and exacerbation of other causes of pain (Whitehead *et al.*, 2007; Kirkness, Yu and Asche, 2008). This relationship is well researched (Kirkness, Yu and Asche, 2008; Hoogeboom, den Broeder, *et al.*, 2012; Siemons *et al.*, 2013; Zambon, Siviero, Denkinger, Limongi, Castell, van der Pas, Otero, Edwards, Peter, Pedersen, Sánchez-Martinez, Dennison, Gesmundo, Schaap, Deeg, van Schoor, Maggi and EPOSA Research Group, 2015). According to Kadam et al, the association of OA with pain-related conditions in the UK population are

nearly two times higher than the control group (Kadam, Jordan and Croft, 2004). People with OA are also reported to exhibit widespread hyperalgesia to mechanical pressure and cold (Moss, Knight and Wright, 2016). Hyperalgesia - an increased sensitivity to painful stimuli - is very common in widespread pain syndromes such as fibromyalgia. The commonality of hyperalgesia in OA and widespread pain syndromes suggests the possible following pathophysiology of developing painful conditions like back pain, fibromyalgia and ankylosing spondylitis (Staud, 2011). It is possible that OA structural joint changes produce chronic stimuli to the nociceptors which decreases the pain threshold (peripheral sensitisation) causing hypersensitivity to pain. Furthermore, these lead to central nervous system (CNS) plasticity and central sensitisation through the stimulation of C-fibres and impairment of descending inhibitory systems (Coderre et al., 1993; Melzack et al., 2001). Other possible reasons for the increased diagnosis of pain-related conditions could be because of the release of NMDA (N-methyle-D-Aspartate) chemicals in OA which activates the COX-2 gene, substance P and nerve growth factor (NGF) which in turn increases the pain sensitivity. Mechanical factors introduced by obesity and OA pain (change in gait, joint deformities) may also contribute to the increased risk of back pain (Wolfe et al., 1996; W. Wang et al., 2016). OA related changes also can be seen in cartilages of spinal facet (apophyseal) joints, which might predispose to the back pain (Ashraf et al., 2014).

The association of OA with gout was stronger before the diagnosis of OA than after, and this bidirectional relationship might in part be explained by the “amplification loop” of cartilage damage enhancing urate crystal deposition and urate crystals causing cartilage damage (Ma and Leung, 2017). Even though epidemiological studies mention the increasing burden of co-existence, a recent systematic review has identified ‘obesity’ as the shared mediator for both hyperuricaemia and OA (Ma and Leung, 2017). Another possible pathological explanation could be that cartilage disruption and exposure of cartilage fragments in OA (i.e. chondrocyte death) leads to local urate generation, which

might deposit on the cartilage and trigger cytokine and protease production (Hwang et al., 2015; Charlier et al., 2016). These chemicals might produce the vicious cycle of OA degeneration and increased joint urate levels. Furthermore, OA cartilage enhances the deposition of urate, as well as calcium, crystals in the joint due to increased promotors and reduced inhibitors of crystal nucleation.

The hazard ratio for osteoporosis following diagnosis of OA to be higher than the odds ratio of OA following diagnosis of osteoporosis, but the evidence and explanations for an association between osteoporosis and OA remain controversial. Most studies have reported high bone density in OA (Dequeker, Aerssens and Luyten, 2003; Im and Kim, 2014), while a few have reported the opposite (Hochberg, Lethbridge-Cejku and Tobin, 2004). This study shows that OA increases the risk of developing osteoporosis. Possible reasons for this could be because of shared epidemiological risk factors or biomechanical factors. Geusens and Bergh have proposed the shared mechanism for OA and osteoporosis to be life-style factors, BMI and osteosclerosis (Geusens and Bergh, 2016). Immobility because of OA pain and obesity could lead to accelerated bone mass loss. Molecular studies have identified 12 specific proteins, of which 8 were closely related to the pathogenesis of osteoporosis and knee OA (Shi and Zhang, 2018). Moreover, the osteoporosis reported in OA varied from joint to joint, being more common in distal joints and spine compared to large weight-bearing joints. Use of ‘osteoporosis’ in any joint in this analysis, which could have overestimated the burden irrespective of the joint involved in OA, which warrants future research.

Care must be taken in interpreting these associations, especially where joint pain is the reason for the consultations since GP diagnoses are predominantly clinical and not pathological. Also, although characteristics of these various MSK conditions differ there is still the possibility of misdiagnosis, especially for atypical cases.

Cardiovascular diseases, such as coronary heart disease and heart failure (Rahman *et al.*, 2013), stroke (Hsu *et al.*, 2017; Swain *et al.*, 2019), PVD (Findlay, 2007) and diabetes

(Louati *et al.*, 2015) are well known to associate with OA. In this study prospective risks of developing diabetes, PVD and heart failure were greater in OA compared to risks of developing OA in people with these conditions. Firstly, OA reduces physical activity levels, which predisposes to hypertension and hyperlipidaemia (Hootman et al., 2003), and physiologically it increases the blood viscosity through endothelial damage (Koenig et al., 1997). The role of inflammatory substances such as CD40 and vascular cell adhesion molecule found in OA (Hoeven et al., 2015) increases the risk of carotid intimal thickening and carotid plaque (Wang et al., 2011) leading to atherosclerosis (Libby, Ridker and Maseri, 2002). Another possible factor could be the use of pain killers such as NSAIDs which also increase the risk of heart disease, hypertension and stroke (Haag, 2008). Insulin resistance and diabetes are part of metabolic syndrome, so the suggested mechanism for the association between diabetes and OA is like the association of metabolic syndrome and OA. There is a bidirectional association of diabetes and OA, i.e. the prevalence of OA in diabetes is increased (Louati et al., 2015) and so is the prevalence of diabetes in OA. However, the latter has been explored in more detail (Schett et al., 2013; Al-Jarallah et al., 2016) and the increased risk of diabetes in an OA population is the least examined. It is possible that OA and diabetes connect through a vicious cycle influencing each other's outcome (King and Rosenthal, 2015). This indicates the role of risk factors such as obesity and hypercholesterolaemia in causing OA, and that screening for metabolic syndrome and CVD may be considered in people presenting with OA (NICE, 2014).

One of the sparsely investigated comorbidities in OA is peripheral vascular diseases (PVD). This interesting association is harder to explain in the absence of any relevant mechanistic studies. Possibly, the vascular pathology such as circulating cytokines in OA also affects the smaller vessels mediating the PVDs (Findlay, 2007). The slowed blood circulation in smaller blood vessels in OA needs to be explored further in detail. However,

Lee et al found a significant association of PVD with use of pain killers in people with OA (Lee et al., 2016).

Even though depression and OA had a significant bidirectional association, a higher risk of depression was seen in people following the diagnosis of OA. A similar finding was seen with sleep disorders. Depression and non-restorative sleep are well recognised to associate with chronic pain experience in OA (Stubbs *et al.*, 2016b). Low affect and non-restorative sleep can reduce descending pain inhibition and cause central sensitisation, and equally chronic pain and reduced participation can cause mood disturbance (Parmelee, Tighe and Dautovich, 2015).

The risks of developing gastritis, GI bleeding, liver diseases and gallstones in OA were high compared to developing OA in these conditions. Gastritis, gastro-intestinal bleeding, liver cirrhosis and peptic ulcer are known comorbidities in OA that may result from NSAID usage (Zak and Pasiyeshvili, 2016). (Papatheodoridis, Sougioultzis and Archimandritis, 2006). (Zak and Pasiyeshvili, 2016; Zak et al., 2019) However, increased recording of incident OA in people with these conditions could result from self-medication for OA pain before presenting to the general practitioner and being diagnosed with OA (i.e., protopathic bias). Interestingly, the risk of OA in liver cirrhosis is reported to be high but the reverse relationship has yet to be established (Arora *et al.*, 2016).

Another interesting finding was that the risk of developing cholelithiasis was higher in the OA compared to the control population, and this association with OA appeared bidirectional. It is hard to explain these associations mechanistically. Possibly, since both of these conditions share common risk factors, sometimes known as the four F's (female, forty years of age, fatty tissue and fertile) (Schirmer, Winters and Edlich, 2005; Njeze, 2013), the comorbidity pattern may reflect merely co-existence due to these risk factors rather than any linked pathogenesis such as genetic, environment and other comorbidities. Studies have shown an association between H. Pylori and gallstones but its association with OA is unclear (Popescu, Andrenscou and Babes, 2018). Also, the

analgesics which are commonly used in OA are not apparently linked with gallstones (Sterling et al., 1995; Pazzi et al., 1998). Further research is needed to explain the biological link between these conditions.

Other comorbidities with significant bi-directional associations with OA were respiratory, hypothyroidism and neurological conditions such as Parkinson's disease, epilepsy, and migraine. Thyroid disease, epilepsy, migraine, and respiratory illness may have earlier age of onset than OA, which could have led to the early recording in the database prior to OA. Also, these comorbidities could be mediated through the systemic inflammation, medication use or other comorbidities in OA.

Evidence suggests a higher prevalence of OA in asthma patients but the high prevalence of asthma in OA still needs to be proved (Mahmood and Malghooth, 2019). This is one of the least studied conditions in patients with OA. Wshah et al in their systematic review found that the prevalence of OA was higher in individuals with COPD (Wshah et al., 2018). The reason for a direct association between COPD and OA is not well understood, but both chronic conditions share common risk factors such as obesity and physical activity.

The four other conditions with bi-directional positive associations in this study were anaemia, BPH, cancer and hearing problems, which have all been reported before (Kramer *et al.*, 2002; Zlateva *et al.*, 2010). Release of inflammatory substances in OA has been linked with sensorineural hearing loss (Takatsu *et al.*, 2005), BPH (Chughtai *et al.*, 2011), cataract(Jonas *et al.*, 2018) and cancer (Ziegler, 1998) and subclinical systemic inflammation may occur for many years before OA becomes symptomatic and clinically apparent. In addition, the rise in incidence of BPH in OA could result from the use of analgesics such as NSAIDs (Nygård et al., 2017). The incidence of these conditions in OA warrants further research. The association between OA and cancer is difficult to explain and has not been studied well. However, both OA and cancer share similar inflammatory mechanisms (Ziegler, 1998). Interestingly, the use of NSAIDs reduces the

spread of metastases in some cancer patients, which is against any potential role of pain medications (Wang et al., 2011; Zhao et al., 2006). Association of OA and hearing loss has not gained the attention of clinicians and researchers, though Kramer et al previously reported a significant association between these two conditions (Kramer et al., 2002). Cartilage damage in the incudomalleolar joint between the malleus and incus bones and the incudostapedial joint between the incus and stapes bones in the ear, which impacts on hearing, has been reported in people with OA (Rawool and Harrington, 2007). Also the low-grade chronic inflammation in OA could lead to sensorineural hearing loss, just as it is reported to do in RA (Takatsu et al., 2005).

4.4.2 Association in prospective analysis only

Dementia was associated with OA only in the prospective analysis. This is similar with a recent systematic review of cross-sectional and case-control studies which reported that people with OA were 20% more likely to have dementia (Weber *et al.*, 2019). Similar findings were reported by a longitudinal study from Taiwan (Huang et al., 2015). As dementia is predominantly an ageing problem, the association in the retrospective study may not have been significant because of the low prevalence of dementia in younger decades and difficulty in detecting OA symptoms and less consultations for OA in people with dementia. One of the possible explanations for the prospective association with dementia could be the role of inflammatory cytokines (IL-1), which result from joint inflammation and can reach the cerebral circulation causing neuro-inflammation (Kyrkanides et al., 2011). The association between OA and SLE is ill searched and difficult to explain, which needs further investigation.

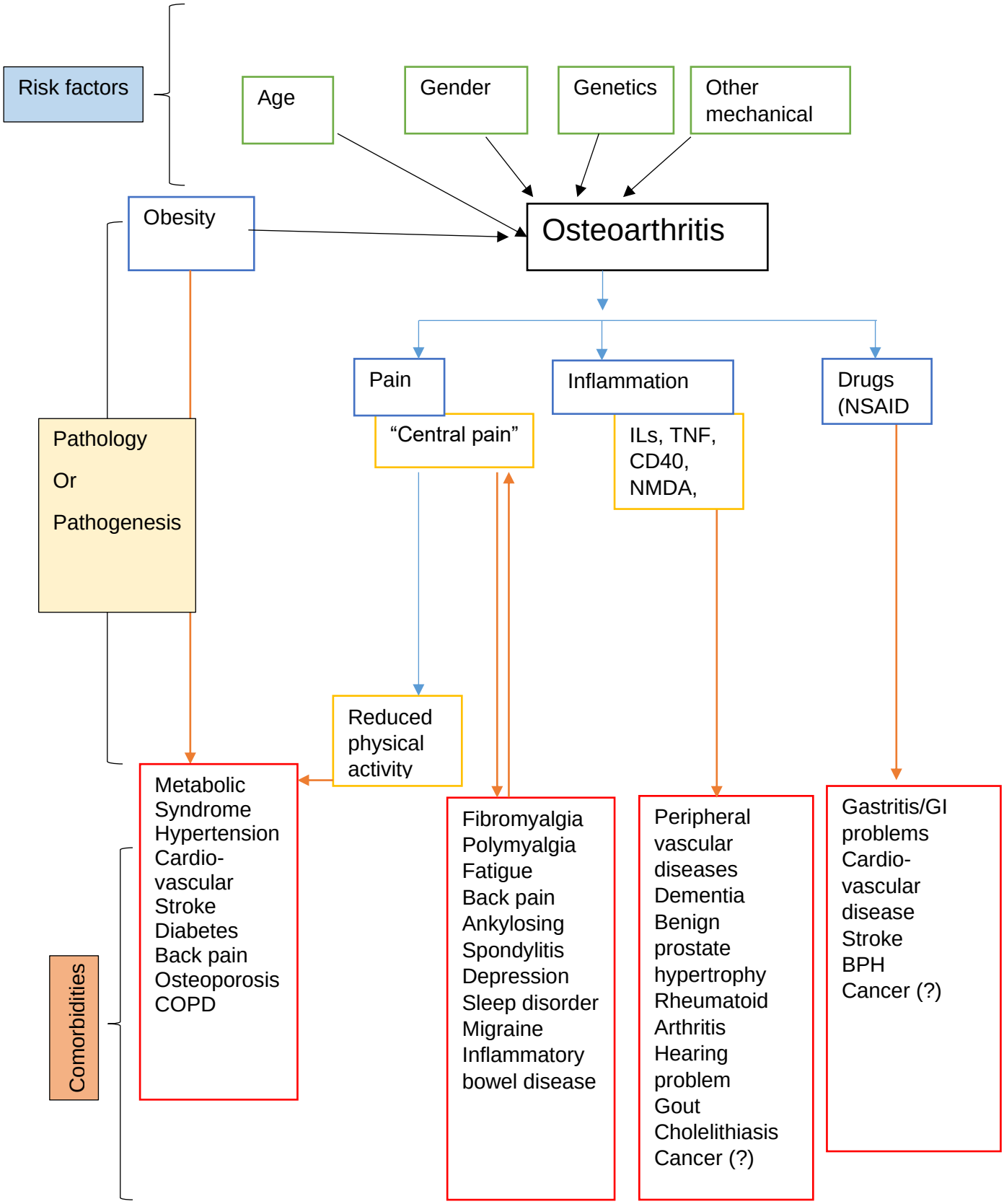
Association of comorbidities with joint specific OA are like the above-mentioned comorbidities. However, some of the additional findings are association of hip OA with chronic kidney disease and migraine, knee OA with COPD and migraine, and hand OA with renal stones and epilepsy, prospectively. These specific associations may be

mediated through the factors already discussed, but further investigation into mechanisms of linkage is still required.

This study suggests that although structural changes of OA may appear relatively limited within the skeleton, pathologically and physiologically, its effect may be seen in almost every organ. Although the burden of comorbidities in RA and gout may be higher than that in OA, the significant associations of multiple chronic conditions with OA found in this study should not be neglected. A Versus Arthritis report on multimorbidity in OA also highlighted the importance of understanding the presence of multiple comorbidities with OA for formulating a 'patient-centred' management plan (Loftis, Ellis and Margham, 2014). Thus, close observation of people with OA through annual assessment in primary care appears warranted, as recommended by NICE (Conaghan *et al.*, 2008). Recently, the European League Against Rheumatism (EULAR) and NICE, have emphasised the importance of diagnosis and management of specific comorbidities and understanding their pattern in OA when managing people with OA (Conaghan *et al.*, 2008).

Figure 4.4.1 depicts the possible pathogenesis of association with comorbidities in people with OA. As discussed above, age, obesity, pain, inflammation, and medication use are thought to be the drivers of incident of new comorbidities and vice versa. The strong association with MSK conditions and pain related diseases are largely explained through central pain mechanism and changes in inflammatory substances. Metabolic pathology related to obesity can cause many chronic conditions including CV and respiratory. Another aspect of OA least explored is the use of medication. Though there is no cure for OA, prescription of drugs such as pain killers and anti-inflammatory reduce the pain symptoms and slows down the disease progression bringing relief to the patient. However, these drugs are not immune to the side-effects. Commonly studied comorbidities associated with the drugs are gastrointestinal and CV. Further research is needed to explore the association with other identified chronic conditions.

Figure 4.4-1 Schematic presentation of OA and possible associated factors with the comorbidities



?Indicates the possible association; IL- interleukins; TNF- Tumour Necrosis factor; NMDA- N-methyl-D-aspartate; CD40- Cluster of differentiation 40; NSAIDs- Nonsteroidal anti-inflammatory drugs

4.4.3 Study limitations

There are several caveats to this study. The possibility of misclassification of OA because of physician diagnosis rather than full clinical and imaging assessment has been emphasised already. Nevertheless, I tried to optimise identification of symptomatic OA cases through strict inclusion and exclusion criteria using similar methodology to that of previous studies (Swain *et al.*, 2020) and there is some reassurance that the codes for hip OA have been shown to have good validity (Ferguson *et al.*, 2019). Misclassification bias for comorbidities is also possible, though most comorbidities in the study have previously been validated (Herrett *et al.*, 2010; Nada F. Khan, Harrison and Rose, 2010). Another important caveat is that risk factors such as diet and physical activity in the analysis could not be included, as these are not routinely recorded within CPRD. Therefore, the estimates from this study may not always relate to direct associations between OA and comorbidities, which could have been mediated through other unrecorded factors. Because of the indolent nature of OA, the recording of the OA in the database is possible to happen long after the initial symptoms and pathological changes in the body. Thus, the temporal association with comorbidity is difficult to assume in this study. The primary aim of the study was to estimate the associations and burden of comorbidities diagnosed prior to and after the diagnosis of OA, rather than to define risk factors. The associations could to some extent be due to ascertainment bias through increased numbers of hospital or GP visits in people with OA, especially for the stronger association with rheumatological conditions. Along with the possible Berkson bias, where patients assessed in hospital undergo more routine testing so may have occult comorbidity diagnosed more often, a chance of collider bias due to sampling design might exist. That means diagnosis of comorbidities might have been influenced by reporting of other chronic conditions, especially by the QOF guidelines where the quality of recording of certain conditions are much better than others. However, the controls were matched having a minimum 36 months of registration and at least one consultation for any reasons. More focus was

given on the possible explanation of the association rather than the causes, which is beyond the scope of this study. The maximum follow-up in this study was for 20 years and the cases and controls were not matched for the length of follow-up which might have influenced the diagnosis of comorbidities.

The sample size for the prospective analysis was nearly 440,000 with equal numbers of OA cases and matched controls and maximum follow-up for up to 20 years, making this the first prospective study to provide such a clear picture of the burden of many comorbidities. The 49 chronic conditions studied using case-control design on the database that represents the general population.

4.4.4 Conclusion

In conclusion, the risk of multimorbidity is higher in people with OA. Musculoskeletal, gastrointestinal, CVD and psychological conditions are associated with OA both before and after diagnosis. Significant associations with gallstones, IBD, BPH, anaemia, hearing problems, liver disease and cancer highlight the discordant comorbidities in OA, which cannot readily be explained mechanistically. Bidirectional associations with multimorbidity and 40 comorbidities suggest the need to identify shared risk factor mechanisms. The temporal associations reported merit further investigation regarding causality and have important clinical implications with respect to optimal management of OA and its potential comorbidities. Future studies should investigate clustering of the comorbidities and shared risk factors.

Summary of Chapter 4

Chapter 4 explored the burden of comorbidities in people with OA compared to the controls (non-OA). Key findings from the chapters are:

- The burden of comorbidities and multimorbidity (additional two chronic conditions) was more in people with OA compared to age and sex matched controls.
- Of 49 chronic conditions examined, nearly 30 had significant association with OA both before and after the diagnosis of OA.
- Musculoskeletal (MSK), gastrointestinal (GI), cardiovascular (CV) and psychological conditions (MH) were associated with OA
- Interestingly, there was a significant association both before and after the diagnosis of OA with anaemia, inflammatory bowel disease (IBD), benign prostatic hypertrophy (BPH), gall bladder stones, liver diseases, cancer, and hearing impairment.

Knowing the burden of the comorbidities in people with OA, the next research question to explore the pattern of co-existence of these chronic conditions in people with OA and their controls. Which means how the people are clustered or grouped based on the co-existence of the diseases. The next chapter (chapter 5) investigates the research question.

5 Chapter 5

Clustering of comorbidities

5.1 Introduction

The prevalence of multimorbidity (≥ 2 long term conditions) is widely being reported worldwide (Nguyen *et al.*, 2019). Multimorbidity affects 30% of adults (≥ 18 years) in the UK (Cassell A. *et al.*, 2018). The association of graded effect of count of multimorbidity and health outcomes has been studied in detail (Brilleman and Salisbury, 2013). However, it is more important to study the types of conditions, rather than just the number that coexist in a single individual because of the following two reasons.

Firstly, it is essential to understand the cluster or grouping of chronic conditions, which can provide clues about concordant or discordant patterns. There are different possibilities to explain the combination of conditions, but certain conditions often co-exist together. These groups of conditions can be concordant, that is, similar in their aetiology and/or treatments, or alternatively discordant, that is, unrelated etiologically and requiring different management approaches (Ricci-Cabello *et al.*, 2015). The nature of clusters of conditions determines the treatment approaches, especially discordant clusters which increase the complexities of treatment and management (Guthrie *et al.*, 2012). Secondly, understanding these patterns would also help identification of 'at-risk' populations for other conditions and the design of appropriate prevention, screening, and management strategies.

Limited knowledge about the clustering of musculoskeletal conditions and its association with sociodemographic and lifestyle risk factors have led to more 'disease specific' management approaches rather than 'person-centred' individualised care (Duffield *et al.*, 2017). This may lead to the provision of contradictory, expensive, resource-draining and disjointed multiple care. Identification of the most commonly occurring clusters of

diseases can reduce disease burdens and health care costs, inform resource planning, and ultimately improve the quality of life of patients.

Previous systematic review revealed the association of OA with multimorbidity, but also with specific conditions such as CVD and depression (Swain *et al.*, 2019). However, there is scant evidence on OA in terms of its multimorbidity pattern and which cluster it belongs to. The common clusters of conditions in people with OA are not known also, whether these differ from the non-OA population. Therefore, the large CPRD GOLD primary care database in the UK was used to explore the common clusters of conditions and the associated factors in the whole population, to compare findings between those with and those without OA.

5.2 Methods

CPRD GOLD database was used where anonymised primary care clinical data are contributed by UK general practices (Herrett, Arlene M Gallagher, *et al.*, 2015).

5.2.1 Source population

Data on a random selection of individuals were acquired from the CPRD. For this study two different population were chosen for: (1) studying the pattern of conditions in the complete CPRD population; and (2) exploring the pattern among the OA and non-OA sub-populations.

For the complete CPRD population patients were aged 20 years and above with valid registered-status in a practice with data classified by CPRD as UTS in January 2017 and a minimum 1 year of registration period with the database.

For the comparison study, the OA cases and non-OA controls were selected using the same methods as described in chapter 4, section 4.2.2 and 4.2.3.

5.2.2 Comorbidity definition and extraction

Comorbidity was defined as the diagnosis of any of the 49 predefined chronic conditions (in the OA group, any additional diseases other than OA) in individuals of both groups. Details of the comorbidity extraction has been given in Chapter 4, section 4.2.5.

5.2.3 Covariates

Information on BMI, alcohol use and smoking status were used with age and gender. Details of the covariates are given in chapter 4 section 4.2.6. In a subgroup analysis the index of multiple deprivation was examined as a risk factor.

5.2.4 Statistical analysis

For this objective, a clustering analysis method was used.

The clustering of the chronic diseases groups patients with similar morbidities. Various methods are available for doing such analyses (Saxena *et al.*, 2017), however, few have methodologies to group the patients rather than variables. One such method used in this study is latent class analysis (LCA) (Nylund *et al.*, 2007, Lanza and Rhoades, 2013). I preferred LCA over other methods because it is data-driven and identifies distinct patterns (Muthen and Muthen, 2000). In comparison to other commonly used clustering methods, such as hierarchical clustering (Bridges, 1966), LCA uses a probabilistic approach which is not sensitive to rotation of factors and does not require any subjective distance measures for pattern determination (Bartholomew, 2008; Collins and Lanza, 2009). Also, it handles categorical variables in better ways and provides greater reproducibility and stability of the latent class solutions (Feuillet *et al.*, 2015).

5.2.5 Selection of statistical methods

Alternative approaches for clustering to select the best possible methods were tested. One of these was K-mode analysis. K-mode is a machine learning technique, which identifies patient groups with distinct profiles (Huang & Ng, 2003). It is an extension of the widely used K-mean method used for continuous outcomes with centroid based algorithms that calculate the distance between the groups using Euclidian distance or dissimilarity matrix (Wang Shunye, 2013). As variables used in this study are

dichotomous (‘yes’ or ‘no’), the distance measures used by K-means were not useful. Huang et al modified K-means method to K-mode for clustering of patients with categorical variables. K-mode determines the clusters based on the number of matching categories between data points, rather than the similarity index used in K-means. The optimal clusters were assessed using the Silhouette Coefficient (SC). The average SC is known as the Silhouette Index (SI) and evaluates the overall quality of separation between the clusters (Rousseeuw, 1987). The SC is calculated using its intra-cluster distance and its nearest cluster distance. The SC ranges from -1 to +1 explaining the least to the best classification. “diceR” package in R was used to choose the best model for clustering (Chiu and Talhouk, 2018). The selection of models from the three major indices are given in Appendix Method 2 (page 353). Appendix Table 10 (page 320), and Appendix Figure 11 and 12 (page 318-319), shows the clustering using machine learning approach (K-mode).

5.2.6 Latent class modelling

LCA was used to explore multimorbidity patterns from one to ten multimorbidity classes. It is a statistical technique for analysis of multivariate categorical data. The latent class model stratifies the data by observed (“manifest”) by unobserved variables (“latent”). The assumption is that the manifest variables are independent but conditional upon values of latent variables, commonly known as “local/ conditional dependence”. Latent class model probabilistically groups each observation into a “latent class," which in turn produces expectations about how that observation will respond on each manifest variable. The grouping is done by weighted sum of manifest variables calculated by the product of the frequency in a cell and the proportion of observation in cell and the probabilities of being in the cell conditional upon the latent variable. Observation with similar set of responses on manifest variables tend to cluster in same latent classes.

One of the important tasks in LCA is to identify the best fit model and the number of latent classes. Established methods were followed to select the best model based on a combination of statistics, specifically, Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), sample size adjusted BIC (aBIC), log-likelihood ratio test,

entropy and clinical judgement (Nylund, Asparouhov and Muthén, 2007). In theory, models with the lowest BIC are thought to be the best. The model selection is alternatively explained by calculating percentage change in log-likelihood ratio of K-class (LL2) with K-1 class (LL1) using the formula (LL2-LL1)*100/(LL1). Additionally, the best model should have entropy more than 0.70 and should make more sense clinically. The latent classes were named after the posterior probabilities (PP) distribution of conditions to each cluster, ideally should be >= 60%. The group with lowest probabilities of all the conditions were named as the 'relative healthy' group. Once the best class was identified, the groups were attached to the original database and descriptive analysis was done for the covariates. Multinomial regression models to explore the risk factors using 'relative healthy' as the reference group to predict association with other latent classes by including the covariates in the model. The analysis was done in R software using "*po*LCA" package(Linzer and Lewis, 2011).

5.2.7 Sensitivity analysis

The whole dataset was divided into training and testing data, constituting randomly selected 80% and 20% of the study population. LCA was performed in both datasets separately. The agreement was measured using Janssen-Shannon index for similarity and the cluster types (Appendix Table 11, page 321). Appendix tables 12 and 13 (pages 322-323) describe the sensitivity findings. The descriptive statistics of clusters from both training and testing data are given in Appendix tables 14 and 15 (pages 324-325).

5.3 Results

5.3.1 Total study population

5.3.1.1 Descriptive statistics

In total 1,425,823 patients had active registration with the CPRD database as of 1st Jan 2017. Of these, nearly 50% were men and 29% were 60 years or older. The mean age of the total population was 54.49 years (SD 16.92). Almost 21% were obese and the mean BMI in the population was 26.91 Kg/m^2 (SD 5.48). The prevalence of multimorbidity was 48% and 34% had three or more chronic conditions. The mean number of morbidities was

2.19 and the range was 0-22. The other sociodemographic details are given in Table 5.3-1.

Table 5.3-1. Characteristics of people with active registration status in CPRD (from 1st Jan-31st Dec 2017)

	Variables	N= 1,425,823 n (%)		Multimorbidity (N= 682,474) n (%)	
Gender	Men	699553	(49.06)	299470	(43.88)
	Women	726270	(50.94)	383004	(56.12)
Age (as of 1st Jan 2017) (years)	20-35	194497	(13.64)	23182	(3.40)
	36-50	387069	(27.15)	127831	(18.73)
	51-65	427493	(29.98)	227428	(33.32)
	66-80	311403	(21.84)	220597	(32.32)
	>80	105361	(7.39)	83436	(12.23)
Smoking	Never smoked	770455	(54.04)	144152	(21.12)
	Current smoker	299849	(21.03)	365278	(53.52)
	Ex-smoker	280807	(19.69)	170330	(24.96)
	Missing	74712	(5.24)	2714	(0.40)
Alcohol use	Never	223076	(15.65)	120610	(17.67)
	Ex-drinker	26119	(1.83)	17657	(2.59)
	Current (1-9)	418150	(29.33)	217314	(31.84)
	Current (>=10)	220244	(15.45)	120217	(17.61)
	Current (Unknown)	314538	(22.06)	167784	(24.58)
	Missing	223696	(15.69)	38892	(5.70)
BMI	Underweight	32193	(2.26)	11197	(1.64)
	Normal	459535	(32.23)	214726	(31.46)
	Overweight	418842	(29.38)	233200	(34.17)
	Obese	295485	(20.72)	183254	(26.85)
	Missing	219768	(15.41)	40097	(5.88)
Age (Mean, SD) in years		54.49	(16.92)		
BMI (Mean, SD)		26.91	(5.48)		
Number of morbidities (Mean, SD)		2.19	(2.50)		

BMI-Body Mass Index; SD- Standard deviation

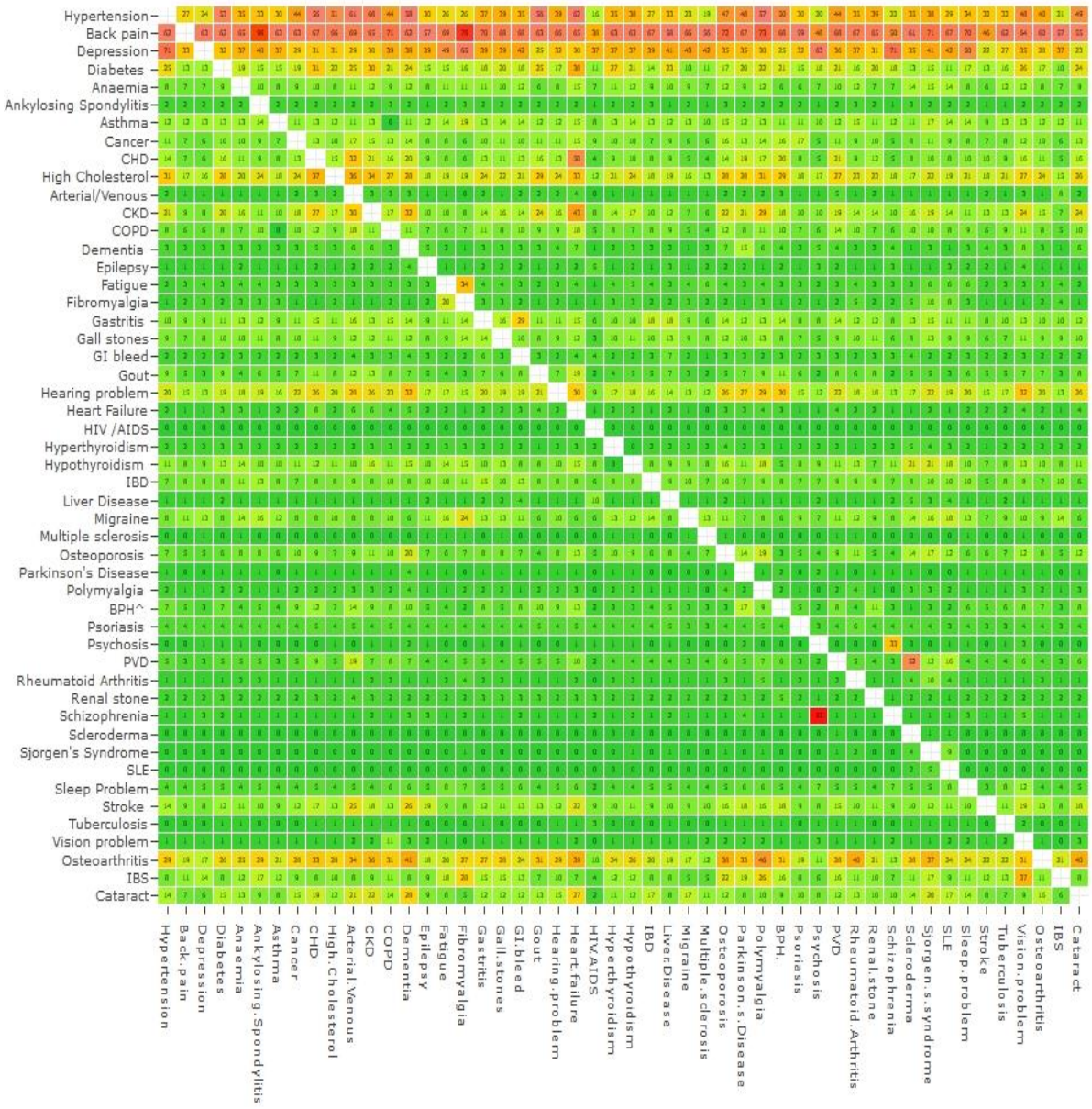
In the study population the prevalence of each chronic condition ever diagnosed is provided in Table 5.3-2. Overall, leading conditions were back pain (38.20%), depression (19.48%), hypertension (18.16%), OA (13.14%) and high cholesterol (10.82%). A similar pattern was seen in women whereas in men the prevalence of hypertension (17.85%) was higher than depression (14.07%). The prevalence of multimorbidity was higher in women (53%) compared to men (43%).

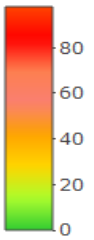
Table 5.3-2. Distribution of chronic conditions in men and women

Conditions	Men (N=699553)	(%)	Women (N=726270)	(%)	Total (N= 1425823)	(%)
Back pain	246185	35.19	298519	41.10	544704	38.20
Depression	98423	14.07	179394	24.70	277817	19.48
Hypertension	124883	17.85	134055	18.46	258938	18.16
Osteoarthritis	73339	10.48	114084	15.71	187423	13.14
High cholesterol	75664	10.82	78581	10.82	154245	10.82
Hearing	64188	9.18	63386	8.73	127574	8.95
Diabetes mellitus	61744	8.83	56060	7.72	117804	8.26
Stroke	53246	7.61	57369	7.90	110615	7.76
Asthma	41135	5.88	60506	8.33	101641	7.13
Irritable bowel syndrome	27767	3.97	73573	10.13	101340	7.11
Migraine	20700	2.96	64048	8.82	84748	5.94
Chronic kidney disease	32873	4.70	46949	6.46	79822	5.60
Benign prostatic hypertrophy	39143	5.60	0	0	39143	5.60
Cataract	30552	4.37	44042	6.06	74594	5.23
Hypothyroidism	12958	1.85	57305	7.89	70263	4.93
Gastritis	33821	4.83	34402	4.74	68223	4.78
Cancer (any)	28226	4.03	36980	5.09	65206	4.57
Chronic heart disease	40490	5.79	23603	3.25	64093	4.50
Gall stones	15381	2.20	44279	6.10	59660	4.18
Inflammatory bowel disease	25582	3.66	34011	4.68	59593	4.18
Anaemia	10713	1.53	45035	6.20	55748	3.91
COPD	23487	3.36	26679	3.67	50166	3.52
Gout	33164	4.74	9654	1.33	42818	3.00
Osteoporosis	5269	0.75	35215	4.85	40484	2.84
Psoriasis	17686	2.53	17645	2.43	35331	2.48
Sleep Disorder	13614	1.95	15963	2.20	29577	2.07
Peripheral vascular disease	12511	1.79	14378	1.98	26889	1.89
Fatigue	6629	0.95	14787	2.04	21416	1.50
Renal stone	10499	1.50	5211	0.72	15710	1.10
Dementia	5766	0.82	9736	1.34	15502	1.09
Hyperthyroid	2804	0.40	12095	1.67	14899	1.04
Gastrointestinal bleed	8087	1.16	5877	0.81	13964	0.98
Fibromyalgia	2016	0.29	10640	1.47	12656	0.89
Ankylosing spondylitis	4167	0.60	8204	1.13	12371	0.87
Epilepsy	5591	0.80	5359	0.74	10950	0.77
Schizophrenia	5060	0.72	5260	0.72	10320	0.72
Heart failure	6077	0.87	4293	0.59	10370	0.73
Rheumatoid arthritis	2974	0.43	6613	0.91	9587	0.67
Polymyalgia	2749	0.39	6050	0.83	8799	0.62
Liver disease	4042	0.58	3631	0.50	7673	0.54
Arterial/venous	4814	0.69	2259	0.31	7073	0.50
Vision	2296	0.33	2870	0.40	5166	0.36
Multiple sclerosis	1384	0.20	3068	0.42	4452	0.31
Psychosis	2190	0.31	1869	0.26	4059	0.28
Parkinson’s disease	2369	0.34	1616	0.22	3985	0.28
Tuberculosis	1683	0.24	2209	0.30	3892	0.27
Sjogren’s syndrome	224	0.03	1488	0.20	1712	0.12
Systemic lupus erythematosus	106	0.02	615	0.08	721	0.05
Scleroderma	77	0.01	380	0.05	457	0.03
HIV infection/AIDS	120	0.02	70	0.01	190	0.01
Multiple Chronic Conditions						
Zero	259384	37.08	207992	28.64	467376	32.78
One	140699	20.11	135274	18.63	275973	19.36
Two	93036	13.30	102936	14.17	195972	13.74
Three	64841	9.27	79405	10.93	144246	10.12
Four	46238	6.61	60525	8.33	106763	7.49
Five or more	95355	13.63	140138	19.30	235493	16.52

AIDS – Acquired immune deficiency syndrome; COPD- Chronic Obstructive Pulmonary Disease; HIV - Human Immunodeficiency virus

Figure 5.3-1 Pattern of comorbidities (numbers in the cell represents among the people having the column conditions what is the percentage with row conditions)





Prevalence (%)

BPH^- Benign prostatic hypertrophy (only for men); CHD- Chronic heart disease; CKD- Chronic kidney disease; COPD- Chronic obstructive pulmonary disease; GI- Gastrointestinal; IBD- Inflammatory bowel disease; IBS- Inflammatory bowel syndrome; PVD- Peripheral vascular diseases; SLE- Systemic lupus erythematous.

The pattern and combination of chronic conditions is depicted in Figure 5.3-1.

As an example of how to interpret this figure, in the top row (back pain) the second cell from the left (hypertension) has a value of 27 and this means that 27% of the people with

back pain have hypertension, whereas the first cell in the second to top row has the value 62 which means that 62% of the people with hypertension have back pain.

5.3.1.2 Clustering of comorbidities

Table 5.3-3 shows AIC, BIC, aBIC, entropy, log likelihood and distribution of probabilities in class for each model. There was an increase in BIC and adjusted BIC values when the number of classes increased from seven to eight. This suggests that the seven-class model was good, but the likelihood ratio tests rejected the seven-class model over six-class ($p<0.001$), which explains why the addition of one class from six to seven was not statistically significant. Also, one of the classes in the seven-class model comprised less than 1% of the sample. Therefore, the six-class model was selected as the optimal solution. All models had relatively low entropy (0.70 or close to this), which indicated that there was some overlap in the classification of classes.

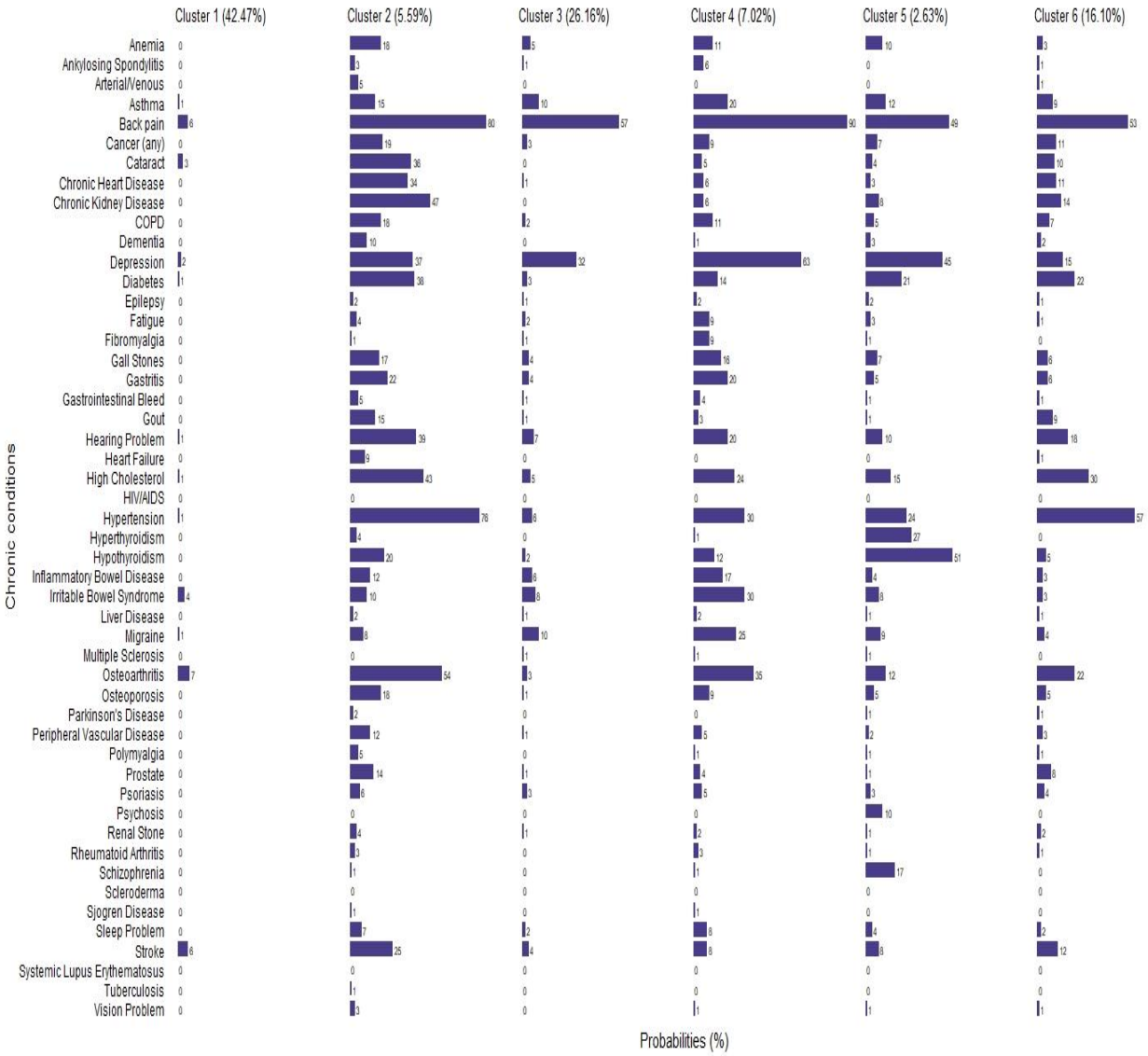
Table 5.3-3 Model statistics for different classes

Model	Log-likelihood	AIC	BIC	aBIC	Likelihood-ratio	Entropy	Parameters
1	-1.1E+07	21961579	21959584	21959434	3462723		47
2	-1E+07	20157016	20153876	20153574	1656326	0.75	95
3	-9958232	19921034	19918515	19918061	1420277	0.65	143
4	-9915683	19836036	19834105	19833498	1335179	0.63	191
5	-9896793	19796128	19797013	19796253	1297398	0.63	239
6	**-9878616**	**19764272**	**19761347**	**19760435**	**1261044**	**0.64**	**287**
7	-9862727	19730398	19734619	19733532	1230675	0.6	335
8	-9854203	19732930	19737756	19736512	1233109	0.58	383
9	-9848211	19706904	19712334	19710935	1206984	0.59	431
10	-9842969	19692013	19698049	19696494	1191996	0.58	479

AIC- Akaike information criteria; Bayesian information criteria, aBIC- Adjusted Bayesian information criteria

Based on the probabilities of class membership, six patterns of multimorbidity clusters were identified as shown in Figure 5.3-2, specifically: (a) relatively healthy class (42.47%) with lowest posterior probabilities of most diseases; (b) a back pain class (26.16%) with highest posterior probabilities of back pain; (c) a metabolic syndrome class (16.10%) with highest probabilities of hypertension, high cholesterol and diabetes; (d) a pain and depression cluster (7.02%) having back pain and depression as leading contributing conditions; (e) a cardiovascular and musculoskeletal (5.59%) cluster with back pain, hypertension and OA as major contributors to the class; and (f) a thyroid cluster (2.63%) with highest posterior probabilities from hypothyroidism.

Figure 5.3-2 Six-class model of multimorbidity pattern



BPH- Benign Prostatic Hypertrophy; COPD- Chronic obstructive pulmonary disease.

Cluster 1 -Relative healthy; Cluster 2 multiple comorbidities led by backpain and hypertension; Cluster 3- Back pain; Cluster 4- musculoskeletal and mental health; Class 5- thyroid cluster; Class 6- metabolic syndrome

The clusters of conditions across different age group are given in the Appendix Table 17-20 (pages 327-330). Appendix Table 16 (page 326) describes the best model for different age groups in OA. The statistics suggest a three class model in the age group 20-39 years and five class clusters afterwards. Details of the posterior probabilities distribution of individual conditions are given in Appendix Table 17-20 (pages 327-330).

5.3.2 OA and Non-OA populations

For this analysis 221,807 people with OA and 221,807 age, sex and practice area matched non-OA controls were used. Study participant characteristics are given in the results section of Chapter 4, Table 4.3-1 (pages 114).

Leading conditions reported in OA cases were back pain (48.02%), hypertension (38.06%), depression (29.95%), high cholesterol (20.22%) and hearing problems (17.41%). A similar pattern was observed in the non-OA population. (Table 5.3-4) The prevalence of multimorbidity was 77% and 70% in the OA and non-OA population, respectively. Single (14.9%) and two morbidities (15.4%) were more common in the non-OA population compared to OA (single-11.5%, two morbidities- 13.1%), whereas the prevalence of four (13.5%) and five (36.8%) morbidities were higher in those with OA compared to non-OA (four morbidities-12.6%, five morbidities-27.0%). (Table 5.3-4)

Table 5.3-4. Distribution of chronic conditions in the OA, non-OA, and overall populations

Conditions	Non- OA (n=221807)	%	OA (n=221807)	%	Total (n=443614)	%
Back pain	75715	34.14	106535	48.02	182250	41.08
Hypertension	78032	35.18	84420	38.06	162451	36.62
Depression	49552	22.34	66431	29.95	116005	26.15
High Cholesterol	40435	18.23	44849	20.22	85263	19.22
Hearing	35733	16.11	38617	17.41	74350	16.76
Diabetes Mellitus	29012	13.08	33781	15.23	62816	14.16
Chronic Kidney Disease	29101	13.12	29678	13.38	58779	13.25
Stroke	27548	12.42	29523	13.31	57049	12.86
Coronary Heart Disease	23622	10.65	28924	13.04	52568	11.85
Asthma	22669	10.22	29234	13.18	51903	11.70
Cancer (any)	20606	9.29	24221	10.92	44849	10.11
Hypothyroidism	17811	8.03	20562	9.27	38373	8.65
Gastritis	15881	7.16	21404	9.65	37308	8.41
COPD	16746	7.55	20428	9.21	37175	8.38
Migraine	15194	6.85	19696	8.88	34912	7.87
Gall stones	13663	6.16	18210	8.21	31851	7.18
Irritable Bowel Syndrome	12710	5.73	18210	8.21	30920	6.97
Anaemia	12931	5.83	16170	7.29	29101	6.56
Inflammatory Bowel Disease	12576	5.67	16214	7.31	28791	6.49
Osteoporosis	12931	5.83	14063	6.34	27016	6.09
Benign prostatic Hypertrophy	12177	5.49	14551	6.56	26750	6.03
Gout	10292	4.64	15371	6.93	25641	5.78
Cataract	11490	5.18	13575	6.12	25064	5.65
Peripheral Vascular Disease	7919	3.57	10358	4.67	18277	4.12
Sleep Disorder	7675	3.46	10003	4.51	17700	3.99
Psoriasis	7253	3.27	8695	3.92	15926	3.59
Dementia	6765	3.05	7209	3.25	13974	3.15
Heart failure	4968	2.24	7541	3.40	12510	2.82
Fatigue	3815	1.72	4946	2.23	8739	1.97
Gastrointestinal bleed	3571	1.61	4880	2.20	8429	1.90
Ankylosing Spondylitis	3349	1.51	4880	2.20	8207	1.85
Hyperthyroid	3704	1.67	4170	1.88	7896	1.78
Polymyalgia	3039	1.37	4436	2.00	7453	1.68
Epilepsy	3150	1.42	3948	1.78	7098	1.60
Renal stone	3261	1.47	3726	1.68	7009	1.58
Rheumatoid Arthritis	1730	0.78	4436	2.00	6166	1.39
Fibromyalgia	1841	0.83	4259	1.92	6122	1.38
Arterial/Venous	2662	1.20	5501	2.48	5944	1.34
Vision problem	2506	1.13	2595	1.17	5102	1.15
Schizophrenia	2396	1.08	2418	1.09	4791	1.08
Liver Disease	1553	0.70	2196	0.99	3726	0.84
Parkinson's Disease	1575	0.71	1930	0.87	3505	0.79
Tuberculosis	1486	0.67	1664	0.75	3150	0.71
Psychosis	887	0.40	843	0.38	1730	0.39
Multiple sclerosis	821	0.37	688	0.31	1508	0.34
Sjogren's syndrome	377	0.17	621	0.28	1020	0.23
Systemic Lupus Erythematosus	155	0.07	244	0.11	399	0.09
Scleroderma	133	0.06	155	0.07	266	0.06
HIV infection/AIDS	22	0.01	22	0.01	44	0.01
Multiple chronic conditions						
Zero	33914	15.29	25508	11.50	59400	13.39
One	33027	14.89	24510	11.05	57537	12.97
Two	34180	15.41	28946	13.05	63126	14.23
Three	32827	14.80	31319	14.12	64147	14.46
Four	27926	12.59	29944	13.50	57847	13.04
Five or more	59932	27.02	81581	36.78	141513	31.90

AIDS – Acquired immune deficiency syndrome; COPD- Chronic Obstructive Pulmonary Disease; HIV - Human Immunodeficiency virus

5.3.3 Clustering of comorbidities in OA and non-OA

Firstly, cluster analysis was done in the separate OA and non-OA populations followed by the same cluster analysis in men and women and different age groups of 20-39, 40-59, 60-79 and >=80 years in both groups.

5.3.3.1 Clustering in OA

Table 5.3-5 summarises the AIC, BIC, aBIC, entropy, log likelihood and distribution of probabilities in each class for each model. It shows, there was a gradual decline in BIC and adjusted BIC values with increasing number of classes. Models for class 7 onwards had one group with <1% sample size, so these were not selected. Between class 6 and five, the change in likelihood ratio was less than 1%. This suggests that the five-class model was good and explains why the addition of one class from five to six was not statistically significant. Therefore, the five-class model was selected as the optimal solution. All models had relatively low entropy (0.60 or thereabouts), which indicated that there was some overlap in the classification of classes.

Table 5.3-5. Model statistics for different classes in OA

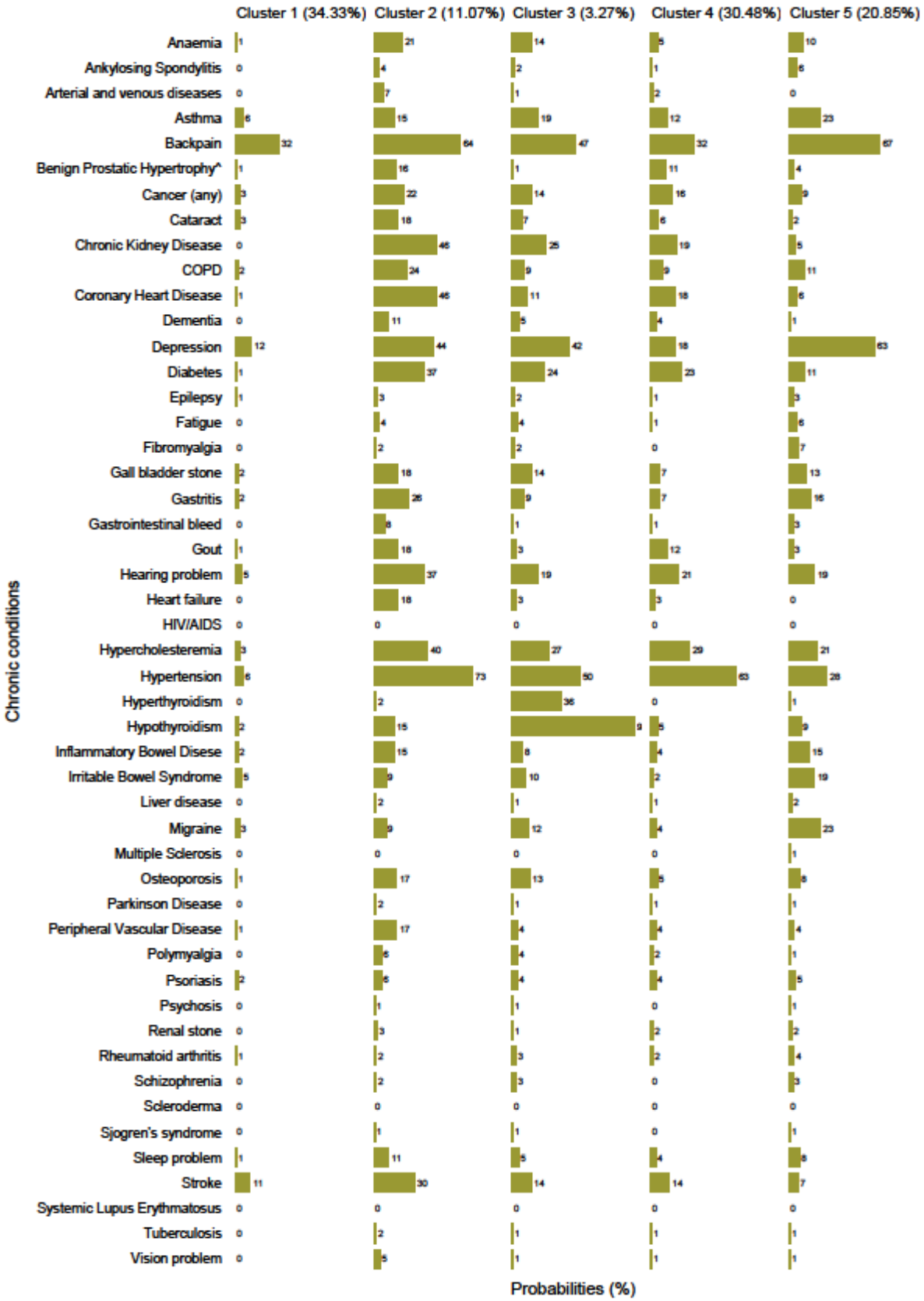
Model	Parameters	log-likelihood	BIC	aBIC	AIC	Entropy
1		-2393224	-	4786905	4786548	
2	93	-2317774	4636791	4636470	4635750	0.69
3	140	-2297448	4596767	4596284	4595200	0.67
4	187	-2288014	4578527	4577882	4576434	0.64
5	**234**	**-2284016**	**4571158**	**4570351**	**4568540**	**0.65**
6	281	-2280824	4565402	4564433	4562258	0.63
7	328	-2277791	4559964	4558833	4556294	0.63
8	375	-2275128	4555266	4553972	4551070	0.62
9	422	-2273244	4552125	4550670	4547403	0.61
10	469	-2271431	4549128	4547511	4543881	0.60

AIC- Akaike information criteria; BIC- Bayesian information criteria, aBIC- Adjusted Bayesian information criteria

Cluster 1 was the relative healthy group sharing maximum class size (34.33%). Cluster 2 had a 11.07% population of total OA and was dominated by hypertension (PP73%) and back pain (PP-64%). In cluster 3, the smallest cluster (3.04%) in the OA, thyroid problem was the leading contributor (PP-90%) among all conditions. Cluster 4 shared one third of the total population size, in which hypertension was the foremost chronic condition (PP-63%). Cluster 5 had the strongest contribution from both back pain (PP-67%) and depression (PP-63%). (Figure 5.3-3)

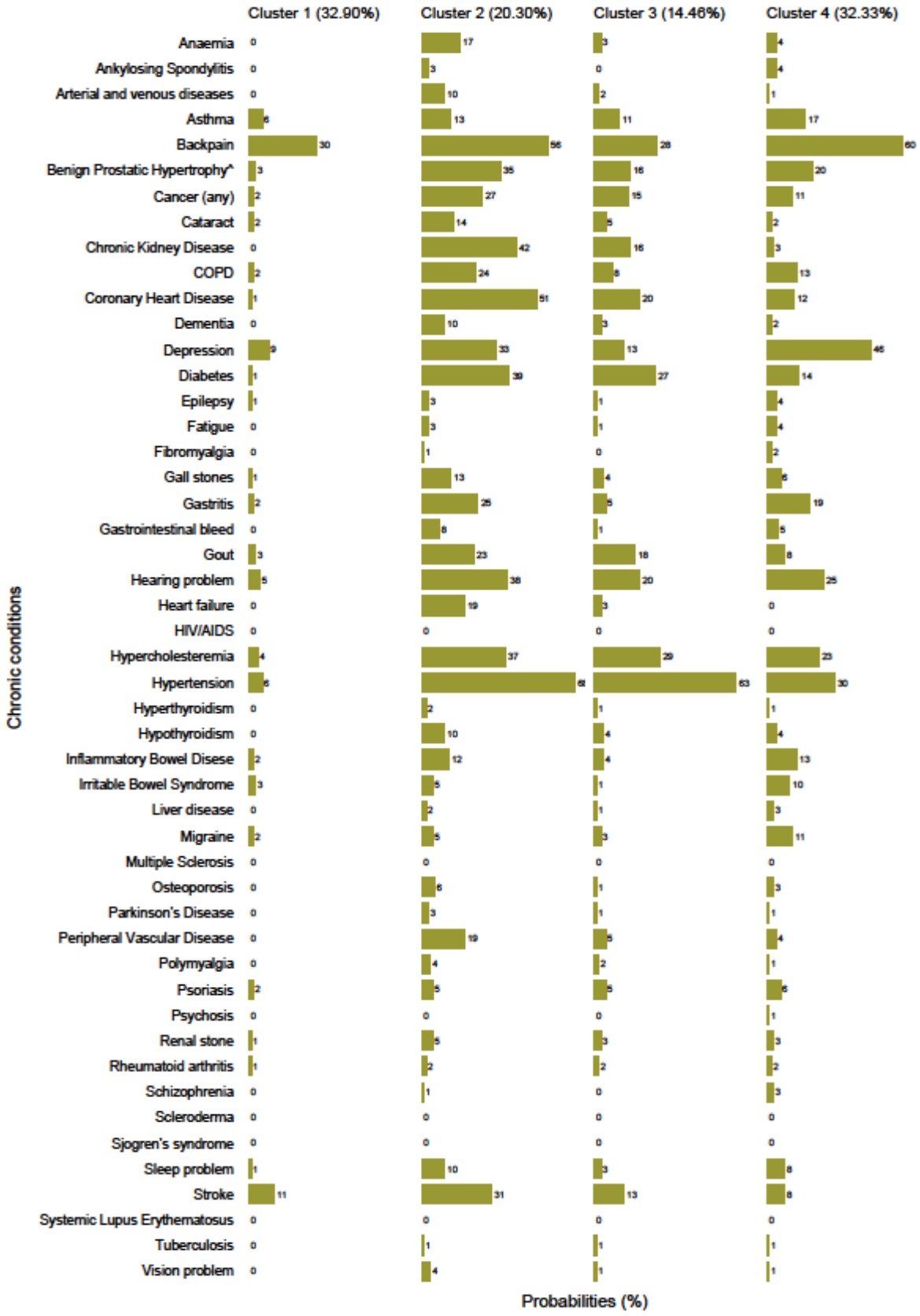
Summary statistics for model selection of clusters by gender is given in Appendix Table 21 (page 325). There were four clusters in men, and five in women. Thyroid leading cluster was more prominent in women, which was absent in men. (Figure 5.3-4 and Figure 5.3-5)

Figure 5.3-3 Five-class model clusters in the OA population overall



Cluster 1- Relative healthy; Cluster 2- CV-MSK; Cluster 3 Thyroid ; Cluster 4 CV; Cluster 5 MSK-MH
CV- Cardiovascular; MH- Mental health; MSK- Musculoskeletal. COPD- Chornic obstructive pulmonary diseases; ^only for men

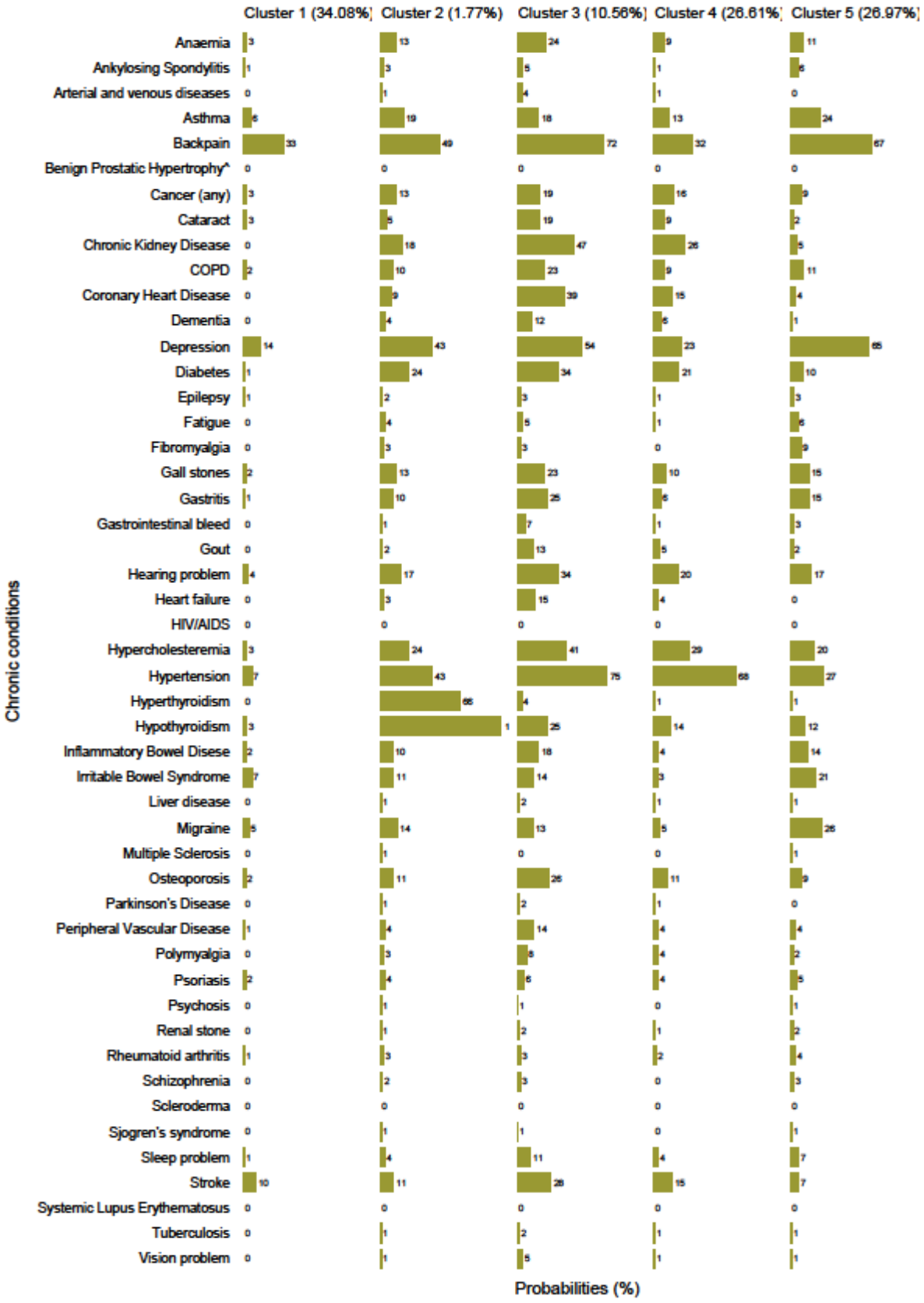
Figure 5.3-4 Four-class model clusters in the OA population in Men



COPD: Chronic obstructive pulmonary disease; ^only for men

Cluster 1- Relative healthy; Cluster 2- CV-MSK; Cluster 3- CV; Cluster 4- MSK-MH
CV- Cardiovascular; MSK-Musculoskeletal; MH-Mental health

Figure 5.3-5 Five-class model clusters in the OA population in women



COPD: Chronic obstructive pulmonary disease; ^only for men

Cluster 1- Relative healthy; Cluster 2- Thyroid; Cluster 3- CV-MSK; Cluster 4- CV; Cluster 5- MSK-MH
CV- Cardiovascular; MSK-Musculoskeletal; MH-Mental health

Similar approaches were used to identify clusters in different age groups within the OA population. Details of the model summary are provided in the Appendix Table 22 (page 332). A three class model was selected for age group 20-39 , and a five class model was selected for 40-59, 60-79 and >=80 years.

In the age group 20-39 years the three class model identified MSK, MH and relatively healthy clusters. (Table 5.3-6) The five clusters in the age group 40-59 years were MSK (31%), CV-MSK (28%), MSK-CV-MH (10%), thyroid (5%) and relatively healthy (26%). Individual probabilities of chronic conditions across the cluster can be seen in Table 5.3-7. In the age group 60-79 years five clusters were found, namely CV (27%), MSK (30%), CV-MSK (14%), MSK-CV-MH (13%), and relatively healthy (16%). Detailed contributions of the chronic conditions are provided in Table 5.3-8. Clusters identified in the elderly age had a slightly different pattern. Prominent clusters were CV-MSK (15%), MSK-CV-MH (17%), CV-Renal (25%), MSK-Hearing (29%), and relatively healthy group (14%). Details of the probailities by conditions are shown in Table 5.3-9.

Table 5.3-6. Posterior probabilities distribution of chronic conditions to three-class model in the OA population (20-39 years)

	Cluster 1 (56.06%)	Cluster 2 (36.32%)	Cluster 3 (7.62%)
Anaemia	2.27	9.98	10.07
Ankylosing Spondylitis	0.52	4.39	1.71
Arterial/Venous	0.02	0.28	0.78
Asthma	10.53	28.99	23.74
Back pain	35.98	**83.77**	47.40
Benign prostatic hypertrophy	0.41	0.61	2.41
Cancer (any)	0.73	2.42	5.12
Cataract	0.12	0.49	0.74
Chronic Heart Disease	0.41	1.13	10.8
Chronic Kidney Disease	0.6	1.34	11.08
COPD	1.06	5.66	6.3
Dementia	0	0.24	0.15
Depression	15.45	54.74	**66.01**
Diabetes	2.25	6.59	39.54
Epilepsy	1.67	3.11	5.93
Fatigue	0.27	5.8	1.75
Fibromyalgia	0.01	8.85	5.45
Gall stones	1.18	9.11	8.91
Gastritis	1.69	12.1	12.75
Gastrointestinal Bleed	0.28	3.34	4.87
Gout	2.42	1.98	13.93
Hearing Problem	4.35	11.65	10.92
Heart Failure	0.02	0.02	1.58
High Cholesterol	2.29	4.99	40.78
HIV/AIDS	0.05	0.04	0.23
Hypertension	4.08	6.99	58.50
Hyperthyroidism	0.23	1.47	5.73
Hypothyroidism	1.38	8.09	16.21
Inflammatory Bowel Disease	3.37	17.26	12.69
Irritable Bowel Syndrome	3.83	17.02	5.71
Liver Disease	0.3	1.32	2.16
Migraine	4.97	29.18	16.32
Multiple Sclerosis	0.14	0.79	0.37
Osteoporosis	0.22	1.1	2.43
Parkinson's Disease	0.03	0.06	0.28
Peripheral Vascular Disease	0.94	3.45	5.27
Polymyalgia	0	0.18	0
Psoriasis	3.55	6.12	7.75
Psychosis	0	0.73	2.6
Renal Stone	0.43	1.51	3.89
Rheumatoid Arthritis	2.37	5.73	5.16
Schizophrenia	0	2.45	5.14
Scleroderma	0.07	0.07	0.24
Sjogren's Syndrome	0.04	0.41	0.74
Sleep Problem	0.75	6.83	6.71
Stroke	6.06	5.28	9.75
Systemic Lupus Erythematosus	0.04	0.2	0.38
Tuberculosis	0.34	0.39	1.93
Vision Problem	0.1	0.71	1.7

COPD- Chronic obstructive pulmonary disease; *Cluster 1 – Relative healthy; Cluster 2 MSK; Cluster 3- MH

MSK – Musculoskeletal; MH- Mental health

Table 5.3-7. Posterior probabilities distribution of chronic conditions to the five-class model in the OA population (40-59 years)

	Cluster 1 (26.29%)	Cluster 2 (27.6%)	Cluster 3 (5.14%)	Cluster 4 (9.79%)	Cluster 5 (31.17%)
Anaemia	0.93	4.12	12.81	16.72	8.12
Ankylosing Spondylitis	0	0.93	2.55	5.13	4.62
Arterial/Venous	0	0.69	0.39	2.8	0.17
Asthma	3.94	13.04	19.7	25.22	20.86
Back pain	21.48	**63.85**	50.4	**91.15**	**84.92**
Benign prostatic hypertrophy	0.58	6.81	0.71	7.5	2.22
Cancer (any)	1.5	8.39	8.8	12.9	6.31
Cataract	0.63	1.03	1.25	3.03	0.72
Chronic Heart Disease	0.23	11.16	4.48	28.26	2.65
Chronic Kidney Disease	0	9.23	12.26	23.89	1.57
COPD	0.83	6.63	8.64	17.87	8.5
Dementia	0.04	0.43	1.49	2.12	0.24
Depression	7.01	24.06	54.53	**67.93**	55.48
Diabetes	0.52	21.54	20.47	41.43	4.29
Epilepsy	0.45	1.45	3.4	3.97	2.68
Fatigue	0.11	0.57	4.67	6.76	4.4
Fibromyalgia	0.02	0.11	4.3	9.38	5.74
Gall stones	0.79	5.04	11.61	19.87	9.7
Gastritis	0.61	5.88	8.21	26.93	11.5
Gastrointestinal Bleed	0.08	1.13	1.34	6.69	2.1
Gout	0.95	12.03	1.97	12.66	1.56
Hearing Problem	2.69	13.19	14.57	23.64	14.73
Heart Failure	0.02	0.91	0.35	5.23	0
High Cholesterol	1.58	29.53	21.13	45.82	11.2
HIV/AIDS	0	0.02	0.03	0.04	0.02
Hypertension	3.2	**53.13**	30.44	**66.92**	13.58
Hyperthyroidism	0.06	0	29.41	1.62	0
Hypothyroidism	1.18	4.33	**61.18**	14.12	5.94
Inflammatory Bowel Disease	0.87	4.97	8.7	20.35	11.17
Irritable Bowel Syndrome	5.83	2.09	11.87	19.28	16.97
Liver Disease	0.11	1.09	1.42	3.61	1.28
Migraine	1.61	5.35	15.26	23.25	21.08
Multiple Sclerosis	0.11	0.17	0.65	0.63	0.71
Osteoporosis	0.34	1.77	4.97	8.01	4.77
Parkinson's Disease	0.06	0.25	0.6	0.48	0.19
Peripheral Vascular Disease	0.39	2.89	3.43	10.89	3.25
Polymyalgia	0.05	0.53	1.18	2.03	0.7
Psoriasis	1.5	0	4.7	6.75	4.29
Psychosis	0.02	2.25	7.51	0.49	0
Renal Stone	0.21	1.89	0.79	3.64	1.54
Rheumatoid Arthritis	0.43	0.12	4.06	4.32	3.01
Schizophrenia	0.03	0.04	12.74	2.16	0.84
Scleroderma	0	0.07	0.16	0.32	0.05
Sjogren's Syndrome	0.02	2.85	1.11	0.73	0.31
Sleep Problem	0.12	9.11	5	12.88	5.76
Stroke	9.48	0.03	9.79	17.41	5.79
Systemic Lupus Erythematosus	0.02	0.73	0.14	0.51	0.18
Tuberculosis	0.1	0.38	0.68	1.28	0.7
Vision Problem	0.06	0.38	1.5	2.04	0.41

COPD- Chronic obstructive pulmonary disease; Cluster 1- Relative healthy; Cluster 2 CV-MSK; Cluster 3 Thyroid (metabolic); Cluster 4 MSK-MH-CV; Cluster 5 MSK

CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Table 5.3-8. Posterior probabilities distribution of chronic conditions to the five-class model in the OA population (60-79 years)

	Cluster 1 (15.72%)	Cluster 2 (14.35%)	Cluster 3 (12.66%)	Cluster 4 (27.10%)	Cluster 5 (30.16%)
Anaemia	0.16	19.05	17.46	5.12	4.59
Ankylosing Spondylitis	0	2.93	5.71	0.32	3.26
Arterial/Venous	0.02	9.1	1.33	1.68	0.82
Asthma	0.98	12.05	23.17	11.92	12.9
Back pain	4.91	**80.88**	**87.98**	53.36	**75.46**
Benign prostatic hypertrophy	7.5	0.58	0.71	2.22	6.81
Cancer (any)	1.25	25.01	16.32	15.76	16.1
Cataract	6.03	15.02	12.69	6.9	4.25
Chronic Heart Disease	0.44	53.58	21.03	18.24	8.38
Chronic Kidney Disease	0	44.71	32.12	26.66	2.93
COPD	0.55	27.42	14.41	7.14	12.24
Dementia	0.17	9.13	9.64	4.04	3.09
Depression	1.16	34.17	**65.82**	13.87	30.15
Diabetes	0.51	40.16	26.14	27.8	6.14
Epilepsy	0.11	2.65	2.8	1.06	1.86
Fatigue	0.04	2.96	6.96	0.65	1.99
Fibromyalgia	0	0.52	5.46	0.03	0.97
Gall stones	0.46	14.52	22.06	6.79	8.5
Gastritis	0.2	25.49	22.48	4.4	10.67
Gastrointestinal Bleed	0	7.67	4.37	0.69	1.97
Gout	0.3	22.55	6.04	12.24	3.58
Hearing Problem	1.4	36.82	26.94	18.89	23.1
Heart Failure	0.1	20.75	4.25	3.33	0.49
High Cholesterol	1.17	41.73	37.75	30.59	18.97
HIV/AIDS	0.01	0.01	0.01	0.01	0
Hypertension	3.97	**71.55**	**64.32**	**73.7**	29.17
Hyperthyroidism	0.03	0.99	9.5	1.98	0.56
Hypothyroidism	0.59	10.16	31.94	10.99	6.51
Inflammatory Bowel Disease	0.1	12.02	16.71	2.97	7.54
Irritable Bowel Syndrome	4.97	5.07	21.66	1.58	8.72
Liver Disease	0.01	2.04	1.86	0.67	0.79
Migraine	0.17	5.62	17.47	3.13	9.39
Multiple Sclerosis	0.01	0.29	0.4	0.15	0.43
Osteoporosis	0.24	9.86	24.4	4.14	9.74
Parkinson’s Disease	0.1	2.49	2.28	1.08	1.45
Peripheral Vascular Disease	0.13	19.56	7.2	4.21	3.08
Polymyalgia	0.1	4.5	6.51	2.57	2.09
Psoriasis	0.29	5.56	5.14	4.05	4.18
Psychosis	0	0.03	2.25	0	0.11
Renal Stone	0.1	4.44	1.81	2.04	1.63
Rheumatoid Arthritis	0.09	1.68	4.02	1.42	1.95
Schizophrenia	0.04	0.53	5.32	0.27	0.67
Scleroderma	0.01	0.18	0.23	0	0.05
Sjogren’s Syndrome	0	0.33	1.38	0.03	0.31
Sleep Problem	0.04	8.22	9.2	2.28	4.22
Stroke	16.56	30.43	18.5	14.86	8.64
Systemic Lupus Erythematosus	0	0.06	0.42	0.04	0.09
Tuberculosis	0.02	1.48	1.49	0.6	1.01
Vision Problem	0.03	3.71	2.67	0.83	0.79

COPD- Chronic obstructive pulmonary disease; Cluster 1- Relative healthy; Cluster 2 CV-MSK; Cluster 3 CV-MSK-MH; Cluster 4 CV; Cluster 5 MSK

CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Table 5.3-9. Posterior probabilities distribution of chronic conditions to the five-class model in the OA population (>=80 years)

	Cluster 1* (14.00%)	Cluster 2* (14.97%)	Cluster 3* (17.45%)	Cluster 4* (25.13%)	Cluster 5* (28.44%)
Anaemia	0.02	17.85	22.74	9.63	8.75
Ankylosing Spondylitis	0	1.89	4.59	0	1.79
Arterial/Venous	0	11.6	3.59	1.1	2.14
Asthma	0.24	10.19	14.9	9.73	9.27
Back pain	1.37	**73.49**	**88.38**	44.14	**61.42**
Benign prostatic hypertrophy	6.81	2.22	0.58	7.5	0.71
Cancer (any)	0.57	27.57	17.21	16.69	21.89
Cataract	28.23	33.62	46.62	27.37	25.06
Chronic Heart Disease	0.59	54.58	34.06	22.88	17.75
Chronic Kidney Disease	0	54.78	46.42	**56.82**	7.33
COPD	0.48	23.14	19.04	5.79	14.59
Dementia	0.73	8.97	18.02	13.34	25.26
Depression	0	20.63	**69.19**	12.74	24.1
Diabetes	0.32	32.93	19.8	23.05	6.82
Epilepsy	0	1.36	2.24	1.67	1.8
Fatigue	0.01	2.1	4.73	0.7	1.36
Fibromyalgia	0	0	1.09	0	0.09
Gall stones	0.09	15.42	19.83	7.44	6.9
Gastritis	0.13	22.91	19.61	4.01	10.62
Gastrointestinal Bleed	0.09	8.49	5.96	1.13	3.71
Gout	0.14	25.27	8.34	9.77	3.83
Hearing Problem	1.06	45.7	43.14	27.59	**47.33**
Heart Failure	0.28	27.87	18.16	10.8	9.21
High Cholesterol	0	29.22	25.85	20.59	6.6
Hypertension	3.7	**72.00**	**71.59**	**80.56**	39.42
Hyperthyroidism	0.02	0.74	8.32	3.53	0.12
Hypothyroidism	0.43	8.97	31.27	15.91	4.7
Inflammatory Bowel Disease	0.04	12.03	15.67	2.41	6.93
Irritable Bowel Syndrome	2.87	4.43	13.35	1.29	3.87
Liver Disease	0.04	0.57	1.01	0.49	0.12
Migraine	0	2.9	8.92	1.84	3.43
Multiple Sclerosis	0	0.05	0.16	0.09	0.08
Osteoporosis	0	5.71	38.82	11.18	13.28
Parkinson’s Disease	0.18	2.10	2.24	0.86	2.98
Peripheral Vascular Disease	0.09	21.67	11.58	4.91	4.66
Polymyalgia	0.06	5.02	11.65	5.84	5.37
Psoriasis	0.15	4.39	5.65	2.33	2.88
Psychosis	0	0	1.19	0	0.36
Renal Stone	0.01	3.44	1.1	0.74	1.52
Rheumatoid Arthritis	0.12	1.15	2.47	0.8	1.16
Schizophrenia	0	0.19	2.43	0.08	1.46
Scleroderma	0	0	0.15	0.02	0.03
Sjogren’s Syndrome	0	0.07	0.81	0.13	0.19
Sleep Problem	0.16	10.63	14.46	4.23	13.11
Stroke	25.19	32.09	31.97	23.43	20.82
Systemic Lupus Erythematosus	0	0	0.37	0	0
Tuberculosis	0.02	1.05	2.14	0.59	1.01
Vision Problem	0.02	5.95	7.52	1.61	5.48

COPD- Chronic obstructive pulmonary disease; Cluster 1- Relative healthy; Cluster 2 CV-MSK; Cluster3 MSK-CV-MH; Cluster 4 CV-Renal; Cluster 5 MSK-Hearing

CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

The distribution of socio-demographic characteristics across the clusters in all OA cases is provided in Table 5.3-10. The proportion of women was higher in all the clusters except in cluster 4 (CV). The mean age was highest in cluster 2 (70.92 years) followed by cluster 4 (65.67 years) and the lowest was found in the relative healthy group (56.83 years). Smoking prevalence was highest (54%) in cluster 2 (CV-MSK). The prevalence of obesity was 37.7% in cluster 3 (metabolic led by Thyroid), followed by 36% in cluster 4 (CV). The mean number of chronic conditions was highest for CV-MSK cluster (8.62; SD 1.97) and thyroid cluster (6.18; SD 2.00). (Table 5.3-10)

Table 5.3-10. Subject characteristics across the clusters in the OA population

	Relative Healthy (N=76150) n(%)	CV-MSK (N=24566) n(%)	Thyroid (N=7268) n(%)	CV (N=67605) n(%)	MSK-MH (N=46262) n(%)
Gender					
Men	33416 (43.88)	11153 (45.40)	823 (11.32)	36331 (53.74)	12189 (26.35)
Women	42734 (56.12)	13413 (54.60)	6445 (88.68)	31274 (46.26)	34073 (73.65)
Age					
<40 years	7254 (9.53)	85 (0.35)	167 (2.30)	984 (1.46)	3772 (8.15)
40-59 years	38906 (51.09)	3635 (14.80)	2583 (35.54)	20123 (29.77)	25866 (55.91)
60-79 years	26774 (35.16)	15923 (64.82)	3827 (52.66)	39292 (58.12)	15537 (33.58)
>=80 years	3216 (4.22)	4923 (20.04)	691 (9.51)	7206 (10.66)	1087 (2.35)
Smoking					
Never smoked	42708 (56.08)	11290 (45.96)	4140 (56.96)	36075 (53.36)	23336 (50.44)
Current smoker	15071 (19.79)	4004 (16.30)	1151 (15.84)	9926 (14.68)	11551 (24.97)
Ex-smoker	18362 (24.11)	9270 (37.74)	1976 (27.19)	21599 (31.95)	11369 (24.58)
Alcohol use					
Never	12763 (16.76)	6609 (26.90)	2090 (28.76)	12830 (18.98)	9811 (21.21)
Ex-drinker	1350 (1.77)	1183 (4.82)	234 (3.22)	1733 (2.56)	1540 (3.33)
Current (1-9)	28957 (38.03)	7452 (30.33)	2471 (34.00)	22473 (33.24)	16229 (35.08)
Current (>=10)	16240 (22.10)	3606 (14.68)	683 (9.40)	14998 (22.18)	7646 (16.53)
Current (Unknown)	16827 (22.10)	5707 (23.23)	1787 (24.59)	15551 (23.00)	11013 (23.81)
BMI					
Underweight	912 (1.20)	468 (1.91)	128 (1.76)	739 (1.09)	784 (1.69)
Normal	25096 (32.96)	6300 (25.65)	1932 (26.58)	16076 (23.78)	14160 (30.61)
Overweight	28665 (37.64)	9219 (37.53)	2466 (33.93)	26452 (39.13)	15909 (34.39)
Obese	21430 (28.14)	8575 (34.91)	2742 (37.73)	24316 (35.97)	15404 (33.30)
Mean age (SD)	56.83 (12.93)	70.92 (10.31)	63.99 (11.91)	65.57 (11.35)	56.13 (11.63)
Mean BMI (SD)	27.67 (5.38)	28.52 (5.67)	28.95 (6.21)	28.80 (5.50)	28.30 (5.97)
Mean CC (SD)	1.11 (0.98)	8.62 (1.97)	6.18 (2.00)	4.17 (1.47)	4.93 (1.70)

BMI- Body mass index; CC- Chronic conditions; SD -Standard deviation
CV-Cardiovascular; MH- Mental health; MSK- Musculoskeletal

Table 5.3-11 describes the association of patient characteristic with cluster membership having the healthy cluster as the reference group. Women had a higher risk than men of being in metabolic cluster (OR 5.45; 95% CI 5.05-5.88) and cluster 5 (OR 2.25; 95% CI 2.19-2.31) when compared to the gender ratio in the healthy cluster. Being a current

smoker had an increased OR when compared to the smoker to non-smoker ratio in the healthy cluster. Obesity was found to be significantly associated with all the clusters when compared to the ratio in the healthy group.

Table 5.3-11. Multinomial regression for association with clusters in the OA population

Variables	Relative Healthy OR (95% CI)	CV-MSK OR (95% CI)	Thyroid OR (95% CI)	CV OR (95% CI)	MSK-MH OR (95% CI)
Gender					
Men	1	Reference	Reference	Reference	Reference
Women	1	0.82 (0.79-0.84)*	5.55 (5.14-5.99)*	0.64 (0.62-0.65)*	2.25 (2.19-2.31)*
Age					
<40 years	1	Reference	Reference	Reference	Reference
40-59 years		7.90 (6.36-9.82)*	2.67 (2.28-3.13)*	3.76 (3.51-4.03)*	1.25 (1.19-1.30)*
60-79 years		52.36 (42.22-64.93)*	5.73 (4.89-6.71)*	11.14 (10.40-11.94)*	1.11 (1.06-1.16)*
>=80 years		94.61 (74.16-114.53)	7.99 (6.70-9.54)*	19.99 (18.46-21.65)*	0.62 (0.57-0.68)*
Smoking					
Never smoked	1	Reference	Reference	Reference	Reference
Current smoker	1	1.52 (1.44-1.58)*	1.13 (1.05-1.21)*	0.99 (0.97-1.03)	1.48 (1.44-1.53)*
Ex-smoker	1	1.78 (1.72-1.84)*	1.38 (1.30-1.46)*	1.16 (1.13-1.19)*	1.27 (1.23-1.31)*
Alcohol use					
Never	1	Reference	Reference	Reference	Reference
Ex-drinker	1	1.70 (1.55-1.86)*	1.44 (1.24-1.68)*	1.27 (1.17-1.38)*	1.59 (1.47-1.73)*
Current (1-9 units)	1	0.51 (0.49-0.54)*	0.64 (0.59-0.68)*	0.76 (0.76-0.81)*	0.79 (0.77-0.82)*
Current (>=10 units)	1	0.48 (0.46-0.51)*	0.52 (0.47-0.57)*	0.89 (0.86-0.93)*	0.81 (0.78-0.84)*
Current (Unknown)	1	0.69 (0.66-0.72)*	0.77 (0.72-0.83)*	0.92 (0.88-0.95)*	0.90 (0.87-0.94)*
BMI					
Normal	1	Reference	Reference	Reference	Reference
Underweight	1	1.55 (1.45-1.58)*	1.40 (1.15-1.70)*	1.21 (1.09-1.34)*	1.36 (1.23-1.50)*
Overweight	1	1.33 (1.28-1.38)*	1.31 (1.23-1.39)*	1.37 (1.33-1.40)*	1.10 (1.06-1.13)*
Obese	1	2.11 (2.03-2.20)*	1.88 (1.77-2.00)*	2.03 (1.97-2.09)*	1.26 (1.22-1.30)*

CI- Confidence interval; OR- Odds ratio; *P value <0.05

CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

5.3.3.2 Clustering in non-OA

Model statistics for different clusters in the non-OA population are given in Table 5.3-12. There was a gradual decline in BIC and adjusted BIC values with increase in number of classes. The model for class 6 onwards had one group with <1% sample size, so those were not selected. The change in likelihood ratio from class four to five was more than 1%. This suggested that the five-class model was the best and this was selected as the optimal solution. All models had relatively low entropy (0.58-0.68), which indicated that there was some overlap in the classification of classes.

Table 5.3-12. Model statistics for different clusters in non-OA

Class	Parameter	log-likelihood	BIC	aBIC	AIC	Entropy
2	93	-2074549	-	4149544	4149195	
3	140	-2002933	4007085	4006770	4006064	0.68
4	187	-1987575	3976984	3976511	3975448	0.64
5	**281**	**-1979733**	**3961915**	**3961282**	**3959863**	**0.60**
6	328	-1975968	3955001	3954210	3952434	0.60
7	375	-1972613	3948908	3947957	3945825	0.62
8	422	-1970319	3944935	3943826	3941337	0.60
9	469	-1968272	3941455	3940186	3937341	0.59
10	516	-1966666	3938858	3937431	3934229	0.58

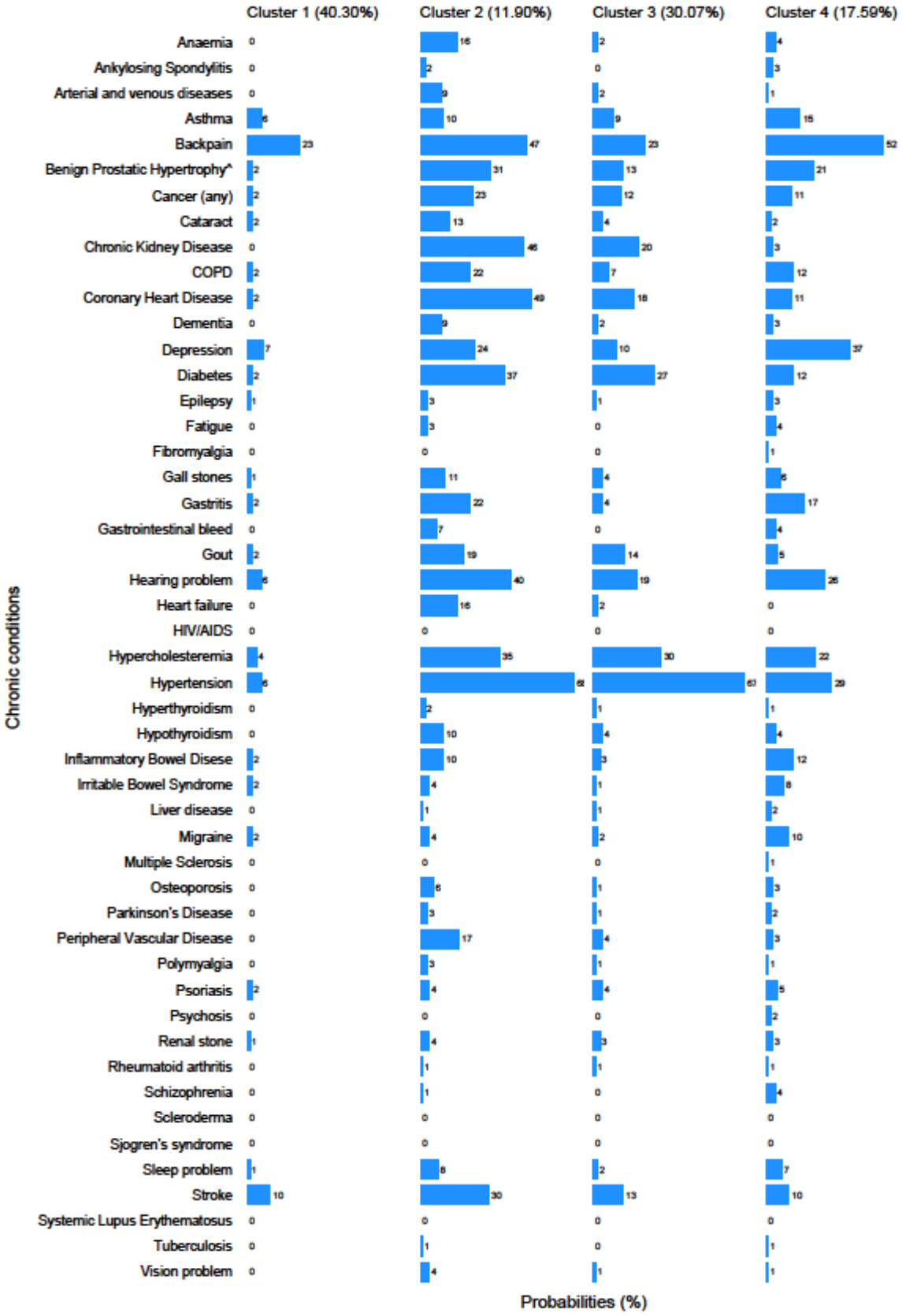
AIC- Akaike information criteria; Bayesian information criteria, aBIC- Adjusted Bayesian information criteria

LCA identified five clusters in the non-OA population, like OA but with different class sizes. Cluster 1 was relatively healthy (40.6%) with lowest contribution of all chronic conditions. Cluster 2 was the smallest cluster and was led by thyroid (PP-54%) and depression (PP-44%). Cluster 3 shared 10.09% of the total population predominantly with hypertension (PP-72%) and back pain (PP-54%). Nearly 30% of the total population was grouped within cluster 4 with the highest contribution from hypertension (PP-66%). Cluster 5 (16%) was led by back pain (PP-61%) and depression (PP-51%). (Figure 5.3-6)

Figure 5.3-6 Five-class model clusters in the non-OA population overall

COPD: Chronic obstructive pulmonary disease; ^only for men
Cluster 1- Relative healthy; Cluster 2- Thyroid; Cluster 3- CV-MSK; Cluster 4- CV; Cluster 5- MSK-MH
CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Figure 5.3-7 Five-class model clusters in the non-OA population in men



COPD: Chronic obstructive pulmonary disease; ^only for men

Cluster 1- Relative healthy; Cluster 3- CV-MSK; Cluster 4- CV; Cluster 4- MSK-MH
CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Figure 5.3-8 Five-class model clusters in the non-OA population in women

COPD: Chronic obstructive pulmonary disease; ^only for men

Cluster 1- Relative healthy; Cluster 2- Thyroid ; Cluster 3- CV-MSK; Cluster 4- CV; Cluster 5- MSK-MH; CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Across gender, in people without OA there were four clusters in men and five in women. Details of the summary statistics of these clusters are given in Appendix Table 23 (page 333). As in people with OA, women had an additional cluster led by thyroid disorders. (Figure 5.3-7 and Figure 5.3-8)

Age group analysis in the non-OA population found three clusters, specifically healthy (59.7%), MSK (29.7%) and MH (10.6%) in the age group 20-39 years. (Table 5.3-13) However, in the age group 40-59 years, five classes were identified, specifically relatively healthy (36%), MSK (32%), CV (22%), MSK-MH (7%) and thyroid (3%). (Table 5.3-14) In the age group 60-79 years the five clusters identified were relatively healthy (29.8%), MSK (26.5%), CV (20%), CV-MSK (14%) and MSK-MH (10%). (Table 5.3-15) Similarly, in the age group of $\geq$ 80 years, five distinct clusters were found, namely relatively healthy (28%), MSK (22%), CV-Renal (19%), CV-MSK-Renal (16%), and Hearing- vision (15%). (Table 5.3-16) Details of the summary statistics of the models are given in Appendix Table 24 (page 334).

Table 5.3-13. Posterior probabilities distribution of chronic conditions to three-class model in the non-OA population aged 20-39 years

	Cluster 1 (59.6%)	Cluster 2 (10.6%)	Cluster 3 (29.7%)
Anaemia	1.96	8.83	8.38
Ankylosing Spondylitis	0.03	0.86	3.24
Arterial/Venous	0	0.41	0.24
Asthma	8.62	19.48	21.9
Back pain	26.86	50.61	**77.24**
Benign prostatic hypertrophy^	0.26	1.83	0.77
Cancer (any)	0.9	4.36	1.92
Cataract	0.16	0.22	0.18
Chronic Heart Disease	0.37	6.32	0.74
Chronic Kidney Disease	0.34	7.62	0.67
COPD	1.07	3.96	4.14
Dementia	0	0.18	0.01
Depression	10.73	**55.74**	44.82
Diabetes	1.52	28.27	2.95
Epilepsy	0.66	3.2	2.24
Fatigue	0.35	2.77	4.13
Fibromyalgia	0	1.5	2.8
Gall stones	0.57	4.97	7.85
Gastritis	0.8	6.75	9.78
Gastrointestinal Bleed	0.23	1.81	2.63
Gout	0.81	4.75	1.09
Hearing Problem	3.34	13.88	9.99
Heart Failure	0.03	0.78	0
High Cholesterol	2.3	23.78	3.3
HIV/AIDS	0.01	0.07	0
Hypertension	2.83	31.12	5.53
Hyperthyroidism	0.07	8.83	0.17
Hypothyroidism	0.71	18.33	3.47
Inflammatory Bowel Disease	2.39	8.65	14.55
Irritable Bowel Syndrome	2.44	6.65	11.45
Liver Disease	0.35	1.71	0.86
Migraine	3.4	12.51	23.21
Multiple Sclerosis	0.19	0.48	0.55
Osteoporosis	0.09	1.23	0.77
Parkinson’s Disease	0	0.07	0
Peripheral Vascular Disease	0.44	2.29	2.38
Polymyalgia	0	0.07	0
Psoriasis	1.8	5.16	4.54
Psychosis	0	5.21	0
Renal Stone	0.31	2.05	1.33
Rheumatoid Arthritis	0.17	0.98	0.76
Schizophrenia	0.07	7.76	0.97
Scleroderma	0.02	0.09	0
Sjogren’s Syndrome	0.03	0	0.15
Sleep Problem	0.62	5.68	3.9
Stroke	6.12	7.14	5.1
Systemic Lupus Erythematosus	0.06	0.06	0.06
Tuberculosis	0.31	1.03	0.67
Vision Problem	0.14	0.68	0.15

COPD- Chronic obstructive pulmonary disease; *Cluster 1- Relative healthy; Cluster 2 MSK-MH; Cluster 3 MSK; ^only for men; MSK- Musculoskeletal; MH- Mental health

Table 5.3-14. Posterior probabilities distribution of chronic conditions to five-class model in the non-OA population aged40-59 years

	Cluster 1 (36.0%)	Cluster 2 (3.18%)	Cluster 3 (31.6%)	Cluster 4 (22.2%)	Cluster 5 (6.98%)
Anaemia	1.25	12.79	6.35	4.50	13.89
Ankylosing Spondylitis	0.00	1.00	3.17	0.93	4.49
Arterial/Venous	0.06	0.27	0.06	1.27	1.81
Asthma	4.78	14.28	15.22	11.16	21.34
Back pain	17.84	58.74	**73.97**	55.39	**81.84**
Cancer (any)	2.16	9.24	6.40	8.72	10.34
Benign prostatic hypertrophy^	0.70	0.44	2.48	6.58	5.17
Cataract	0.58	1.01	0.52	1.23	1.85
Chronic Heart Disease	0.73	3.68	1.77	14.36	15.77
Chronic Kidney Disease	0.24	10.23	1.04	14.15	15.33
COPD	1.11	5.16	6.36	7.78	16.00
Dementia	0.03	0.27	0.26	0.55	2.03
Depression	6.18	36.75	41.03	19.94	**72.47**
Diabetes	1.26	21.49	3.25	26.20	24.71
Epilepsy	0.62	1.61	1.65	1.44	4.78
Fatigue	0.14	3.83	2.65	0.62	10.00
Fibromyalgia	0.00	1.43	1.63	0.00	9.35
Gall stones	1.04	8.55	6.34	5.26	17.43
Gastritis	0.58	4.78	7.76	6.17	24.32
Gastrointestinal Bleed	0.08	0.56	1.30	0.94	6.41
Gout	0.87	1.30	1.35	9.62	4.37
Hearing Problem	3.15	11.98	13.02	14.41	20.22
Heart Failure	0.00	0.50	0.00	1.64	1.68
High Cholesterol	3.30	19.09	10.33	35.63	33.86
HIV/AIDS	0.00	0.00	0.01	0.00	0.05
Hypertension	6.04	30.44	12.14	**62.84**	44.56
Hyperthyroidism	0.16	33.84	0.19	0.37	2.39
Hypothyroidism	1.05	**90.75**	4.33	4.26	14.48
Inflammatory Bowel Disease	0.85	5.84	9.10	4.73	19.10
Irritable Bowel Syndrome	3.06	7.85	10.70	1.90	22.45
Liver Disease	0.12	0.48	0.83	1.03	3.47
Migraine	1.86	14.11	14.72	4.98	23.92
Multiple Sclerosis	0.16	0.52	0.83	0.24	0.88
Osteoporosis	0.44	6.36	3.96	2.19	8.92
Parkinson's Disease	0.03	0.28	0.22	0.37	0.93
Peripheral Vascular Disease	0.24	2.22	2.07	3.73	7.41
Polymyalgia	0.08	0.71	0.33	0.72	1.08
Psoriasis	1.45	3.90	4.07	4.43	5.46
Psychosis	0.10	0.00	0.00	0.00	5.41
Renal Stone	0.29	0.82	1.21	2.96	3.21
Rheumatoid Arthritis	0.15	2.14	1.18	0.77	1.76
Schizophrenia	0.31	0.77	0.54	0.38	10.19
Scleroderma	0.01	0.20	0.08	0.00	0.19
Sjogren's Syndrome	0.00	0.46	0.24	0.04	0.58
Sleep Problem	0.27	3.73	3.69	2.33	11.83
Stroke	8.24	7.75	6.18	10.25	13.17
Systemic Lupus Erythematosus	0.01	0.18	0.09	0.01	0.45
Tuberculosis	0.18	0.65	0.75	0.56	1.14
Vision Problem	0.08	0.30	0.35	0.69	1.94

COPD- Chronic obstructive pulmonary disease; *Cluster 1- Relative healthy; Cluster 2- Metabolic; Cluster 3- MSK; Cluster 4- CV; Cluster 5 MSK-MH; ^only for men

CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Table 5.3-15. Posterior probabilities distribution of chronic conditions to five-class model in the non-OA population aged 60-79 years

	Cluster 1 (29.8%)	Cluster 2 (26.5%)	Cluster 3 (19.7%)	Cluster 4 (14.1%)	Cluster 5 (9.93%)
Anaemia	0.34	4.08	4.52	16.99	12.82
Ankylosing Spondylitis	0.00	2.46	0.00	2.99	3.96
Arterial/Venous	0.05	0.73	1.28	7.71	1.00
Asthma	1.77	11.68	9.10	11.53	16.07
Back pain	5.12	**65.93**	39.58	**73.60**	**74.42**
Benign prostatic hypertrophy^	0.70	6.58	0.44	2.48	5.17
Cancer (any)	1.81	14.67	12.97	21.23	14.03
Cataract	4.77	3.88	6.17	13.73	8.99
Chronic Heart Disease	0.83	8.01	16.00	47.88	14.11
Chronic Kidney Disease	0.00	2.55	30.89	48.75	28.45
COPD	0.93	10.78	6.95	22.51	13.09
Dementia	0.47	2.98	3.68	8.59	9.75
Depression	1.94	23.31	10.27	28.53	**53.02**
Diabetes	0.64	6.14	27.38	37.78	20.73
Epilepsy	0.23	1.62	1.09	2.32	2.52
Fatigue	0.02	1.57	0.43	2.87	6.34
Fibromyalgia	0.00	0.48	0.00	0.32	3.95
Gall stones	0.46	7.35	5.55	13.58	15.99
Gastritis	0.27	9.39	2.96	23.68	14.26
Gastrointestinal Bleed	0.00	1.47	0.38	6.91	3.42
Gout	0.19	2.83	9.22	17.60	1.90
Hearing Problem	2.24	23.89	17.81	36.87	25.63
Heart Failure	0.03	0.40	2.36	14.34	2.32
High Cholesterol	1.68	18.95	30.99	40.75	31.60
HIV/AIDS	0.00	0.00	0.01	0.00	0.00
Hypertension	5.06	29.19	**75.68**	**73.51**	54.34
Hyperthyroidism	0.09	0.01	1.82	1.38	11.56
Hypothyroidism	0.95	4.41	10.50	10.32	35.42
Inflammatory Bowel Disease	0.21	6.72	2.37	11.75	11.26
Irritable Bowel Syndrome	3.32	6.98	0.68	6.02	15.50
Liver Disease	0.07	0.64	0.49	1.41	1.41
Migraine	0.62	7.89	2.51	6.27	15.09
Multiple Sclerosis	0.10	0.57	0.14	0.18	0.92
Osteoporosis	0.78	9.52	4.60	11.40	23.09
Parkinson’s Disease	0.10	1.31	0.74	1.95	2.13
Peripheral Vascular Disease	0.26	2.96	3.35	17.22	4.66
Polymyalgia	0.18	1.57	1.89	3.41	5.29
Psoriasis	0.57	4.15	3.55	4.91	4.18
Psychosis	0.05	0.00	0.00	0.01	3.64
Renal Stone	0.11	1.95	1.59	4.03	0.77
Rheumatoid Arthritis	0.14	0.87	0.68	1.20	2.09
Schizophrenia	0.16	0.40	0.19	0.36	8.16
Scleroderma	0.00	0.07	0.02	0.10	0.34
Sjogren’s Syndrome	0.01	0.16	0.05	0.28	1.09
Sleep Problem	0.02	3.72	1.90	7.07	8.01
Stroke	14.39	8.67	14.14	28.97	17.28
Systemic Lupus Erythematosus	0.00	0.05	0.03	0.16	0.31
Tuberculosis	0.11	1.00	0.53	1.65	1.13
Vision Problem	0.05	0.68	0.99	4.03	2.30

COPD- Chronic obstructive pulmonary disease; *Cluster 1- Relative healthy; Cluster 2-MSK; Cluster 3-CV; Cluster 4-CV- MSK; Cluster 5-MSK-MH; ^only for men

CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Table 5.3-16. Posterior probabilities distribution of chronic conditions to five-class model in the non-OA population aged ≥ 80 years

	Cluster 1 (28%)	Cluster 2 (22.0%)	Cluster 3 (15.6%)	Cluster 4 (19.2%)	Cluster 5 (15.1%)
Anaemia	0.41	15.22	20.51	8.27	7.37
Ankylosing Spondylitis	0.00	3.42	3.32	0.00	0.88
Arterial/Venous	0.00	1.39	7.79	1.66	2.60
Asthma	0.87	12.33	9.99	8.26	8.58
Back pain	2.53	**72.76**	**72.32**	34.35	41.47
Benign prostatic hypertrophy^	6.58	0.70	2.48	0.44	5.17
Cancer (any)	0.86	14.74	20.87	15.63	19.26
Cataract	24.49	38.45	38.40	23.20	**43.88**
Chronic Heart Disease	0.27	14.86	51.76	20.98	16.33
Chronic Kidney Disease	0.11	32.93	**61.53**	**52.40**	8.27
COPD	0.69	10.29	21.03	5.49	13.57
Dementia	1.96	27.58	13.29	11.64	33.71
Depression	0.59	39.67	25.44	11.44	14.06
Diabetes	0.87	9.97	32.06	26.35	7.60
Epilepsy	0.13	2.37	1.82	1.15	1.79
Fatigue	0.04	3.83	2.74	1.13	0.88
Fibromyalgia	0.00	0.51	0.36	0.00	0.02
Gall stones	0.02	10.61	15.74	5.95	7.54
Gastritis	0.06	9.84	22.73	3.65	10.11
Gastrointestinal Bleed	0.00	4.09	7.69	1.07	3.53
Gout	0.06	1.75	17.82	7.88	3.88
Hearing Problem	1.83	44.84	48.16	25.87	**47.30**
Heart Failure	0.04	4.97	23.66	7.04	6.56
High Cholesterol	0.30	15.01	32.11	21.86	4.03
HIV/AIDS	0.00	0.00	0.03	0.00	0.00
Hypertension	5.65	58.96	**76.19**	**83.76**	35.56
Hyperthyroidism	0.06	6.01	3.70	2.52	0.00
Hypothyroidism	0.98	22.15	16.98	16.40	3.08
Inflammatory Bowel Disease	0.00	9.18	12.15	2.51	6.16
Irritable Bowel Syndrome	2.27	9.62	4.96	1.24	2.18
Liver Disease	0.06	0.36	0.91	0.11	0.42
Migraine	0.06	7.59	4.97	1.78	1.94
Multiple Sclerosis	0.00	0.00	0.16	0.05	0.08
Osteoporosis	0.54	37.95	15.27	8.24	9.67
Parkinson's Disease	0.15	2.82	1.87	0.90	1.98
Peripheral Vascular Disease	0.00	4.97	16.70	5.46	3.56
Polymyalgia	0.25	8.19	6.24	3.10	2.35
Psoriasis	0.00	4.05	4.42	2.29	2.43
Psychosis	0.00	2.65	0.00	0.00	0.00
Renal Stone	0.00	0.67	2.85	0.44	2.24
Rheumatoid Arthritis	0.03	1.26	1.07	0.41	0.73
Schizophrenia	0.08	5.33	0.33	0.19	0.29
Scleroderma	0.00	0.17	0.05	0.00	0.00
Sjogren's Syndrome	0.00	0.82	0.21	0.10	0.00
Sleep Problem	0.59	13.96	14.84	3.84	11.41
Stroke	21.24	26.03	36.05	21.09	18.09
Systemic Lupus Erythematosus	0.00	0.14	0.04	0.00	0.00
Tuberculosis	0.02	2.16	0.93	0.58	0.78
Vision Problem	0.06	8.17	7.99	2.60	5.35

COPD- Chronic obstructive pulmonary disease; *Cluster 1- Relative healthy; Cluster 2- MSK; Cluster 3- CV-MSK-Renal; Cluster 4- CV-Renal; cluster 5- Hearing-vision problem; ^only for men

CV- Cardiovascular; MSK- Musculoskeletal

The distribution of socio-demographic characteristics across the clusters in people without OA is provided in Table 5.3-17. The proportion of women was higher in all the clusters except in metabolic syndrome. Nearly 68% in the metabolic syndrome and 86% in the CVD and MSK clusters were aged 60 years or more, whereas the other three clusters had younger populations. Smoking prevalence was highest (52%) in the CV-MSK cluster (cluster n). The prevalence of obesity was 26% in CV cluster and 25% in the thyroid-MH cluster, followed by 23% in the CV-MSK cluster. The mean number of chronic conditions was highest for the CV-MSK (7.78; SD 1.84) and thyroid-MH(5.46; SD 2.01) clusters. (Table 5.3-17)

Table 5.3-17. Socio-demographic distribution across clusters in the non-OA population

Variables	Relative Healthy (n=89993)	Thyroid (n=7545)	CV-MSK (n=22390)	CV (n=66223)	MSK-MH (n=35656)
Gender					
Men	38510 (42.79)	1432 (18.98)	10363 (46.28)	33000 (49.83)	10590 (29.70)
Women	51483 (57.21)	6113 (81.02)	12027 (53.72)	33223 (50.17)	25066 (70.30)
Age					
<40 years	9224 (10.25)	295 (3.91)	47 (0.21)	880 (1.33)	2572 (7.21)
40-59 years	46747 (51.95)	3000 (39.76)	2633 (11.76)	20061 (30.29)	18838 (52.83)
60-79 years	30472 (33.86)	3597 (47.67)	14452 (64.55)	38548 (58.21)	13273 (37.23)
>=80 years	3550 (3.94)	653 (8.65)	5258 (23.48)	6734 (10.14)	973 (2.73)
Smoking					
Never smoked	52782 (58.66)	4173 (55.32)	10700 (47.79)	37516 (56.66)	18710 (52.48)
Current smoker	17743 (19.72)	1550 (20.55)	3372 (15.06)	9280 (14.01)	8292 (23.36)
Ex-smoker	19459 (21.62)	1821 (24.14)	8316 (37.14)	19420 (29.33)	8651 (24.26)
Alcohol use					
Never	14616 (16.24)	2053 (27.21)	5692 (25.43)	12423 (18.76)	6607 (18.54)
Ex-drinker	1444 (1.60)	344 (4.56)	1018 (4.55)	1599 (2.42)	944 (2.65)
Current (1-9)	35200 (39.12)	2494 (33.06)	7000 (31.27)	22544 (34.05)	13142 (36.87)
Current (>=10)	18808 (20.90)	819 (10.85)	3301 (14.75)	13746 (20.76)	6552 (18.38)
Current (Unknown)	19906 (22.12)	1835 (24.32)	5376 (24.01)	15895 (24.01)	8397 (23.56)
BMI					
Underweight	1869 (2.08)	248 (3.29)	720 (3.22)	1034 (1.56)	939 (2.63)
Normal	39992 (44.46)	2858 (37.89)	7862 (35.12)	21168 (31.97)	14741 (41.35)
Overweight	32538 (36.17)	2592 (34.36)	8529 (38.10)	26914 (40.65)	12440 (34.90)
Obese	15549 (17.29)	1845 (24.46)	5277 (23.57)	17096 (25.82)	7527 (21.12)
Mean age (SD)	56.30 (12.93)	62.27 (12.63)	72.23 (10.03)	65.42 (11.23)	57.18 (11.73)
Mean BMI (SD)	26.02 (4.72)	26.87 (5.54)	26.83 (5.09)	27.42 (5.01)	26.45 (5.18)
Multimorbidity (SD)	0.99 (0.94)	5.24 (1.94)	7.84 (1.87)	3.63 (1.44)	4.48 (1.51)

BMI- Body mass index, SD- Standard deviation
CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Table 5.3-18. Multinomial regression for associations between patient factors with cluster membership in the non-OA population

	Relative Healthy	Thyroid OR (95% CI)	CV-MSK OR (95% CI)	CV OR (95% CI)	MSK-MH OR (95% CI)
Gender					
Men	1	Reference	Reference	Reference	Reference
Women	1	3.03 (2.85-3.23)*	0.81 (0.78-0.83)*	0.75 (0.73-0.77)*	1.86 (1.81-1.91)*
Age					
<40 years	1	Reference	Reference	Reference	Reference
40-59 years		1.97 (1.74-2.23)*	10.89 (8.15-14.55)*	4.39 (4.09-4.71)*	1.42 (1.36-1.49)*
60-79 years		3.61 (3.19-4.08)*	92.06 (69.05-122.73)*	13.38 (12.45-14.37)*	1.56 (1.49-1.64)*
>=80 years		5.26 (4.55-6.08)*	319 (239.32-427.73)*	22.81 (21.03-24.75)*	0.97 (0.89-1.05)*
Smoking					
Never smoked	1	Reference	Reference	Reference	Reference
Current smoker	1	1.43 (1.34-1.52)*	1.36 (1.30-1.42)*	0.87 (0.84-0.89)*	1.43 (1.39-1.48)*
Ex-smoker	1	1.38 (1.30-1.46)*	1.88 (1.81-1.95)*	1.17 (1.14-1.20)*	1.34 (1.30-1.39)*
Alcohol use					
Never	1	Reference	Reference	Reference	Reference
Ex-drinker	1	1.97 (1.75-2.26)*	1.76 (1.60-1.93)*	1.27 (1.17-1.38)*	1.50 (1.38-1.64)*
Current (1-9)	1	0.60 (0.56-0.64)*	0.58 (0.56-0.61)*	0.81 (0.78-0.83)*	0.88 (0.85-0.91)*
Current (>=10)	1	0.50 (0.46-0.55)*	0.55 (0.52-0.58)*	0.93 (0.89-0.96)*	0.93 (0.89-0.97)*
Current (Unknown)	1	0.76 (0.71-0.82)*	0.77 (0.74-0.81)*	0.99 (0.95-1.02)	0.97 (0.93-1.01)
BMI					
Normal	1	Reference	Reference	Reference	Reference
Underweight	1	1.38 (1.19-1.59)*	1.40 (1.27-1.54)*	0.94 (0.87-1.02)*	1.23 (1.13-1.34)*
Overweight	1	1.27 (1.20-1.35)*	1.41 (1.35-1.46)*	1.52 (1.49-1.57)*	1.10 (1.07-1.14)*
Obese	1	1.85 (1.74-1.97)*	2.34 (2.24-2.45)*	2.39 (2.32-2.46)*	1.34 (1.29-1.39)*

BMI- Body mass index, SD- Standard deviation; *P value<0.05
CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

Table 5.3-18 describes the association of patient characteristics with cluster membership with healthy cluster as the reference group. Women had a higher risk than men of being in the thyroid-MH cluster (OR 3.01; 95% CI 2.83-3.20) and MSK-MH (OR 1.86; 95% CI 1.81-1.91) clusters when compared to the gender ratio in the healthy cluster. Being an ex-smoker increased the OR of being in all clusters when compared to the smoker to non-smoker ratio in the healthy cluster. Obesity and overweight were found to be significantly associated with all the clusters when compared to the obesity to normal ratio in the healthy group cluster.

5.3.4 Sensitivity analysis

The sensitivity analysis revealed equal number and type of clusters in both the training and testing dataset. Identified cluster are relative healthy, MSK, CV, CV-MSK, MSK-MH and metabolic (thyroid). (Appendix Tables 12 and 13, page 322-323)

5.4 Discussion

This study used a patient-centred approach for examining clustering of the chronic conditions rather than the disease centred approach. Firstly, the whole population was examined to understand the clustering of other conditions with OA. Later, clusters within OA and non-OA population were explored separately. The following are the key findings from the study. (1) There was a 54% probability that OA would be clustered with hypertension and back pain and a 35% probability that OA would be in the back pain and depression cluster. (2) Within both OA and non-OA people, clusters were centred around back pain, hypertension, depression, and thyroid problems. Within OA, the cluster size of CV-MSK, CV, and MSK-MH was more compared to that in non-OA. (3) In people with OA, age was strongly associated with MSK-MH (OR- 1.12; 95% CI 1.12-1.13), women had higher association with thyroid-MH cluster (OR- 5.45; 95% CI 5.05 -5.88), and obesity was associated with all the clusters with higher risk towards CV-MSK cluster (OR- 2.49; 95% CI 2.39-2.59), and CV cluster (OR- 2.19; 95% CI 2.14-2.27). (4) ex-smoker and ex-drinkers had strong association with all the clusters, while current alcohol users were protective towards each cluster. Similar associations were found in the non-OA group.

5.4.1 Clustering of OA in total study population

For the clustering analysis in the total study population a sample of 1.4 million people was used. To my knowledge, this is the first population based study to examine 50 chronic conditions, including OA. Two recently published studies on clustering of multimorbidity from a UK GP database (Zhu *et al.*, 2020) and UK Biobank (Zemedikun D.T. *et al.*, 2018) did not include OA as a chronic condition. It was unclear whether OA was grouped in the most commonly reported 'painful' clusters by Zhu et al. (Zhu *et al.*, 2020) but OA was the fourth most chronic condition (13.14%) reported in this study among adults aged 20 years or more. Thus, ignoring this most common arthritis may not provide appropriate clinically

relevant patterns. Of six clusters identified, the interest for this study was to find the affinity of OA towards certain clusters. OA contributed mostly to the CV-MSK cluster (5.6%) followed by the MSK-MH cluster (7%). Some previous studies have included arthritis and rheumatological and musculoskeletal diseases (RMD). Islam et al found that arthritis with depression created a distinct cluster among people aged 50 years or above (Islam *et al.*, 2014). A similar pattern was reported by Simoes et al who found the clustering of RMDs with depression among adults aged 18 years and above (Simoes D. *et al.*, 2018). As reported from the Spanish GP database study that examined the population aged 65 years or above, I did not find a distinct musculoskeletal (MSK) cluster (Guisado-Clavero *et al.*, 2018). LCA was used in the Newcastle 85+ study which identified clustering of OA and hypertension together (Collerton *et al.*, 2016). Previous clustering analyses have varied widely in statistical methods used, age group of study population, the total number and types of chronic conditions included and the reporting of the conditions. In this study most of these MSK conditions coexisted with hypertension or depression and OA was not an exception. But looking at the clustering in different age groups, a clear shift in pattern for OA is observed. In the younger age group (20-39 years) OA did not appear in any clusters and in the age group 40-59 years it had the highest probability to be present with depression. However, in later age it had equal probability to contribute to MSK, CV and depression clusters. The coexistence of OA with depression at each age group accords with findings from other studies. However, the distinct clustering with CV is of interest. The higher association with CV or metabolic syndrome and painful musculoskeletal conditions is well established (Hall *et al.*, 2016) and has been explained through linkage of obesity, ageing, physical inactivity and subclinical inflammation leading to physiological changes (Prior *et al.*, 2012; Rahman *et al.*, 2013; Kim *et al.*, 2016). Studies have shown the increased reporting of OA in MSK and vice versa (U. Kadam, Jordan, and Croft, 2004). Thus, perhaps people with both MSK and CVD are at increased risk of having OA within their clusters. A lot of research has been done to explain the nature of pain in OA in relation to 'central sensitization' (Dua A. B *et al.*, 2012; Lluch *et*

al., 2014) and the shared chronic pain mechanisms in MSK conditions may explains the possible clustering. Even though the association of depression and OA were not conclusive (Stubbs *et al.*, 2016b), a strong association was found both before and after the diagnosis of OA (Chapter 4, Figure 4.3-3, page 125). It may be that depression is not directly associated with structural OA, but it can mediate the pain experience through sleep disturbances and pain sensitisation. The coexistence of OA with CVD and depression increases the challenges for management of all these conditions. People with OA should be assessed for CVD and psychological problems, and vice versa. Also, the clustering with CVD warrants further research to explain the pathophysiological association with OA and metabolic syndrome.

5.4.2 Clustering in people with OA

The five identified clustering patterns in OA for all the population from this study were a relatively healthy group, a cluster led by hypertension only (CV) cluster, a hypertension and back pain together (CV-MSK) cluster, a thyroid only (metabolic) cluster, and a combined back pain and depression (MSK-MH) cluster. The largest cluster seen in people with OA after relative healthy group is led by CV followed by the MSK-MH cluster. This suggests that one third of the population with OA live with hypertension and nearly one fourth with depression and other MSK disorders. This together represents more than 50% of total OA population indicating the common occurrence of other conditions. In a comparatively smaller sample of 769 patients, a previous study reported hypertension to be clustered with OA compared to other rheumatic musculoskeletal diseases (RMDs) (Ziade *et al.*, 2020). A gender difference was found in the pattern of clusters within OA. Men had four clusters compared to women among whom the additional thyroid cluster was more prominent. Also, across age groups the clusters varied in numbers, type, and size. In the younger group aged 20-39 years, a clear group of MSK and MH was seen. Whereas, in the middle age group of 40-59 years thyroid appeared as one of the clusters.

Park et al reported that arthritis was clustered with thyroid and this cluster had a significant association with female (Park, Lee and Park, 2019). The change in clusters with age is understandable because of the strong associations of chronic conditions and multimorbidity with age. However, identification of depression and back pain clusters separately in younger age suggests a high burden of these two conditions. The CV clusters evolved with increasing age and were present in combination with back pain and other conditions in the later age group. This complexity of clusters gives idea about the possible patterns of co-existence of two or more chronic conditions. Their pathophysiology needs to be explored further. In older age (>80 years) chronic kidney disease and hearing problems became more prominent. Zhu et al also reported hearing problems as one of the leading conditions in one of their identified clusters (Zhu *et al.*, 2020). This could be because of the high prevalence of such conditions in the elderly population and survivor bias from CV.

The demographic characteristics distribution across different clusters in OA partially explains the observed pattern. For example, people in the CV-MSK cluster were on average nearly 14 years older than those in the relative healthy cluster and the MSK-MH cluster. This suggests the combination of CV and MSK conditions is common in the older population, whereas the MSK-MH cluster had the youngest age which reinforces the findings from the age specific clusters. Overall, the percentage of women was high in thyroid-MH cluster which validates the gender specific analysis that had an additional cluster for thyroid diseases. The class size of the CV-MSK cluster was higher for men compared to women, whereas the percentage of women was higher in the thyroid-MH, and MSK cluster. A similar finding was reported previously, where in women back pain was more frequent with depression while in men back pain was more frequent with CV (Scherer *et al.*, 2016).

It is interesting to know the burden of chronic conditions in each group. The mean number of chronic conditions was highest for the CV-MSK cluster and lowest for the CV group.

The burden of multimorbidity in CV-MSK cluster can be either related to age (as the cluster had the oldest age population) or the coexistence of these two conditions might have led to the appearance of other conditions. The second highest burden of multimorbidity was seen for the thyroid-MH cluster. Even though the cluster is led by thyroid, the contribution of hypertension and other CV is still high compared to other clusters. The association of all risk factors studied such as age, smoking, obesity, and alcohol use with cluster membership was the highest for the CV-MSK cluster. All these risk factors are well known for CV. This suggests the people of this clusters are with unhealthy lifestyle behaviours can be identified as the 'high risk' group from early dates. Women had a higher likelihood of being in thyroid-MH clusters. This supports the results from multimorbidity in a chronic pain study (Scherer *et al.*, 2016; Park, Lee, and Park, 2019). Obesity was found to be strongly associated with clusters led by hypertension. As a well-established risk factor for both CVD and MSK, obesity can be considered to have a significant role in developing such a cluster.

5.4.3 Clustering in people without OA

A similar clustering pattern in the non-OA group across both gender and age was that of OA. The purpose of doing cluster analysis in the non-OA group was to identify any pattern that was different from OA. Even though the pattern in non-OA was like that in OA, two major notable differences are the class sizes and predominance of hypertension in most of the clusters. The reason of not finding very drastic differences in clustering pattern was due to the distribution of the chronic conditions in both groups. In OA and non-OA, the sequence of the leading conditions was the same except for few variations. In OA, the frequency of back pain was the highest whereas, in non-OA hypertension was top of the list. In non-OA there were nearly 7% of more people in the relatively healthy cluster compared to the OA group. This indicates the burden of chronic conditions was less in the non-OA group at all ages and in men and women, which underpins the findings from the

previous chapter. Also, the class size the cluster led by depression and back pain was higher in the OA group than that of non-OA, which indicates the nexus of chronic pain and depression in OA. A similar pattern was reported from the UK CPRD database (Cassell A. *et al.*, 2018). In the elderly population, the non-OA group had a separate cluster of hearing and vision problems and a low contribution from 'back pain' to each cluster.

In terms of membership of single conditions within clusters in both the OA and non-OA groups, back pain, depression, and hypertension played a central role in identification of clusters. Similar patterns have been reported in previous studies (Prados-Torres *et al.*, 2014; Deruaz-Luyet A. *et al.*, 2017; Zemedikun D.T. *et al.*, 2018; Zhu *et al.*, 2020). The systematic review of 14 studies on multimorbidity clusters found three prominent clusters of CVD, MH and MSK, similar to this study (Prados-Torres *et al.*, 2014). In the OA population the class size of chronic pain either with hypertension or depression was high compared to non-OA. Both chronic pain and depression are recognised as having a major impact on health service use (Payne *et al.*, 2013). Findings from this study emphasise the importance of ensuring psychological care and management of pain in younger life (Das, Naylor and Majeed, 2016). Chronic pain presence in most of the clusters and having highest multimorbidity burden with CV, needs further investigation. So, along with CV and depression, chronic pain should be considered as one of the major chronic conditions for multimorbidity clusters.

5.4.4 Strength and limitations

The clustering analysis in this study was done among 1.4 million patients and examined 50 chronic conditions including OA. This is the first study in this area to use such a large sample size, such an extensive disease list, and to include OA as a condition. The use of a broad list of conditions, large database and real presentation of the health system strengthens the study. The gender and age strata specific clusters provide further insight into the changing nature of the pattern of clusters during the life course. Also, inclusion of

the adult population aged 20 years or more, compared to other studies which largely focus on the elderly population, which makes results from this study more representative to the total population. Although various clustering methods have been devised, I compared many of these including the machine learning approach and found LCA provided the best fit model. Another strength of this study is the examination of the association of baseline risk factors with identified clusters which provided clinically important information. Even though no wide diverse clusters were found, as reported in some other studies, the distribution of age and multimorbidity across the clusters provides internal validity.

There are some limitations to this study, most of which are inherent in electronic health record research. There is the possibility of misdiagnosis, mis-recording and misclassification bias for OA and other conditions studied. However, all possible steps were taken to use the validated codes and those not validated were screened thoroughly by this study team. As data were from diagnoses recorded in a GP database the findings represent the burden on the health care system rather than the burden within the whole community. The chance of under/over reporting of conditions are inevitable in such databases. Clustering of binary data (disease yes/no) is an evolving methodological field of research.

5.4.5 Conclusion

Identified multimorbidity clusters provide information about the burden of the conditions in people with OA and in the non-OA control group. Firstly, the metabolic pathology of OA should be examined to understand its strong affinity toward CV. Within OA, the distinct groups represent the clear burden of conditions, especially CV, MH, and MSK. The pattern of MSK and MH clusters in the younger age group justifies future study of longitudinal changes in clusters. Also, outcomes associated with cluster membership can be studied to emphasize the severity of the clusters in each group. Improving care in

multimorbidity is a difficult challenge which has not been fully successful because of the complexity of disease patterns (Smith *et al.*, 2012, 2016). One reason for the failure of previous interventions is that multimorbidity is heterogeneous, with very different diseases, needs and outcomes in different groups of patients (Salisbury *et al.*, 2018).

Summary of Chapter 5

In the previous chapter (Chapter 5) we explored the pattern of existence of the chronic conditions in people with OA and matched controls.

Key findings from the chapter are:

- OA is more likely to be present in people with hypertension and back pain (54%), followed by people with back pain and depression (35%).
- Hypertension, back pain, depression, and thyroid diseases were the leading conditions cluster wise. However, in people with OA a greater number of people had co-existence of cardiovascular and musculoskeletal, cardiovascular only, and musculoskeletal and depression together.
- Age was significantly associated with cardiovascular and musculoskeletal cluster, while women had more likelihood of belonging to thyroid cluster.
- Obesity, ex-smoker, and ex-drinkers had strong association with all the clusters, while current alcohol users were protective towards each cluster.

As the clusters centred around cardiovascular, musculoskeletal, depression and thyroid diseases are established through this chapter, the next question it brings to understand how people move from one cluster to another over time. Which clusters identified at the time of diagnosis of OA remain stable (does not change in size) and which does? Chapter 6 explores the transition of people across the cluster with time.

6 Chapter 6

Transition of comorbidity clusters

6.1 Introduction

In Chapter 5, the clustering pattern of comorbidities at the index date is reported. In both OA and non-OA groups, 5 clusters were identified. As the study population is dynamic and has been followed up to a maximum of 20 years, it is possible to examine whether identified clusters at the index date remain the same or change throughout follow-up. There are high possibilities of changes in the clusters over time because of ageing and the diagnosis of new conditions. Moreover, the dynamic properties of change in clusters are time-dependent and latent in nature and may be influenced by other factors not captured or recorded in the database. Thus, the basic assumption is that the clusters can be influenced not only by known and observable factors, but also by unknown latent conditions which change with time.

Also, the clusters identified after the diagnosis of OA are related, so there is possible transition of people from one cluster to other. Understanding the proportion and probabilities of people transition from one cluster to another would be beneficial for clinicians. This would also help to understand the evolving path of each clusters. To my knowledge, no studies have explored the dynamic changes in clusters over time not even the repeated cross-sectional nature.

In the previous chapter (chapter 5) it was seen, the nature of the clustering of comorbidities varied across different age group. In the younger population, psychological and musculoskeletal conditions were common while with increase in age, cardiovascular conditions were becoming prominent. Even though the study

population in each of the age group studied were not constant, the change in clusters suggest the dynamic nature of the comorbidities coexistence worth investigating. Also, it was not clear how these clusters change in an individual after the diagnosis of OA. This is only possible by following a group of people after the diagnosis of OA.

Therefore, latent transition analysis (LTA) and repeated measures of LCA (RMLCA) were used to estimate the probabilities of transition of clusters over time. Using this methodology, the objectives of this study were to determine: (1) the probability that an individual will be in different latent clusters over time; (2) whether there are changes between latent clusters across time; (3) the probability of moving from one cluster at time t to another in time t+1; and (4) whether any change in latent clusters with time are influenced by the baseline characteristics?

6.2 Methods

6.2.1 Study subjects

Data identified for LCA analysis (Chapter 4, section 4.2.2) in both OA and non-OA individuals were used for this analysis. Each person in the database was followed until 31st December 2017 or until they died or left the database. All available data at each time point was used considering the missing data at each follow-up to be missing at random (MAR), an assumption required for the analysis. The missingness of data could be due to death, transfer out from the database or no observation being recorded. Details of the individual at each time point are provided in Figure 6.3-1. Very few individuals were available at the 20 years follow-up after index date. Therefore, the transition model was analysed up to 15 years after the index date to avoid biases due to small sample sizes at 20 years.

6.2.2 Measurements and covariates

Forty-nine previously used chronic conditions were used for LTA. Covariates such as age, gender, BMI, smoking, alcohol, and time were also used in the model. The model was fitted to a patient database of OA and non-OA, separately with 49 chronic conditions coded as no '1' or yes '2'. The status of chronic conditions was recorded at the index date and at each five years interval afterwards (i.e., at 5, 10 and 15 years). The reasons for selecting a five year gap were to reduce the number of transition points to avoid complexity, and because the chronical nature of the comorbidities. Every 5 years may be more adequate than every year to catch the significant changes.

6.2.3 Statistical analyses

LTA was used for the main analysis. LTA was primarily developed for the analysis of longitudinal data and to deal with categorical response variables. Generally, it can be considered as an extension of LCA allowing each subject to move between the clusters over time. These models use time-specific discrete latent variables (Ryoo *et al.*, 2018). Ryoo et al proposed six steps for doing LTA variables (Ryoo *et al.*, 2018), specifically:

1) Explore the cross-sectional data using LCA
2) Test longitudinal measurement invariance using LTA
3) Define latent statuses
4) Test latent statuses and transition probabilities
5) Include covariates
6) Include distal outcomes

Of these six steps, step 6 is not applicable to the current study. There is no outcome of interest for this study except to understand the nature of the transition.

LMest package was used from R to do LTA. Advantages of this package are: (1) it can deal with univariate and multivariate categorical outcomes; (2) it allows for missingness, dropout, under the assumption of missing-at-random (MAR); (3) it is computationally easy; and (4) random effects can be added for Latent Markov models. LMest uses a log-

likelihood maximization procedure as does LCA for parameter estimation. So, both BIC and AIC criteria were used to choose the best model and the change in log-likelihood ratio and percentage of observation in each class. LTA provides three parameters - conditional response probabilities, initial probabilities, and transition probabilities. Conditional probabilities explain the characteristics of latent classes, i.e., the probability of one condition being present given the condition belongs to the class. Initial probabilities describe the latent class structure and sizes at baseline. Lastly, transitional probabilities, which are the interest of this analysis, give information on the dynamics of the latent process at each time point.

Step-1: The LTA approach used here is quite like models used elsewhere. Before running the model, the assumption of non-varying number of clusters was checked using RMLCA at each time point and the number of classes identified were matched with the statistical parameters obtained from LTA.

Step-2: As same number of clusters and similar clusters at each time were obtained, this meets the assumption of longitudinal measurement invariance. So, no further statistical test was done for this purpose.

Step-3: Latent classes were defined based on the conditional probabilities of chronic conditions in each class. Both initial probabilities and latent classes at each time point were estimated.

Step-4: Latent status and transitional probabilities were estimated using the LTA model for each time point.

Step-5: Along with time, age at each follow-up time, gender, smoking, alcohol, and BMI were included as covariates in the model while determining LTA. Association of covariate on transition probabilities could not be estimated due to different sample sizes at each time point.

6.2.4 Sensitivity analyses

Sensitivity analyses was done among OA and non-OA group without any comorbidity at the index date using RMLCA, since there were high possibilities of a change in number of comorbidity clusters over time as they had none at the index date.

Repeated measures latent class analysis (RMLCA)

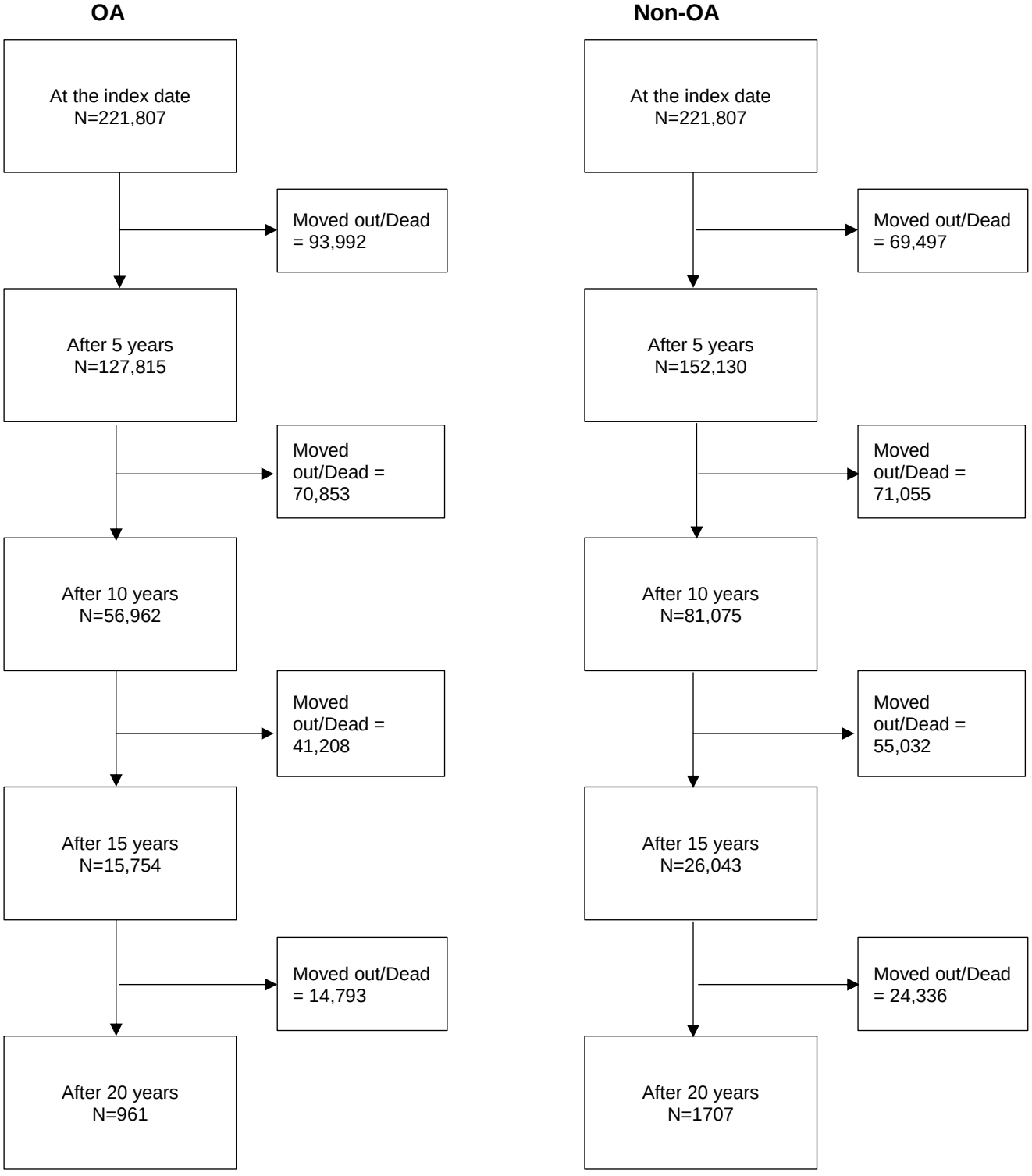
Another method used to explore the transition over time is RMLCA. It is an ideal method when the number of latent classes or clusters differ over time (i.e., not constant). It estimates LCA at each observation period cross-sectionally not considering the transition part. This means multiple LCA models are run for each of the time points and later these are merged with the patient unique identification number in the database to compute the change in clusters. Technically, this describes the changes in clusters more qualitatively than quantitatively. Individuals with similar clustering patterns are expected to be members of the same latent class. But at the same time, it captures and identifies the estimated size of subpopulations in each latent class over time.

While estimating LCA at each time point, the same population was used as described in Table 1. A group for the missing data was created to understand the pattern of clusters. The change of individuals within each cluster at each time point was estimated by calculating the pattern of the trajectory. The plot was created using ‘ggparallel’ in R. The parallel diagram depicts the clusters and the size of the link from one time to another and indicates the volume of individuals who have moved from one cluster to another. Also, the pattern of 20 leading trajectories have been reported having a proportion of >= 1%. Model selection criteria for RMLCA were similar that reported in the previous chapters on LCA.

6.3 Results

There was a high attrition rate in the OA group compared to the non-OA group at each time point. After five years from the index date, nearly 43% and 32% were missing in the OA and non-OA groups respectively, which further increased at each follow-up time.

Figure 6.3-1. Flow chart showing numbers at each time point in both OA and non-OA populations



The baseline sample characteristics have been described previously in Chapter 4, Table 4.3-1. A summary of the key population characteristics at each follow-up time is given in Table 6.3-1. In both OA and non-OA groups the percentage of women at baseline was 57.7% which declined faster in the OA compared to non-OA group over time. Mean age of individuals at the index date was nearly 61 years in both groups. However, at each follow-up time the mean age of the non-OA was higher compared to the OA group. Mean BMI was higher in the OA group at the index date and remained nearly consistent throughout follow-up in both groups. However, the mean number of comorbidities in the OA group at index date was 2.45 compared to 1.84 in the non-OA group, and it continued to increase to 4.02 and 3.41 after 20 years in the OA and non-OA groups, respectively. (Table 6.3-1)

Table 6.3-1. Study population characteristics at each follow up time

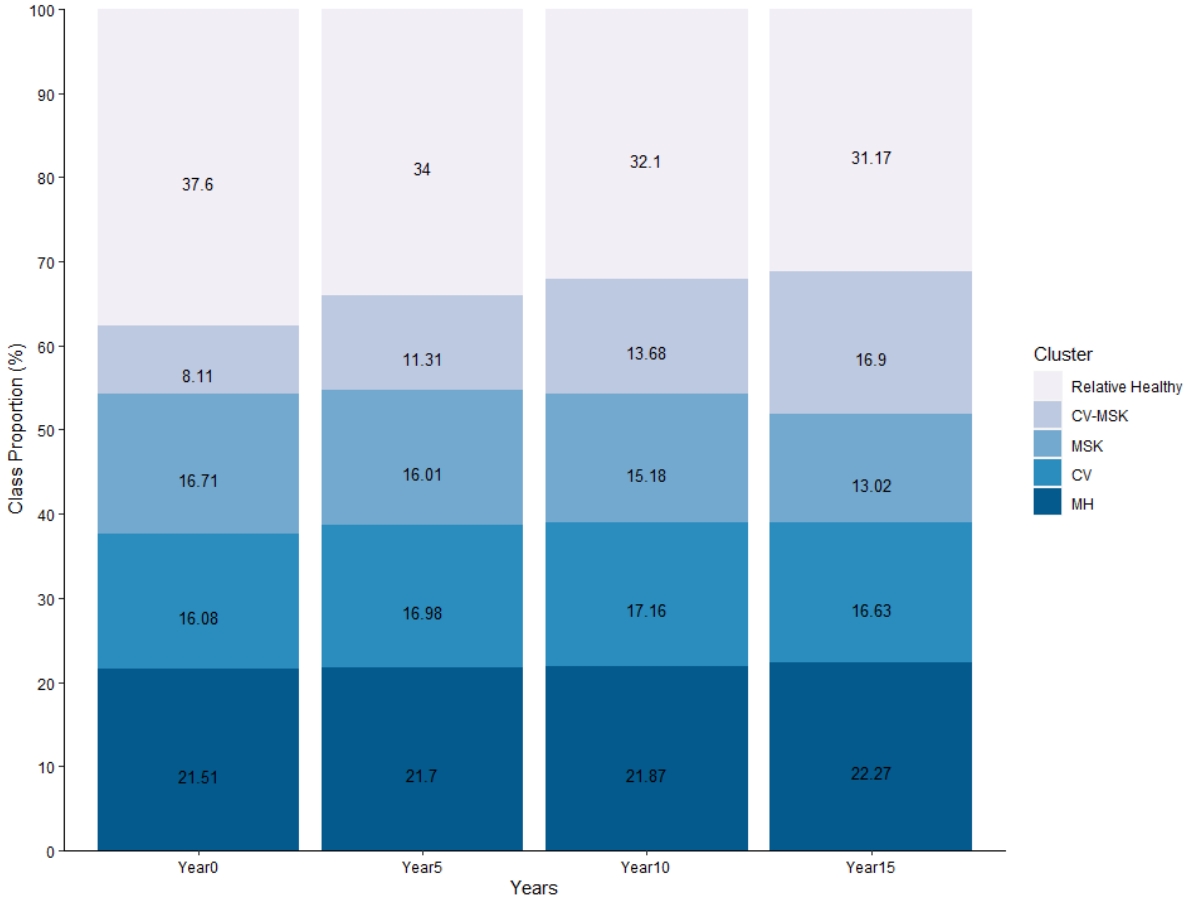
Variable	At index date	After 5 years	After 10 years	After 15 years	After 20 years^
OA	**N= 221807**	**N=127815**	**N=56962**	**N=15754**	**N=961**
Gender (% Female)	127906 (57.67)	73530 (57.53)	32421 (56.92)	8712 (55.30)	516 (53.69)
Age (Years) (Mean, SD)	61.14 (13.03)	64.71(12.68)	67.48(12.12)	69.66(11.66)	71.54(10.88)
Body mass index* (Mean, SD)	28.28 (5.63)	28.39 (5.63)	28.54 (5.63)	28.69 (5.65)	28.85 (5.54)
Number of comorbidities (Mean, SD)	2.45 (2.15)	3.07 (2.43)	3.60 (2.68)	4.01 (2.86)	4.02 (2.96)
Non-OA	**N= 221807**	**N=152130**	**N=81075**	**N=26043**	**N=1707**
Gender (% Female)	127912 (57.67)	87262 (57.36)	46063 (56.82)	14492 (55.65)	939 (55.01)
Age (Years) (Mean, SD)	60.98 (13.15)	65.25 (13.12)	68.75 (12.86)	71.26 (12.49)	73.36 (11.37)
Body mass index* (Mean, SD)	26.62 (4.98)	26.58 (4.94)	26.59 (4.92)	26.64 (4.87)	26.69 (4.90)
Number of comorbidities (Mean, SD)	1.84 (1.88)	2.36 (2.15)	2.83 (2.38)	3.21 (2.55)	3.41 (2.67)

SD- Standard deviation; ^not included in the analysis later; *baseline values

6.3.1 LTA defined latent classes (initial and conditional probabilities) in OA

Appendix Table 25 (page 335) shows fit indices for different number of classes based on the LTA models. Based on the model statistics selected the five-class model to be the best for both the OA and non-OA groups, which is similar to clusters that were found in LCA. (Appendix Table 25, page 335) Five identified clusters were named as per the posterior probability distribution of the chronic conditions. Initial probabilities of clusters at baseline identified five clusters. Cluster 1 was the relatively healthy (37.6%) group with lower probability of each conditions. Cluster 2 was dominated by back pain and hypertension (8.11%) (known as CV-MSK), cluster 3 had higher contribution from back pain (16.71%) (MSK cluster), cluster 4 was predominantly hypertension (16.08%) (CV cluster), and cluster 5 was led by depression (21.51%) (MH cluster). Details of the contribution from each condition are given in Appendix Table 26 (page 336). The percentage in the bracket against each cluster represents the initial probabilities at the index date. Figure 6.3-2 depicts the estimated probabilities of clusters at each time point in OA. Amongst the five clusters, CV-MSK cluster increased in size at follow-up time from 8.11% at index date to 16.9% after 15 years. A marginal increase in cluster size was seen for CV and MH clusters, whereas in both the relatively healthy cluster and MSK cluster the frequency reduced over time.

Figure 6.3-2. Different latent classes at each time point in the OA population.



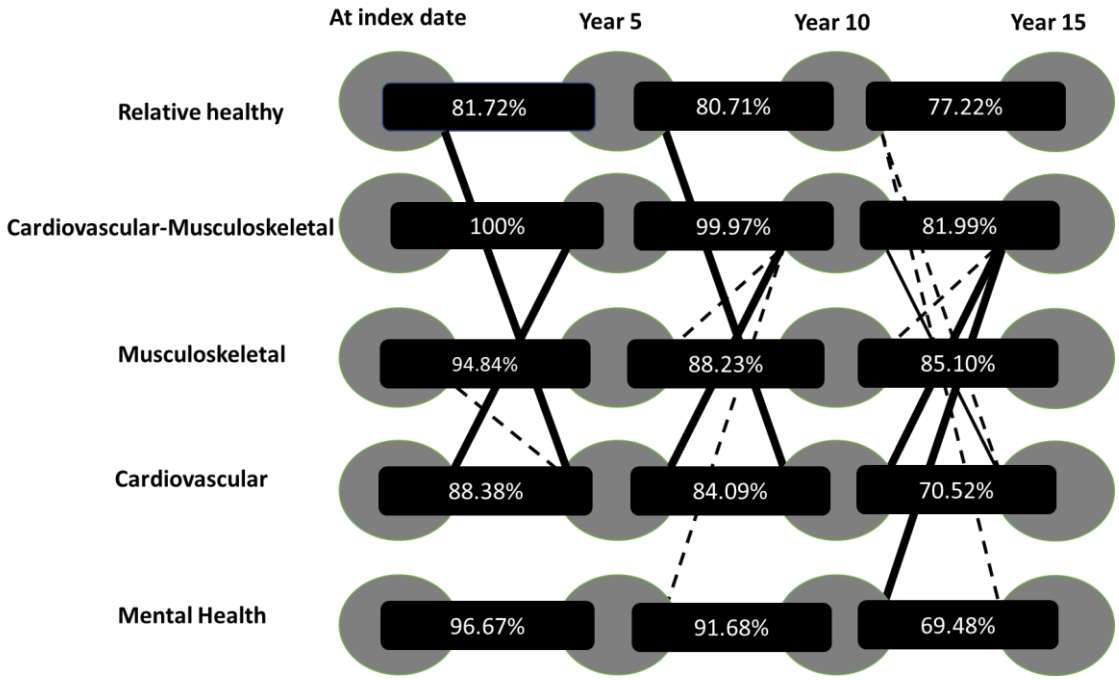
CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

6.3.1.1 Transition between clusters and trajectory paths in OA (transitional probabilities)

The LTA fitted model was used to classify individuals at each time-point according to their maximum posterior estimated class probability. Figure 6.3-3 depicts the path of transition of each cluster with time in the OA group. Detailed transition is provided in Appendix Table 28 (page 338).

The most common paths were those with membership of the same cluster over adjacent years. For example, all the people in CV-MSK cluster at index date stayed in the same cluster at year 5.

Figure 6.3-3. Estimated frequency of cluster transitions in the OA population.



The circles represent clusters (not proportionate to size). Thickness of each black line/bar is proportional to the estimated transition frequency. Dashed line represents 5-10% frequencies. For clear presentation, transition frequencies <5% are not shown here. For detail, please refer to Appendix Table 28 (page 338).

From the index date to year 5 all the transitions into subsequent clusters were stable, while nearly 15% in each moved from CV to CV-MSK and from relative healthy to cluster CV. During the year 5 to year 10 period, besides moving to the same clusters nearly 20% individuals from CV, MH and relative healthy clusters moved to CV-MSK and 15% moved from relative healthy to CV. Transition during years 15 to 20 to the same clusters was less in MH and CV clusters compared to the rest. Nearly 40% of individuals moved from CV, relative healthy clusters, and MH to CV-MSK cluster. After year five, the most common path was towards CV-MSK cluster from each cluster and at each time point nearly 12% of individuals moved from relatively healthy to CV cluster.

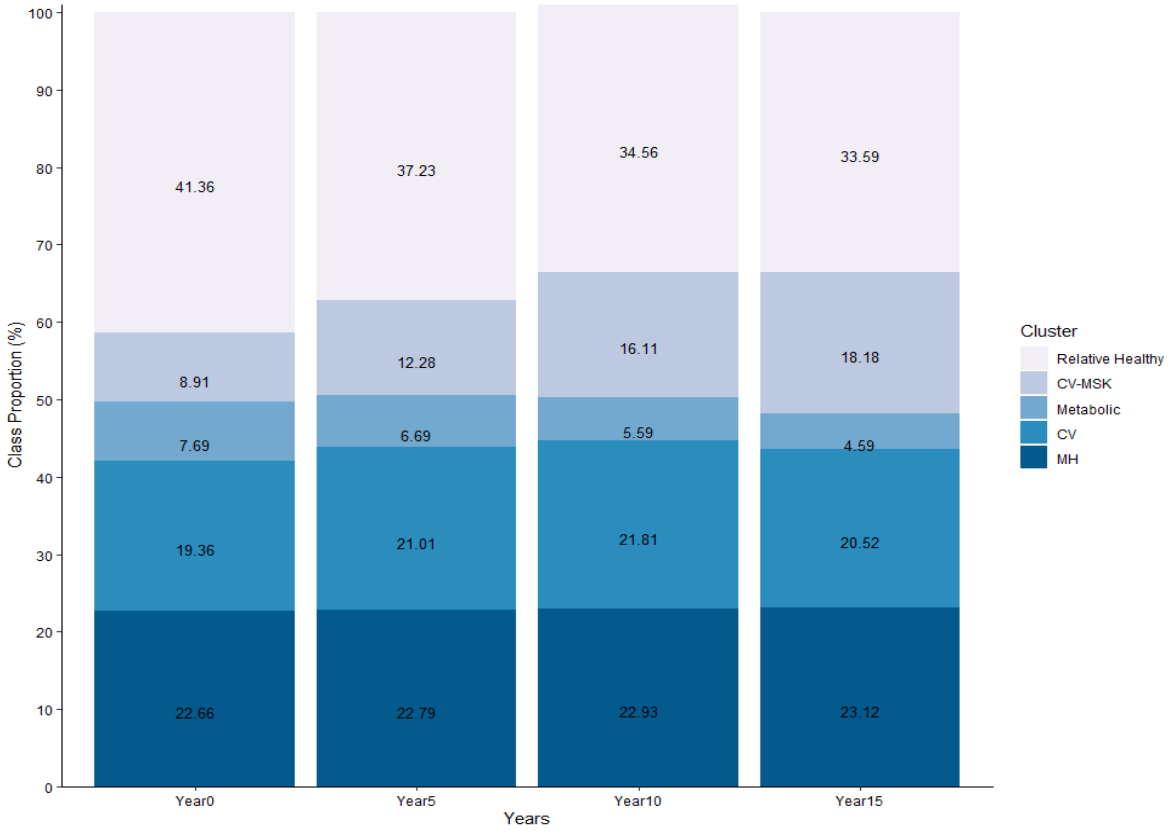
Of the total transition pattern, 30% happened in the relative healthy cluster path, 20% occurred in MH cluster path and 12% were in CV clusters path. Nearly 8% of total transition had from MSK cluster to CV-MSK cluster at year 10 or 15. Transitions more than 1% are given in Appendix Table 30 (page 334).

6.3.2 Different latent classes at each time point in the non-OA population

Five clusters were identified in the non-OA group at the index date. These were cluster 1 (41.36%) which was the relative healthy group, cluster 2 (8.91%) which was dominated by back pain and hypertension (CV-MSK), cluster 3 (7.69%) which was led by thyroid disorder (hence known as metabolic cluster), cluster 4 (19.36%) which was largely led by hypertension (CV cluster), and cluster 5 (22.67%) which was led by depression (MH cluster).

Figure 6.3-4 presents the size of different clusters identified at each time in the non-OA group. Metabolic cluster size reduced from 7.7% to 4.6% after 15 years and relative healthy cluster size decreased by 7% after 15 years. The size of MH cluster remained mostly constant at each time, whereas the size of CV-MSK and CV increased from the index date. Details of the distribution of the conditions in each class is given in Appendix Table 27 (page 337).

Figure 6.3-4 Different latent classes at each time point in the non-OA population.

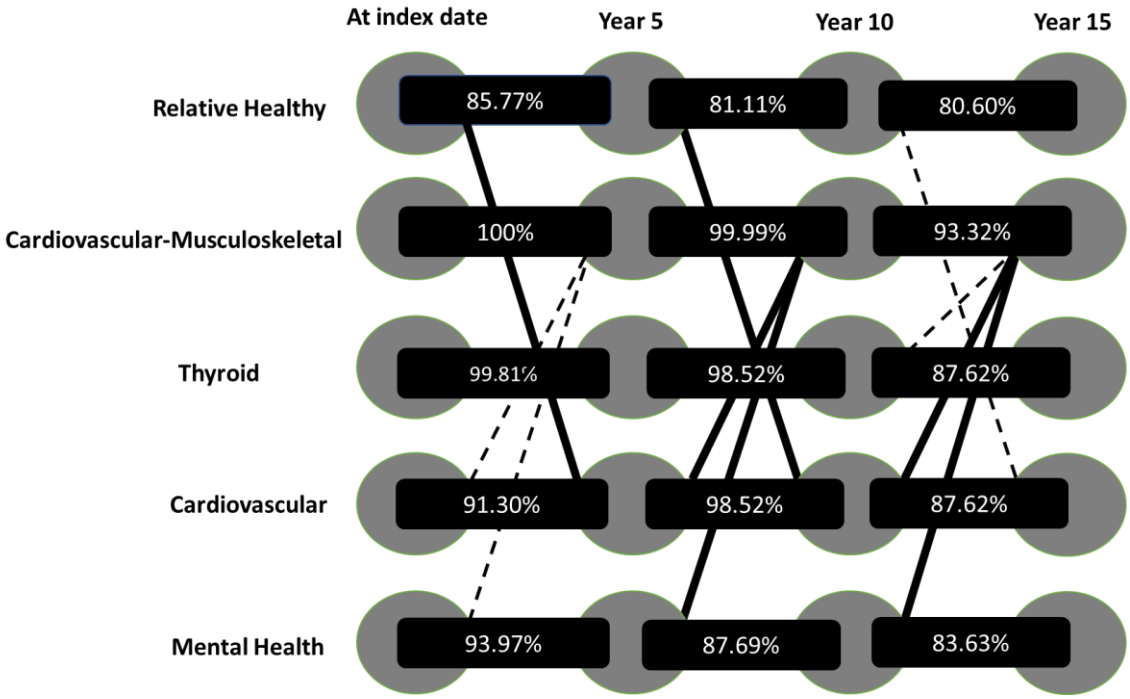


CV- Cardiovascular; MSK- Musculoskeletal; MH- Mental health

6.3.2.1 Transition across clusters in the non-OA population

In the non-OA group, less frequent transition was seen compared to the OA group. Similar trajectory paths seen in OA group were seen but became more distinct after year 10. From the index date to year 5, 15% of individuals moved from relative healthy to CV cluster and 10% combined moved from CV and MH clusters to cluster CV-MSK. CV-MSK cluster was seen to be more stable during the transition period. During years 5 to 10 and 10 to 15, nearly 30% of individuals moved from CV and MH clusters to CV-MSK cluster at each phase. During the period year 10 to 15 a small proportion of <5% were seen to move from relative healthy to CV cluster. (Figure 6.3-5) Details of the transition are given in Appendix Table 29 (page 339).

Figure 6.3-5 Estimated frequency of cluster transitions in the non-OA population

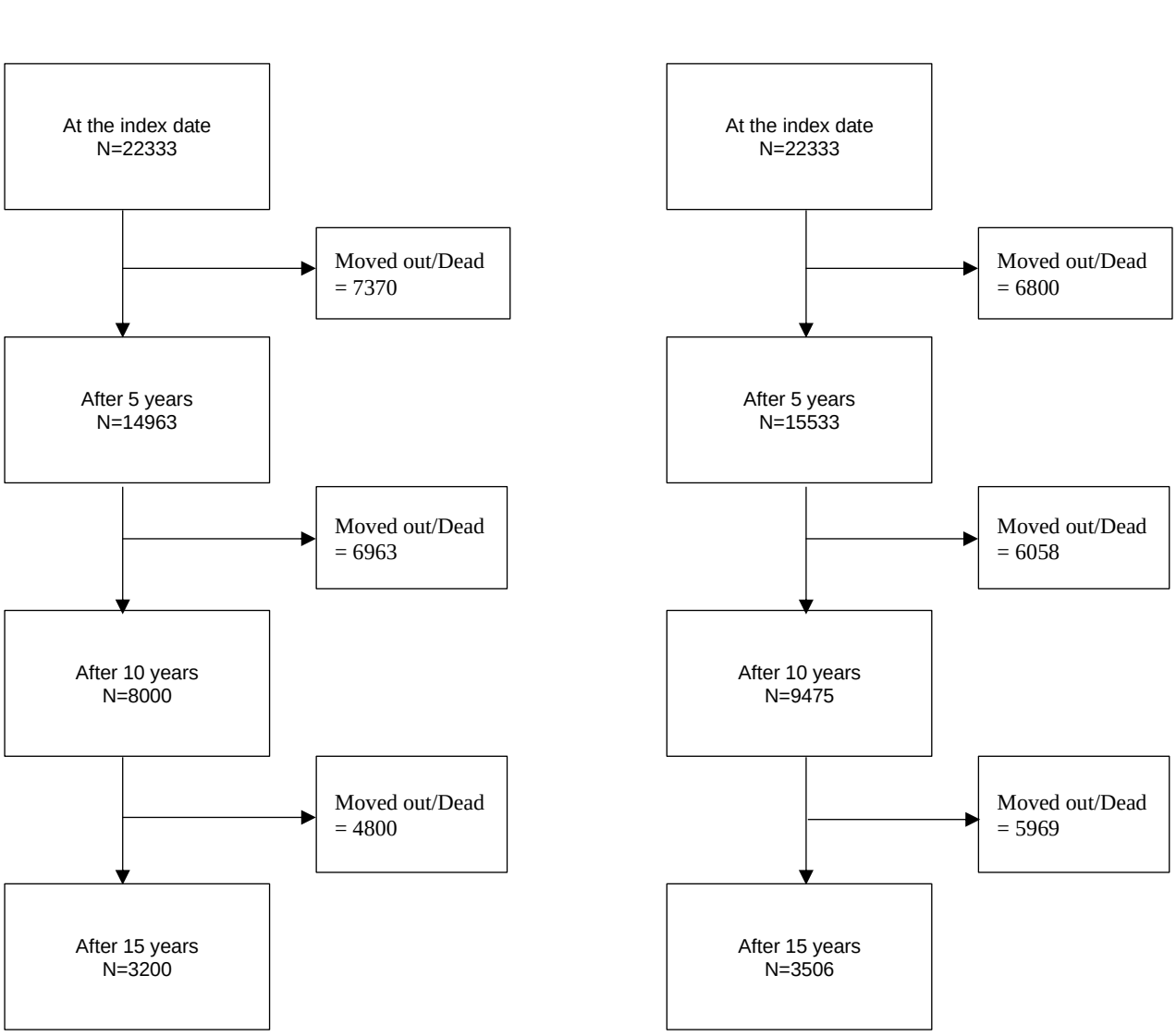


The circles represent clusters (not proportionate to size). Thickness of each black line/bar is proportional to the estimated transition frequency. Dashed line represents 5-10% frequencies. For clear presentation, transition frequencies <5% are not shown here. For detail, please refer to Appendix Table 29 (page 339).

6.3.3 Sensitivity analysis

Transition analysis was done in the second patient cohort i.e., OA cases and controls without any comorbidities at the index date and matched for age, sex, and practice.

Figure 6.3-6 Flow chart showing numbers at each time point (sensitivity analysis) in both groups



An equal number (n=22,333) of cases and controls were followed until 20 years. Details of the number of individuals at each time point are given in Figure 6.3-6. After 15 years, only 38% of OA and 42% of non-OA patients were left in the group. The attrition rate was higher in the OA group compared to the non-OA group.

At the beginning of the follow-up 52.4% were female in each group which reduced to 50% in both OA and non-OA group after 15 years from index date. Similarly, mean age at baseline increased from 57 years to 66 years in OA and to 67 years in the non-OA group. At index date, mean BMI was higher in the OA group which continued to increase at each follow-up time, while in the non-OA group it remained nearly constant. The mean number of comorbidities increased subsequently in both the groups. (Table 6.3-2)

Table 6.3-2 Descriptive statistics for baseline characteristics of study populations (sensitivity analysis)

Variable	At index date	After 5 years	After 10 years	After 15 years
OA	**N=22333**	**N=14963**	**N=8000**	**N=3200**
Gender (% Female)	11711 (52.43)	8084 (51.32)	4216 (49.99)	1644 (48.67)
Age (Years) (Mean, SD)	56.71 (13.55)	60.26 (13.38)	63.19 (12.81)	65.97 (12.30)
Body mass index (Mean, SD)*	28.44 (5.68)	28.55 (5.63)	28.70 (5.63)	28.85 (5.65)
Number of comorbidities (Mean, SD)	0 (0)	0.43 (0.80)	0.84 (1.30)	1.27 (1.72)
Non-OA	**N=22333**	**N=15533**	**N=9475**	**N=3506**
Gender (% Female)	11711 (52.43)	8425 (51.53)	4997 (50.03)	1823 (49.25)
Age (Years) (Mean, SD)	56.53 (13.58)	60.55 (13.46)	64.01 (13.37)	66.54 (13.12)
Body mass index (Mean, SD)*	26.80 (5.05)	26.73 (5.00)	26.74 (4.99)	26.84 (4.98)
Number of comorbidities (Mean, SD)	0(0)	0.32 (0.66)	0.64 (1.07)	1.05 (1.52)

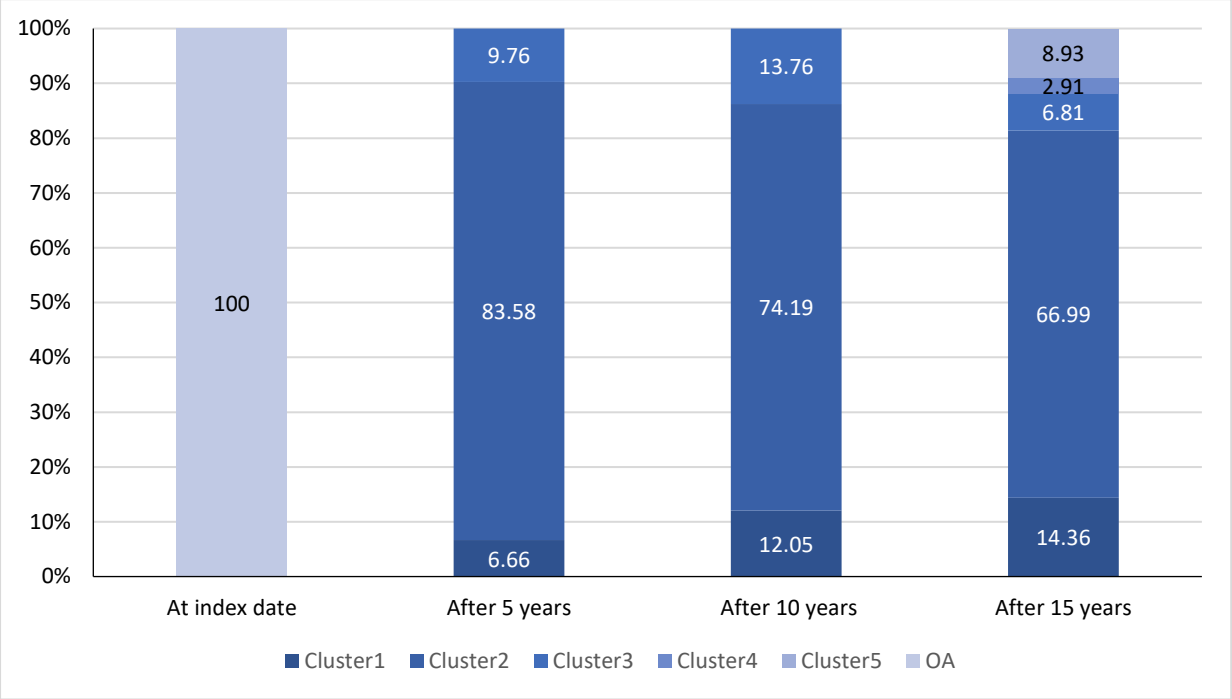
SD- Standard deviation; *value at baseline

6.3.3.1 Clusters in OA at each time point

Clusters identified among OA at each time point through LCA are given in Figure 6.3-7. Three clusters at 5 and 10 years and five clusters at 15 years were selected based in the statistical parameters. (Appendix Table 31, page 341). Three identified clusters at 5 and 10 years were CV, relative healthy and MSK. The cluster size of CV cluster increased from 6.66% at the index date to 14.36% at 15 years, while the cluster size of the relatively healthy

group reduced at each follow-up time. Details of the posterior class distribution for each time point are given in Appendix Table 33-35 (page 343-345).

Figure 6.3-7. Different clusters in OA at each time point (Sensitivity analysis)



Cluster 1- CV; Cluster 2- Relatively healthy; Cluster 3- MSK; Cluster 4- CV-MSK and Cluster 5 – MSK-MH
CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

Figure 6.3-8 represents the top 20 transitions pattern across the clusters in OA. Leading trajectory paths identified in the group are given in Table 6.3-5. Nearly 47% of the total transition occurred within the healthy group in subsequent years. Followed by moving to CV cluster 3.35% and 2.80% and MSK to MSK to MSK-MH (2.30%).

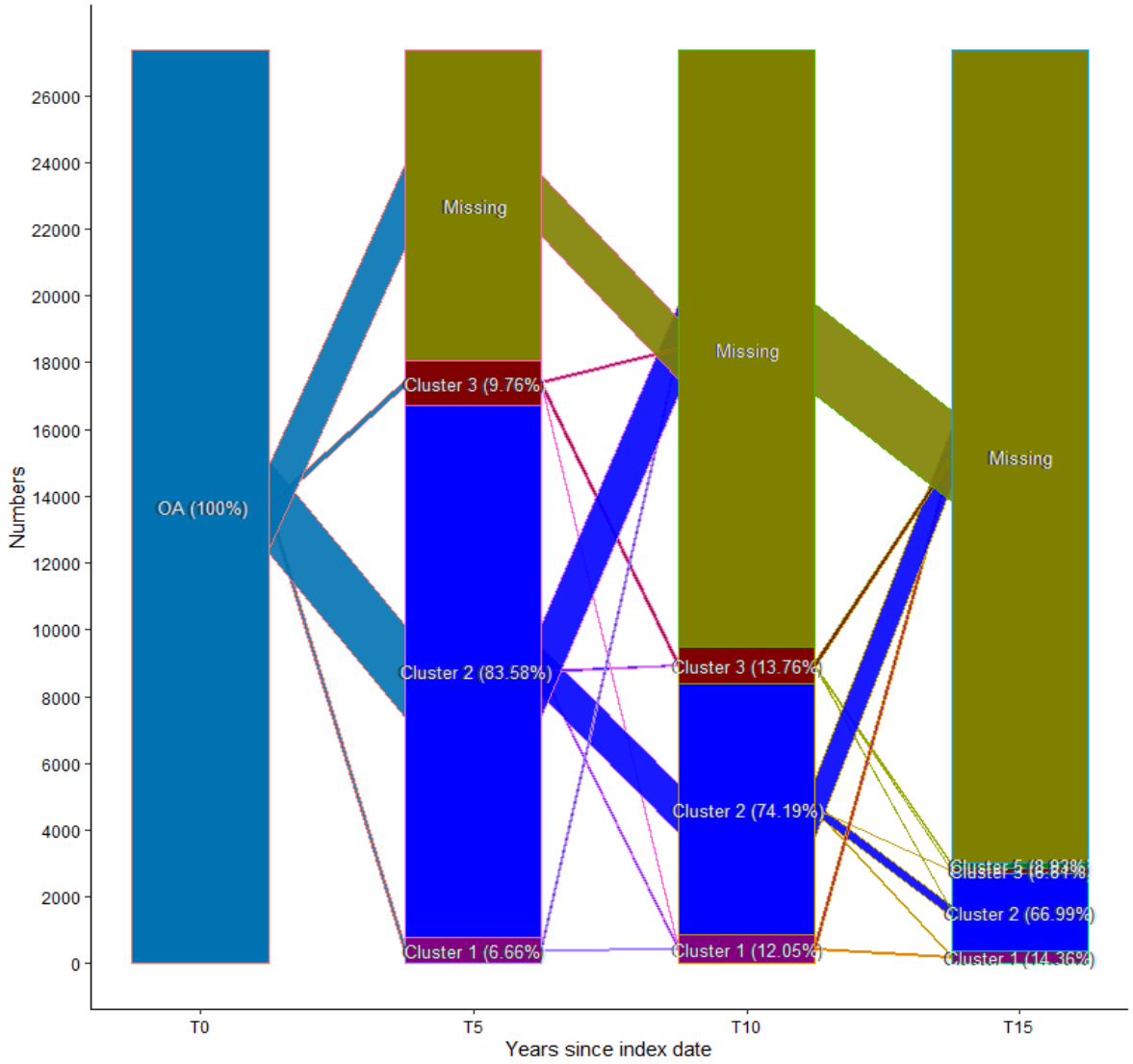
Table 6.3-3 Leading transition paths in OA (sensitivity analysis)

	Year 5	Year 10	Year 15	n	Transition %
Path 1	Healthy (83.58%)	Healthy (74.19%)	Healthy (66.99%)	2262	47.37
Path 2	Healthy (83.58%)	Healthy (74.19%)	CV (14.36%)	160	3.35
Path 3	Healthy (83.58%)	CV (12.05%)	CV (14.36%)	134	2.80
Path 4	MSK (9.76%)	MSK (13.76%)	MSK-MH (8.93%)	110	2.30
Path 5	CV (6.66%)	CV (12.05%)	CV (14.36%)	98	2.05
Path 6	MSK (9.76%)	MSK(13.76%)	MSK (6.81%)	61	1.27
Path 7	Healthy (83.58%)	MSK (13.76%)	MSK-MH (8.93%)	61	1.27
Path 8	Healthy (83.58%)	Healthy (74.19%)	MSK (6.81%)	50	1.04
Path 9	Healthy (83.58%)	MSK (13.76%)	MSK (6.81%)	48	1.00
Path 10	Healthy (83.58%)	MSK (13.76%)	Healthy (66.99%)	48	1.00

Names in each cell at each time point represent the cluster and the percentage represents the class size.

CV- Cardiovascular; MH- Mental health; MSK- Musculoskeletal

Figure 6.3-8. Transition of individuals across clusters in OA (sensitivity analysis)



Cluster 1- CV; Cluster 2- Relatively healthy; Cluster 3- MSK; Cluster 4- CV-MSK; and Cluster 5 – MSK-MH. Thickness of line represents the size of the transition and the colour represents the cluster colours. Percentages in the bracket represents the cluster size.

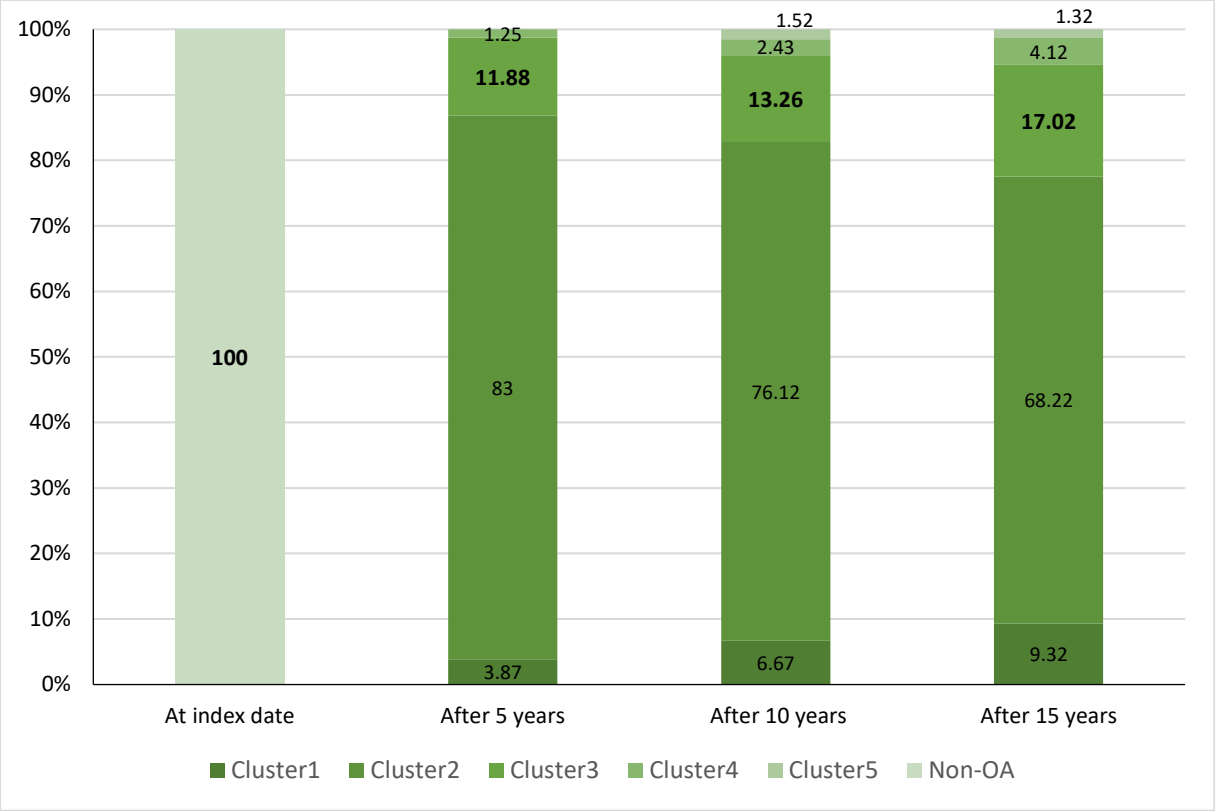
CV- Cardiovascular; MH- Mental health; MSK- Musculoskeletal

6.3.3.2 Clusters in the non-OA population at each time point

Clusters identified in the non-OA population at each time point through LCA are given in Figure 6.3-9. Four clusters at 5 years and five clusters at 10 and 15 years were selected based on the statistical parameters. (Appendix Table 32, page 342) The four clusters identified at 5 years were: cluster 1 (led by hypertension), cluster 2- relatively healthy, cluster 3 (led by back pain) and cluster 4 (led by hypertension and back pain). The size of cluster 1

(led by hypertension) increased from 3.87% at index date to 9.32% at 15 years. Similarly, cluster 2 increased from 11.88% at year 5 to 17.02% at year 15, while the size of the relatively healthy cluster reduced at each follow-up time. Details of the posterior class distribution for each time point are given in Appendix Table 36-38 (page 346-348).

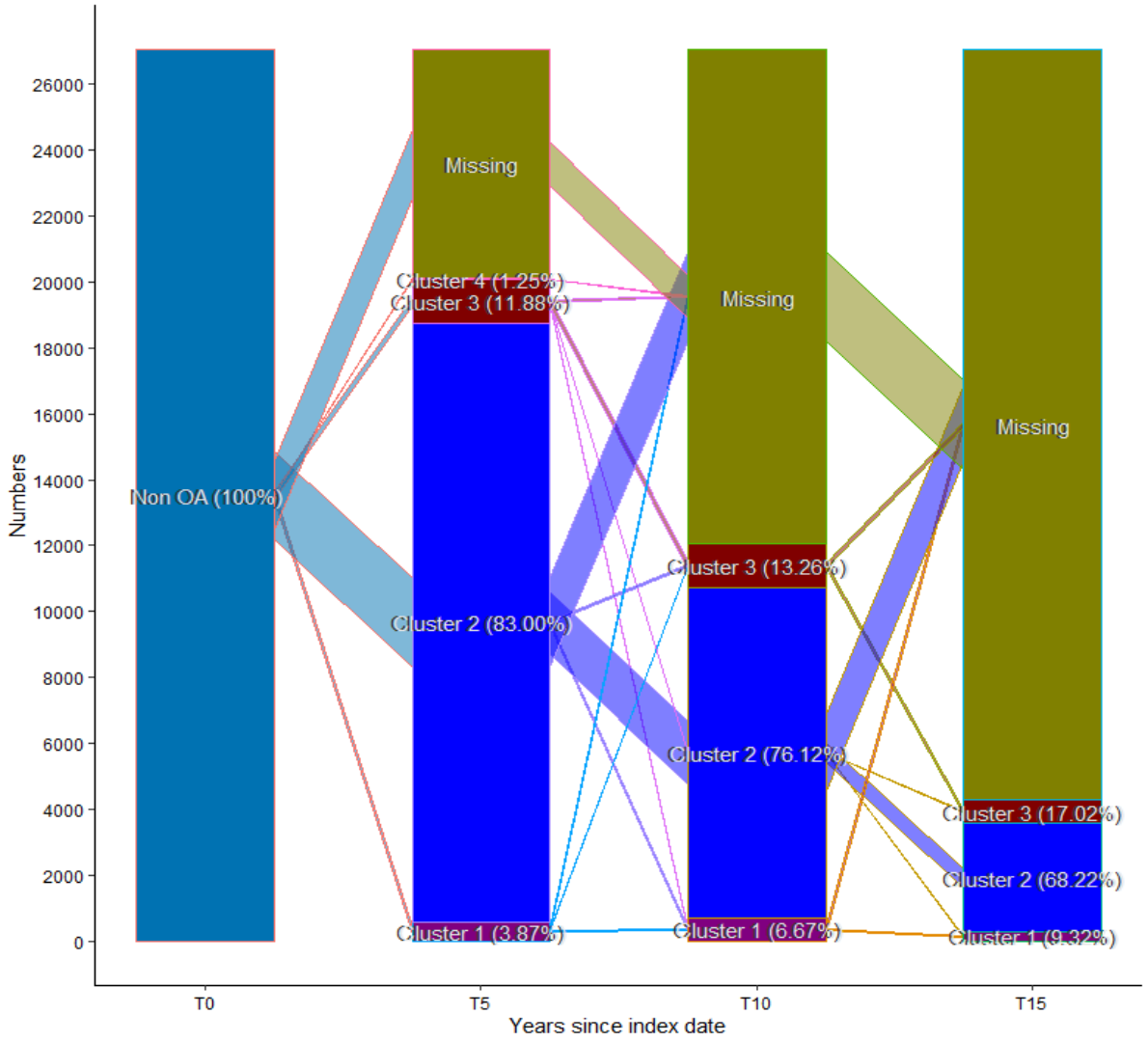
Figure 6.3-9 Different clusters in the non-OA group at each time point (sensitivity analysis)



Cluster 1- CV; Cluster 2- Relatively healthy; Cluster 3- MSK; Cluster 4- CV-MSK; and Cluster 5- Thyroid
CV- Cardiovascular; MSK- Musculoskeletal

Figure 6.3-10 represents the top 20 transitions pattern across the cluster in the non-OA population. Leading trajectory paths identified in the group are given in Table 4. Nearly 69% of the total transition was within the relatively healthy cluster in subsequent years. Transition within cluster 3 in subsequent years was second highest with 6.55% followed by moving to cluster 3 from healthy groups. (Table 6.3-6)

Figure 6.3-10 Transition of individuals across cluster in the non-OA group (sensitivity analysis)



Cluster 1- CV; Cluster 2- Relatively healthy; Cluster 3- MSK; Cluster 4- CV-MSK; and Cluster 5- Thyroid
CV- Cardiovascular; MSK- Musculoskeletal

Table 6.3-4 Leading transition paths in the non-OA population (sensitivity analysis)

	Year 5	Year 10	Year 15	N	%
Path 1	Healthy (83.00%)	Healthy (76.12%)	Healthy (68.22%)	3316	69.44
Path 2	MSK (11.88%)	MSK (13.26%)	MSK (17.02%)	313	6.55
Path 3	Healthy (83.00%)	MSK (13.26%)	MSK (17.02%)	225	4.71
Path 4	Healthy (83.00%)	Healthy (76.12%)	MSK (17.02%)	156	3.26
Path 5	Healthy (83.00%)	CV (6.67%)	CV (9.32%)	109	2.28
Path 6	CV (3.87%)	CV (6.67%)	CV (9.32%)	91	1.90
Path 7	Healthy (83.00%)	Healthy (76.12%)	CV (9.32%)	64	1.34
Path 8	Healthy (83.00%)	CV (6.67%)	Healthy (68.22%)	36	0.75
Path 9	MSK (11.88%)	CV (6.67%)	CV (9.32%)	29	0.60
Path 10	MSK (11.88%)	Healthy (76.12%)	Healthy (68.22%)	26	0.54

Names in each cell at each time point represent the leading conditions in the cluster and the percentage represents the class size. CV- Cardiovascular; MSK- Musculoskeletal

6.4 Discussion

This study identified latent clusters and transition pattern of comorbidities in OA and non-OA controls using LTA. LTA differs from LCA by using longitudinal approaches considering the disease status at each follow-up time, rather than cross-sectional or at a single time point in LCA. This dynamic nature of LTA makes it possible to explore the transition and explains the life course changes in clusters. Tracing the evolution of multimorbidity clusters and their clinical trajectories over time in OA and non-OA led to three major findings. The first was over 15 years, cluster size change and peoples' transitions from one cluster to another generated a well-defined dynamic clinical trajectory. In OA, most of the identified clusters at the index date were stable during the study period. The large dynamic transition in the OA group started to appear after five years, by which time nearly 30% of people had moved towards the CV-MSK clusters and CV cluster only. The class size of CV-MSK clusters increased over time while the size of the MH cluster remained nearly constant. Thirdly, in the non-OA group the transition was less frequent compared to OA and mainly was towards the CV-MSK cluster, which was more prominent after 10 years of index date.

Studies on multimorbidity have explored the pattern and progress of clusters of chronic conditions in different settings, countries, and populations but mostly have been cross-sectional in nature (Prados-Torres *et al.*, 2014). Some studies have focused on a single index disease and its comorbidities (Xu *et al.*, 2018). Recently three studies have explored the transition pathways in multimorbidity clusters (Jensen *et al.*, 2014; Guisado-Clavero *et al.*, 2018; Vetrano *et al.*, 2020) but so far, no studies have been done for multimorbidity trajectory paths in OA. Also, the mentioned previous studies varied widely due to the nature of the data, study population, number of chronic conditions included and methodology. For example, Vetrano et al examined the transition path of multimorbidity among older people aged 50 years or more (Vetrano *et al.*, 2020). Absence of available literature on transition of multimorbidity in OA makes any comparisons difficult.

6.4.1 Transition in the OA population

This longitudinal clustering analysis done by LTA provided five clusters in both the OA and non-OA groups. Firstly, there was higher attrition rate in the OA compared to the non-OA group. The higher attrition could be due to the higher mortality rate in the OA population (Hawker *et al.*, 2014; Barbour *et al.*, 2015) and ageing. The reduction in population size at 10 and 15 years of follow-up could be explained by the long follow-up period. Even though different method was used compared to the LCA in the previous chapter, the consistency of the clustering patterns is encouraging. The class size of the relatively healthy group was the highest (37.6%) which decreased over time indicating the transition of people from this cluster to other disease specific clusters. This is probably again could be due to ageing. Quite stable transition patterns were seen for MSK, and CV-MSK cluster. This means that most people belonging to these clusters at baseline tend to remain in the same clusters until 15 years. Stability of these clusters explains the non-evolving patterns. This could be due to the early and accurate diagnosis of conditions at baseline, and it is possible that these people have undergone thorough screening for other conditions, or they have developed the diseases much earlier due to some other reasons or risk factors. However, some clusters were highly dynamic in nature and evolved with time, for example the CV, MH, and the relatively healthy cluster. All the clusters had overlapping conditions and were dominated by a few diseases. Thus, the evolving dynamic clusters may be influenced by many factors such as other conditions with low prevalence, and biological, pharmacological and socio-psychological factors which increase the susceptibility of future diagnosis (Calderón-Larrañaga *et al.*, 2019).

After five years from the diagnosis date of OA, nearly 15% of people from the relatively healthy group moved to the CV cluster suggesting dominance of hypertension in later age. People with OA are reported to have a higher risk of developing CV, which is supported by the change in clusters in this study (Hall *et al.*, 2016). A similar explanation applies to the transition of people from the clusters led by MSK or CV only to the complex cluster of CV-

MSK. Extensive scientific evidence supports the association of chronic pain in OA (Scherer *et al.*, 2016) and coexistence of MSK and CV (Goodman *et al.*, 2016). Apart from the explanation above, the transition to CV clusters could be due to shared risk factors such as obesity and ageing and to treatment choices such as NSAIDs especially cox-2 inhibitors (Zhang *et al.*, 2019). Another major transition was seen after 10 years from the index date from the MH cluster to the CV-MSK cluster. Coexistence of depression in OA and other painful conditions is common (Bair *et al.*, 2003). The incidence of CVD in depression is well studied (Lespérance *et al.*, 2002) and the path of OA to depression and at a later stage to CVD may be explained by the theory of inflammation, and other factors (Vaccarino *et al.*, 2007). Observed transition path suggests accrual of multiple chronic conditions after OA diagnosis and the evolution of clusters. Large number of transitions to other clusters could be due to initial missed diagnosis or under-diagnosis such as for depression in other clusters. Also, due to the overlapping nature of conditions in each cluster the names of the clusters used needs cautious interpretation.

The associated risk factors for transition probabilities could not be estimated due to large attrition. One of the major requirements for such an analysis is ‘complete-case’ observation for all the time points. However, selecting only people with complete observations for each time point (“survivors”) might lead to selection bias.

6.4.2 Transition in the non-OA population

In the non-OA group, five clusters were found with different cluster sizes and with thyroid disease as a new cluster (i.e., metabolic) compared to the OA group. Slightly more than 40% of the total population was in the relatively healthy cluster at baseline which was higher than that of the OA group. The size of the CV-MSK cluster increased at each time point from the index date. This indicates that a large number of people were having the combination of two conditions at a later age, as would be expected from ageing which is a recognised risk factor for both CV and MSK diseases. In the OA group MSK was more prominent from the

index date, whereas, in the non-OA group it developed later. The clusters in the non-OA group were more stable than OA, which means once assigned to a cluster, the person tends to remain in the same cluster for a longer period. As with OA, the transition path was either from the CV or MH clusters to the cluster led by both CV-MSK, together. However, the transition became more frequent after 10 years from the index date. The mediating disease for these people was hypertension. That is, they moved to the CV cluster first, then moved on to the CV-MSK cluster. The less frequent transition pathways seen in non-OA contrasted to OA could be due to the disease itself or the consequences. However, the delayed transition pattern in non-OA compared to OA group indirectly suggests some underlying role of OA in early evolution of clusters. The transition paths relay the importance of both physical and psychological conditions and their progress also the crucial coexistence of painful MSK conditions. Current evidence on the nexus of these conditions is confined to cross-sectional data (Zhu *et al.*, 2020).

In the sensitivity analyses, the clusters in both the groups changed in numbers, types, and sizes at each time point. Even though RMLCA method was used, a similar pathway was reported in both the OA and non-OA groups. People in the relatively healthy group moved to the cluster led by hypertension in later life then to the cluster led by hypertension and back pain together. Whereas people belonging to the back pain cluster moved to the cluster led by both back pain and depression later. This reflects consistency with the main analysis. As the population used for the sensitivity analysis did not have any chronic conditions at the index date, the clusters of MSK become more frequent compared to other conditions. Xu et al reported that the development of CVD in middle-aged women during 20 years follow-up was higher among those with pre-existing arthritis and/or mental disorders. (Xu *et al.*, 2018) In both the analyses, the accumulation of comorbidities becomes more evident with time, which in turn increases the complexity of management, exposure to pharmacological effects, and reduces the functional impairment and increases load to the health system. Similar

findings have been suggested by a few published previous multimorbidity transition analysis (Jensen *et al.*, 2014; Ibarra-Castillo *et al.*, 2018, p. ; Vetrano *et al.*, 2020).

6.4.3 Strengths and limitations

The major strengths of this study are the inclusion of many chronic conditions (n=49), use of a representative wide age-range of adults, long follow-up period and use of a large GP database. Even though it was possible to estimate the disease clustering at every year, five years follow-up was done to allow sufficient time for diagnosis of new comorbidity. This is the first study to explore the transition pattern in OA and matched controls. LTA, which is a robust method was used to account for the dynamic change in clusters and the missing data. Another advantage of LTA is that it provides a probability of membership for each condition to each cluster, which is the true clinical feature than a distinct exclusive group. Each person was assigned a probability of belonging to a cluster. Inclusion of a wide variety of conditions allowed examination of clusters and trajectories centred around both physical and mental health conditions. The advantage of sensitivity analysis was it could find the full trajectory of cluster formation from people at risk. This sub-cohort had low attrition rate(death or moved out) which allows better estimate for the transition.

There are several limitations to the study. The first is the possible misclassification bias mentioned in previous chapters. Also, consideration of the diseases without considering the severity or chronicity of the condition, but it is likely that the more complex or severe the disease is the chances of developing comorbidities increase. Nevertheless, the evolution of disease patterns in this study covers an important knowledge gap in OA. Secondly, the high dropout rate of participants due to death or leaving the practice did not allow for calculation of transition probabilities associated risk factors. Even though LTA has inbuilt methods to adjust for missingness at random, the availability of complete data would have allowed exploration of further risk factors. Heterogeneous clusters were not found as in other studies and the cluster trajectories used in this study were centred around three common conditions

namely, hypertension, back pain, and depression. Non finding of exclusive patterns of other conditions could be because of the population structure, younger population, and high prevalence of these diseases. But the clusters evolution reported in this study has covered both physical and mental conditions making it more holistic.

6.4.4 Clinical implication

Over their life course, individuals develop multiple diseases. Understanding the diseases clusters, and importantly pathways of these over subsequent years not only help in understanding the complexity and dynamic evolution of multimorbidity clusters but also informs clinicians and health policy makers to plan better management and resource allocation. This study identifies the people at risk of progressing to complex severe disease clusters that may associate with worse outcomes. Reported clusters of conditions here is based on the patients rather than diseases, which provides crucial information for a person-centred care approach. Nearly one third pf people remained in the 'relatively healthy' group with the lowest count of comorbidities. Results from this study can encourage the planning of future randomised clinical trials toward the better management of multimorbidity clusters in OA. Also, this can help for economical calculation for the prediction of burden of diseases.

6.4.5 Conclusion

In conclusion, clusters of multimorbidity in OA and non-OA are characterized by great dynamism and complexity but can still be tracked over time. Large database with a wide range of conditions allowed to find and map evolution of clusters. Few definite pathways were found such as developing a single chronic condition at a young age and later moving to complex clusters. These could be due to shared risk factors, pathophysiology, drug use, or merely unrelated coexistence. Future studies can be focused within each cluster to examine the biological and physiological linkages in these conditions. Also, the outcomes of these evolving clusters must be studied to determine the severity. Last but not the least the identified clusters and their possible transitions can guide every health care practice level for better tailoring of the target population in future interventions for comorbidities in OA.

Summary of Chapter 6

Chapter 6 described the transition or the movement of people from one identified cluster to another over time among people with OA and the matched controls. The key messages are:

- Nearly 30% of the people with OA after five years of the diagnosis move to the cluster of cardiovascular only or cardiovascular-musculoskeletal.
- The size of the cardiovascular only or cardiovascular-musculoskeletal clusters in people with OA increases over time, whereas number of people in depression cluster remains almost unchanged.
- Among people without OA the transition was less frequent compared to OA and mainly was towards the cardiovascular-musculoskeletal cluster, which was more prominent after 10 years of index date.

As seen, the development of cardiovascular and cardiovascular-musculoskeletal disease clusters growing by size in both the group, indicates the people with OA develop or get diagnosed with cardiovascular sooner compared to their non-OA counterparts.

However, the progression with the number of comorbidities is not understood clearly. That is how does the progression of multimorbidity happens in people with OA and non-OA and how can the population be grouped based on the rate of progression is answered in chapter 7.

7 Chapter 7

Trajectories of multimorbidity

7.1 Introduction

As demonstrated in Chapter 4, Table 4.3-6, the risk of multimorbidity was nearly 1.3 times greater in people with OA compared to those without. One of the interests I had was the trajectories of cumulative multimorbidity. This means how the accumulation of chronic conditions grows over period after the index date. Identifying such clusters or group of individuals would help to prepare the prevention strategy in advance. Understanding the development of multimorbidity in people with OA could help to identify long-term associated outcomes, prognostic factors, and design interventions. Studies have measured multimorbidity at single time points and examined associations and clusters. A study from the USA found the rate of increasing multimorbidity varied rapidly over 5-6 years (Quinones A.R. *et al.*, 2011). No studies have attempted to describe the accumulation of morbidities based on many diverse conditions and to identify whether there are distinct trajectories of multimorbidity over time in OA using primary care consultations.

In the UK, a multimorbidity trajectory was reported in a consultation database including 37 conditions (Strauss *et al.*, 2014). However, such trajectories in the OA population have not been reported. Therefore, latent class growth analysis (LCGA) was used to group (cluster) people into distinct trajectories of multimorbidity using a primary care database and a wide range of 49 conditions after the index date. Therefore, the aim of this chapter is to identify the trajectories of the accumulation of multimorbidity over time in both OA and non-OA group and their associations with patient characteristics.

7.2 Methods

7.2.1 Study subjects

The same population described in Chapter 6 for OA and non-OA transition modelling was used. Details are provided on pages 185.

7.2.2 Measurements and covariates

The outcome in this study was the number of comorbidities present in people with OA and in the non-OA group at and after the index date. The term multimorbidity was used as numbers to describe the burden of comorbidities. Covariates considered were the same as described in previous chapters such as age, sex, smoking, alcohol, BMI, and multiple deprivation index.

7.2.3 Statistical analysis

There are various approaches to examine how time and age influence changes in multimorbidity score across the adult life span. One approach is to use LCGA to examine changes and identify the groups. Latent growth curve modelling allows study of the sample as a single population, with the ability to examine model-implied changes and to assess whether there are between-person differences in level and rates of change over time (Grimm, Ram, & Estabrook, 2017).

The guideline for reporting on latent trajectory studies (GRoLTs checklist) was used to assess the feasibility of the trajectory modelling in the study population (van de Schoot *et al.*, 2017). (Appendix Table 39, page 349).

For assessment of trajectory, all the registered patients were followed from their first date of registration until up to 20 years. Twenty years follow-up was done due to maximum data availability. The outcome was the cumulative number of chronic conditions over the years. Trajectories of multimorbidity were assessed using LCGA. LCGA models were fitted starting with a one-cluster model, assuming that all subjects have the same trajectory, and then successively increasing the number of clusters until most of the heterogeneity in the data

was explained (Muthén and Asparouhov, 2009). Counts in each period were assumed to be Poisson distributed. Cubic growth curves were applied for all clusters identified within the LCGA models. For each model, people were assigned to the cluster where their posterior probability of membership was highest (the maximum probability assignment rule). Hence, people could only belong to one cluster.

LCGA can be considered as an extension to the fixed effect growth model. Fixed effect growth models are used to measure between group variability, whereas random effect models are used to address the within group variability. LCGA using a random effect growth model is a type of growth mixture model (GMM). For this study, as participants have different ages at the index date, age was used as a random effect in the model. A series of trajectory models of multimorbidity as a function of age, with a class number ranging from 2 to 10, were assessed using the lcmm (version 1.7.9) package in R (version 3.5.0) (Proust-Lima *et al.*, 2020). The age of the participants was centred at the median age of the population and divided by 10 to reduce problems associated with high ages in quadratic and cubic terms in the model. Three possible polynomial specifications of the longitudinal response of multimorbidity as a function of age, namely linear, quadratic, and cubic were considered, to allow for non-linear patterns in both fixed and random effect components. For each model, class-specific variance covariance random-effects was considered, which allowed for between subjects' trajectory variability to differ between classes. To avoid convergence towards local maxima, all models were rerun several times with different starting values and initial values obtained via grid searching (with a maximum of 15 iterations from 30 random vectors of values from the 1-class model).

The optimal number of classes in each of the above three methods was decided using a combination of statistics Bayesian Information Criteria (BIC), sample size adjusted BIC (SBIC), log-likelihood ratio test (LLRT), entropy for classification quality, minimum of 1% total patients in each cluster and clinical judgement. Within the datasets, conditions were present (i.e., recorded) or not by definition, so missing data methods were not needed for cluster

analysis. The optimal model is that which has the lowest BIC value while the LLRT assesses whether adding one further cluster significantly improves the model fit. The model selection is alternatively explained by examining if the model with K-class is better than K-1 class by calculating percentage change in log-likelihood ratio of these two models using the formula (LL2-LL1)*100/(LL1). Additionally, the best model should have entropy more than 0.70 and should make more sense clinically. The clusters were named after the three most contributing chronic conditions (posterior probabilities) in each cluster. Once the best class was identified, the groups were attached to the original database and descriptive analysis was done for the covariates. Multinomial regression model was used to explore the risk factors using the 'relatively healthy' cluster as the reference group.

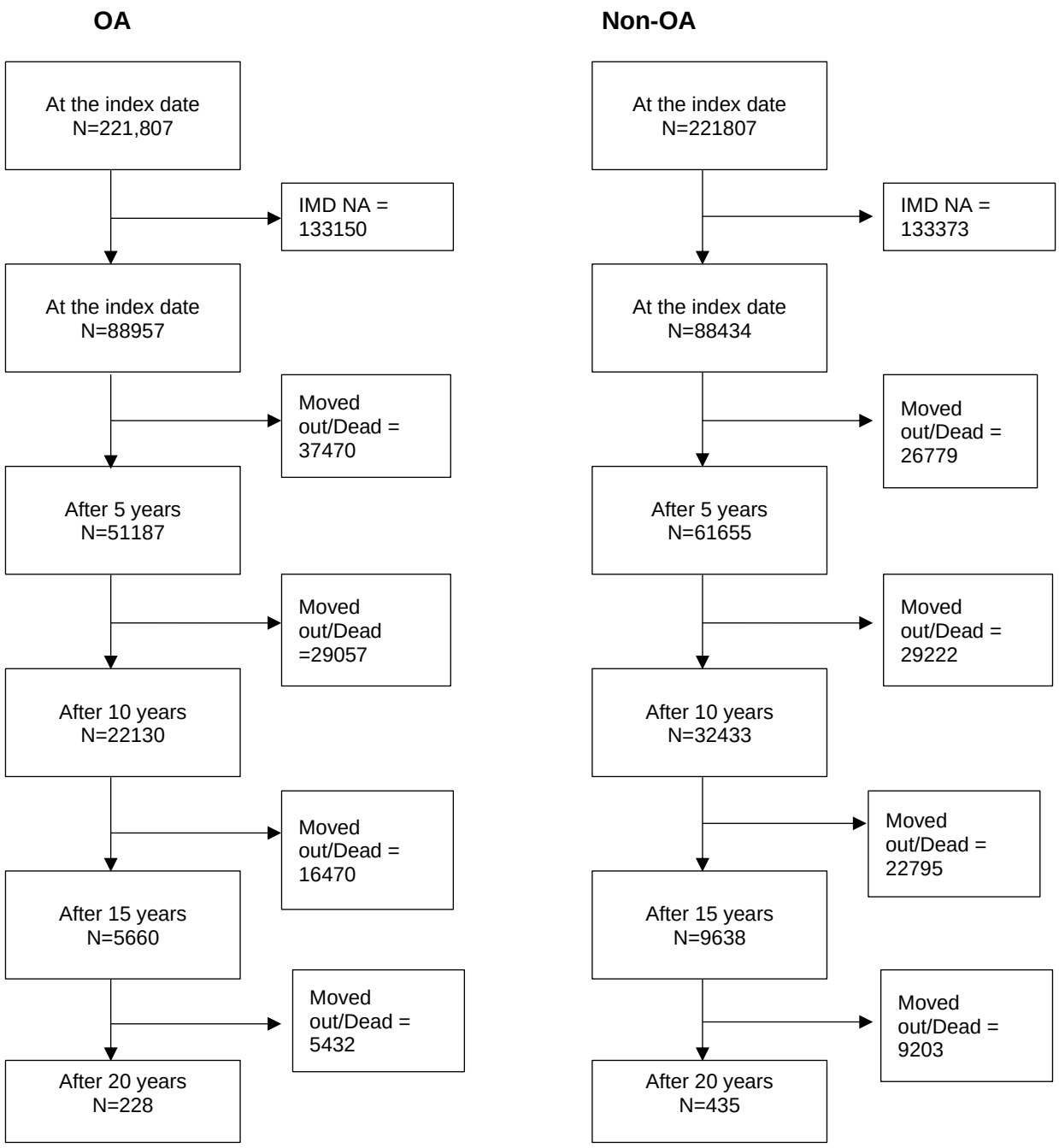
7.2.4 Sensitivity analysis

Sensitivity analysis was done among the subgroups of people with OA and controls without any comorbidities at the index date. A similar method was used to find the trajectory groups.

7.3 Results

Of the 221,807 people with OA, information on the English index of multiple deprivation was available for 88,957. In the non-OA group 88,434 of 221,807 people had data on the English index of multiple deprivation. Only people with complete information on the deprivation index were included in the analysis. The patient flow diagram for each follow-up time is given in Figure 7.3-1.

Figure 7.3-1. Flow diagram showing the number of people at each time of follow up



IMD- Index of multiple deprivation; NA- Not available

7.3.1 Descriptive statistics

Table 7.3-1 describes the sample characteristics at each time point. Nearly 57% were female at index date in both groups, which stayed the same in the non-OA group and reduced to 53% in the OA population after 20 years. The mean age was increasing over time in both the groups, with a younger population being in the non-OA group. The prevalence of obesity (measured at baseline) in the OA group was constantly high than in the non-OA group at each time. A wide difference was observed in prevalence of multimorbidity. Nearly 85% of the OA population had multimorbidity at the index date compared to 51% in the non-OA group. However, it increased in both groups at each time point.

Table 7.3-1 Description of characteristics of the OA and non-OA populations at each time point

Variable	At index date	After 5 years	After 10 years	After 15 years	After 20 years
OA	**N= 88957**	**N=51187**	**N=22130**	**N=5660**	**N=228**
Gender (% Female)	51528 (57.92)	29495 (57.62)	12563 (56.77)	3136 (55.41)	122 (53.51)
Age at index date in years (Mean, SD)	61.27 (13.09)	64.77 (12.74)	67.55 (12.12)	69.70 (11.68)	71.42 (23.08)
Obesity (%)*	28105 (31.59)	16551 (32.33)	7372 (33.31)	1924 (33.99)	76 (33.33)
Multimorbidity (%)	75231 (84.57)	46127 (90.11)	20556 (92.89)	5351 (94.54)	213 (93.42)
Number of comorbidities (Mean, SD)	3.56 (2.11)	4.22 (2.39)	4.75 (2.59)	5.21 (2.78)	5.35 (2.89)
Non-OA	**N= 88434**	**N=61655**	**N=32433**	**N=9638**	**N=435**
Gender (% Female)	51246 (57.95)	35514 (57.60)	18509 (57.07)	5428 (56.32)	248 (57.01)
Age at index date in years (Mean, SD)	61.04 (13.17)	65.32 (13.18)	68.84 (12.88)	71.36 (12.58)	73.42 (11.52)
Obesity (%)*	18369 (20.77)	12685 (20.57)	6622 (20.42)	2008 (20.83)	77 (17.70)
Multimorbidity (%)	44740 (50.59)	38017 (61.66)	22442 (69.19)	7214 (74.85)	333 (76.55)
Number of comorbidities (Mean, SD)	1.97 (1.89)	2.52 (2.13)	3.02 (2.35)	3.43 (2.50)	3.75 (2.63)

SD- Standard deviation; * at the baseline

7.3.2 Trajectories of multimorbidity in OA

For the OA group, the five class model provided the best fit based on the statistical parameters. Average posterior probability (PP) of people ranged from 0.66 to 0.99 across each cluster. Also, the lowest BIC and AIC value suggested a five class solution to be the best one. (Table 7.3-2)

Table 7.3-2. Summary statistics of the latent class growth analysis across 20 years

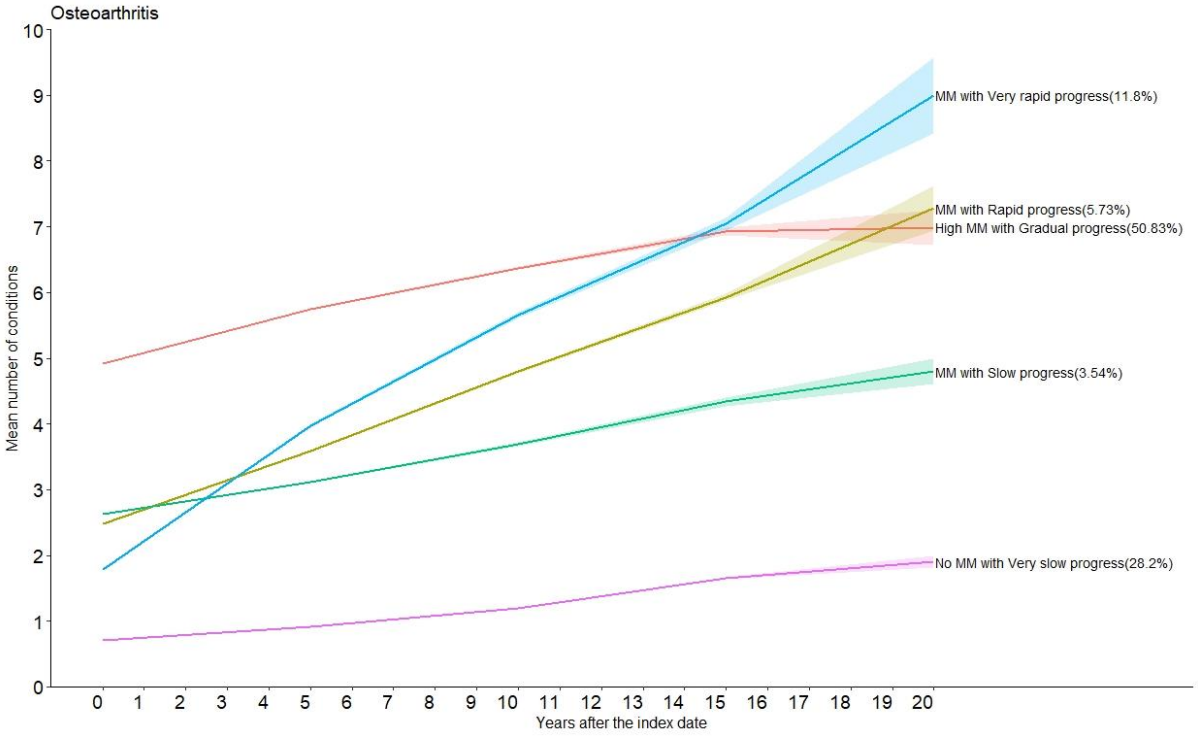
Number of clusters	Log likelihood	Number of parameters	BIC	AIC	Entropy (%)
1	-296515	12	593166.3	593053.5	100
2	-294226	15	588623.6	588482.6	78
3	-293946	18	588097.4	587928.3	76
4	-293997	21	588267.9	588042.3	72
5	-293323	24	586885.6	586688.3	74
6	-293379	27	586985.7	586703.5	70
7	-293694	31	587068.3	586832.1	68
8	-293753	33	587357.1	586957.3	67
9	-293892	36	587525.7	587203.5	65
10	-293894	39	587532.9	587285.8	65

BIC- Bayesian information criteria, AIC- Akaike information criteria

Cluster specific trajectories suggested these can be described as per the growth pattern. In people with OA after the index date, five patterns were seen: multimorbid with very rapidly progressing (11.8%), multimorbid at index date with rapidly progressing (5.73%), high multimorbid at index date but gradual progress (50.83%), multimorbid with slow progress (3.54%) and non-multimorbid with very slow progressing (28.2%). One cluster (28.2%) had zero to one chronic condition at the index date and continued to have less than 20 chronic conditions until 20 years of follow up. This group was referred to as ‘very slow progressing’ for further analysis. Nearly 12% of people had <2 chronic conditions at the index date but within five years they had nearly 4 which increased sharply to nearly 9 after 20 years. This group was termed the very rapidly progressing group and can be considered as high-risk group of developing multimorbidity. Half of the population had an average of five conditions at the index date, and this continued to increase slowly over the years. (Figure 7.3-2)

Another 5.7% of the OA population had multimorbidity at the index date which was doubled at each five years of follow up.

Figure 7.3-2. Clusters of multimorbidity trajectories over time in the OA group



Solid line represents the observed mean number of chronic conditions at each time. Shaded area represents the confidence interval. MM-multimorbid at the index date

Table 7.3-3 describes the characteristics of variables at the index date across the identified clusters. The highest proportion of women was found in the cluster of ‘ multimorbid with gradual progress’ (60%). Multimorbid with very rapid progress and rapid progress clusters had higher mean age compared to other groups. The prevalence of ex-smokers (31.50%) and ex-drinkers (3.26%) was highest in the ‘multimorbid with very rapid progress’ group. Nearly 34% in the ‘ multimorbid with gradual progress’ and 32% in ‘multimorbid with very rapid progress’ cluster were obese. In the ‘ multimorbid with gradual progress’ and ‘ multimorbid with very rapid progress’ clusters 16% belonged to the highest deprivation index, which was high compared to other clusters.

Table 7.3-3 Descriptive characteristics of baseline variables across the clusters in the OA group

Variables	Non-multimorbid with very slow progress n=24679	Multimorbid with slow progress n=3214	Multimorbid with gradual progress n=49213	Multimorbid with rapid progress n=3872	Multimorbid with very rapid progress n=7979
Gender					
Men	11212(45.43)	1443(44.90)	19470(39.56)	1732(44.73)	3572(44.77)
Women	13467(54.57)	1771(55.10)	29743(60.44)	2140(55.27)	4407(55.23)
Age	60.77(13.01)	60.51(13.91)	61.47(13.12)	61.11(13.10)	62.00(12.69)
Smoking					
Never smoked	13834(56.06)	1758(54.70)	25352(51.51)	2112(54.55)	4084(51.18)
Current smoker	4337(17.57)	565(17.58)	9358(19.02)	618(15.96)	1382(17.32)
Ex-smoker	6508(26.37)	891(27.72)	14503(29.47)	1142 (29.49)	2513(31.50)
Alcohol use					
Never	4088(16.56)	543(16.89)	10034(20.39)	685(17.69)	1674(20.98)
Ex-drinker	483(1.96)	68(2.12)	1553(3.16)	110(2.84)	260(3.26)
Current (1-9)	9261(37.53)	1180(36.71)	16630(33.79)	1329(34.32)	2668(33.44)
Current (>=10)	5589(22.65)	730(22.71)	9713(19.74)	853(22.03)	1607(20.14)
Current (Unknown)	5258(21.31)	693(21.56)	11283(22.93)	895(23.11)	1770(22.18)
BMI					
Normal	8089(32.78)	1066(33.17)	13499(27.43)	1092(28.20)	2252(28.22)
Underweight	320(1.30)	37(1.15)	690(1.40)	61(1.58)	134(1.68)
Overweight	9753(39.52)	1250(38.89)	18057(36.69)	1503(38.82)	3049(38.21)
Obese	6517(26.41)	861(26.79)	16967(34.48)	1216(31.40)	2544(31.88)
Multiple deprivation					
IMD1 (Lowest)	5996(24.30)	837(26.04)	11018(22.39)	926(23.92)	1703(21.34)
IMD2	5728(23.21)	768(23.90)	11001(22.35)	921(23.79)	1914(23.09)
IMD3	5327(21.59)	705(21.94)	10175(20.68)	802(20.71)	1689(21.17)
IMD4	4318(17.50)	489(15.21)	9016(18.32)	683(17.64)	1415(17.73)
IMD5 (Highest)	3310(13.41)	415(12.91)	8003(16.26)	540(13.95)	1258(15.77)

BMI- Body mass index; IMD- Index of multiple deprivation.

7.3.3 Factors associated with trajectory groups

Multinomial regression model findings are provided in Table 7.3-4. Women had 1.2 times higher risk of being in the multimorbid with gradual progress cluster compared to men. Being ex-smokers and ex-drinkers increased the risk by 1.2 times to be in all the clusters except for 'non- multimorbid and slow progress' compared to non-smokers and non-drinkers, respectively. Obesity had the highest association with the 'multimorbid with gradual progress' cluster (OR 1.56; 95% CI 1.49-1.62) followed by the multimorbid with very rapid progress (OR 1.37; 95% CI 1.28-1.47) and multimorbid with rapid progress (OR 1.36; 95%

CI 1.24-1.49) clusters. Also, being underweight had significant associations with the multimorbid with gradual progress (OR 1.17; 95% CI 1.02-1.34), multimorbid with very rapid progress (OR 1.40; 95% CI 1.27-1.67) and multimorbid with rapid progress (OR 1.39; 95% CI 1.05-1.84) clusters. Higher deprivation index score had significant association with the multimorbid with gradual progress (OR 1.19; 95% CI 1.15-1.24) and multimorbid with very rapid progress (OR 1.15; 95% CI 1.05-1.23) clusters,.

Table 7.3-4 Factors associated with clusters from LCGA in OA

Variables	Non-multimorbid with very slow progress OR (95% CI)	Multimorbid with slow progress OR (95% CI)	Multimorbid with gradual progress OR (95% CI)	Multimorbid with rapid progress OR (95% CI)	Multimorbid with very rapid progress OR (95% CI)
Gender					
Men	1	Reference	Reference	Reference	Reference
Women	1	1.03 (0.96-1.08)	1.28 (1.24-1.33) *	1.04 (0.97-1.11)	1.02 (0.98-1.06)
Age	1	1.00 (0.99-1.01)	1.00 (1.00-1.01) *	1.00 (1.00-1.01) *	1.01 (1.01-1.02) *
Smoking					
Never smoked	1	Reference	Reference	Reference	Reference
Current smoker	1	1.03 (0.95-1.11)	1.23 (1.18-1.29) *	0.94 (0.86-1.05)	1.11 (1.03-1.19) *
Ex-smoker	1	1.09 (1.01-1.19) *	1.25 (1.21-1.29) *	1.14 (1.05-1.23) *	1.30 (1.22-1.38) *
Alcohol use					
Never	1	Reference	Reference	Reference	Reference
Ex-drinker	1	1.04 (0.79-1.37)	1.29 (1.16-1.44) *	1.35 (1.08-1.68) *	1.24 (1.06-1.46) *
Current (1-9)	1	0.93 (0.84-1.04)	0.78 (0.75-0.82) *	0.86 (0.78-0.96) *	0.72 (0.67-0.78) *
Current (>=10)	1	0.95 (0.84-1.08)	0.79 (0.75-0.84) *	0.93 (0.83-1.04)	0.71 (0.66-0.78) *
Current (Unknown)	1	0.97 (0.86-1.10)	0.89 (0.85-0.94) *	1.01 (0.91-1.13)	0.83 (0.78-0.87) *
Body mass index					
Normal	1	Reference	Reference	Reference	Reference
Underweight	1	0.88 (0.62-1.24)	1.17 (1.02-1.34) *	1.39 (1.05-1.84) *	1.40 (1.27-1.67) *
Overweight	1	0.98 (0.89-1.06)	1.14 (1.10-1.19) *	1.13 (1.04-1.23) *	1.12 (1.08-1.17) *
Obese	1	1.00 (0.91-1.10)	1.56 (1.49-1.62) *	1.36 (1.24-1.49) *	1.37 (1.28-1.47) *
Multiple deprivation					
IMD3	1	Reference	Reference	Reference	Reference
IMD1 (Lowest)	1	1.06 (0.96-1.18)	1.02 (0.99-1.06)	1.05 (0.98-1.12)	0.94 (0.87-1.01)
IMD2	1	1.01 (0.91-1.13)	1.02 (0.98-1.08)	1.09 (1.02-1.16) *	1.07 (0.99-1.16)
IMD4	1	0.85 (0.75-0.96) *	1.05 (1.01-1.09) *	1.02 (0.95-1.09)	1.01 (0.93-1.07)
IMD5 (Highest)	1	0.93 (0.82-1.06)	1.19 (1.15-1.24) *	1.07 (0.99-1.16)	1.15 (1.05-1.23) *

CI- Confidence interval; IMD- Index of multiple deprivation; OR- Odds ratio; *p value <0.05

7.3.4 Trajectories of multimorbidity in the non-OA group

In the non-OA group, four trajectory clusters were found based on the summary statistics shown in Table 7.3-5.

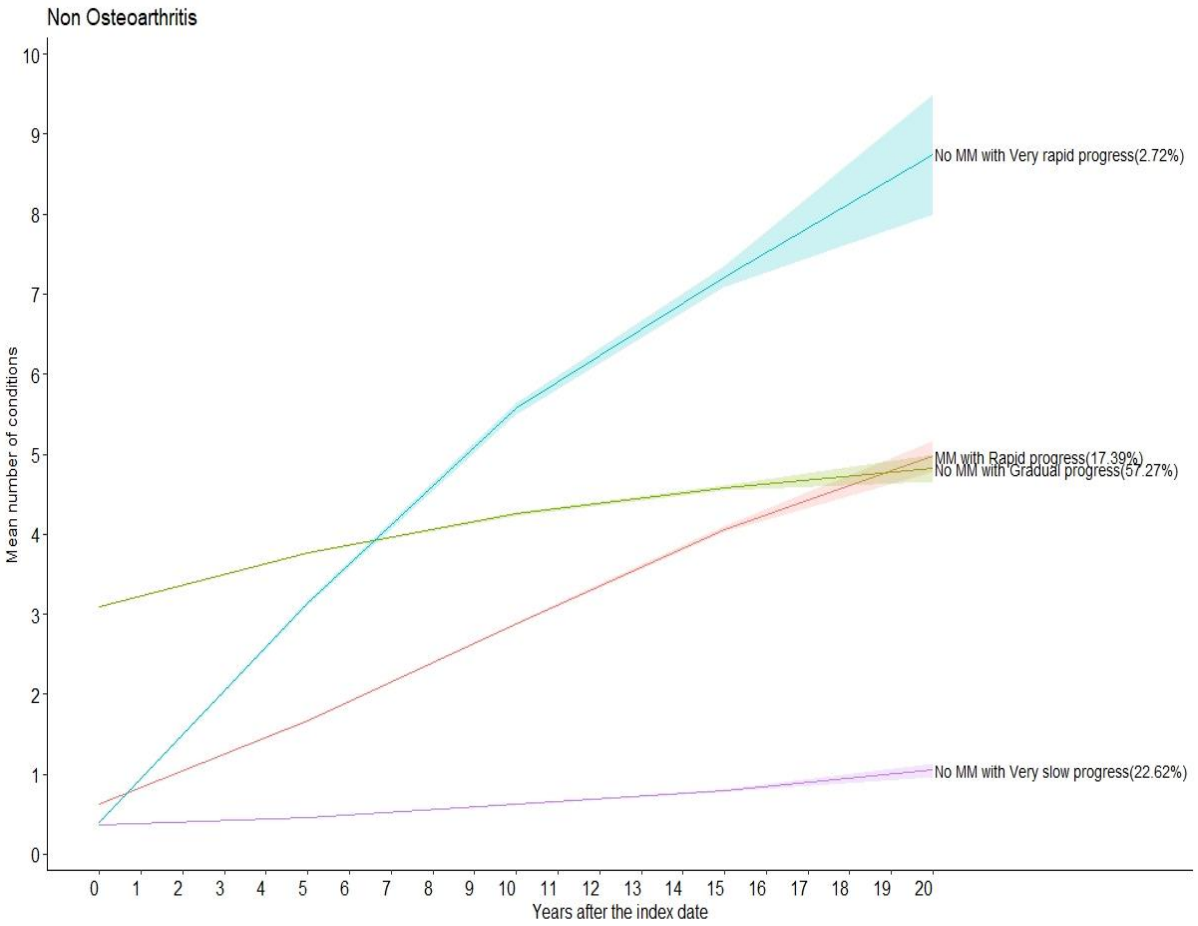
Table 7.3-5 Summary statistics of the latent class growth analysis across 20 years in non-OA.

Number of clusters	Log likelihood	Number of parameters	BIC	AIC	Entropy (%)
1	-305288	12	610712.2	610599.6	100
2	-302407	15	604984.4	604843.6	81
3	-301535	18	603274.4	603105.4	79
4	-300848	21	601935.7	601738.5	78
5	-305304	24	610881	610655.7	76
6	-305130	27	610568.4	610314.9	75
7	-306312	30	610681.8	610479.5	75
8	-307175	33	611068.7	610932.3	73
9	-308285	36	611154.1	611349.1	70
10	-308276	40	611152.4	611337.4	69

BIC- Bayesian information criteria; AIC- Akaike information criteria

The clusters were very rapidly progressing (2.72%), rapidly progressing (17.39%), gradual progressing (57.27%) and very slow progress (22.62%). The cluster of relative healthy always had less than one chronic condition in 20 years. Nearly 3% of the population in the very rapid group had zero chronic conditions at the index date which then it increased suddenly to five at the end of 10 years and to 8 after 20 years. The average number of chronic conditions in the gradually progressing group at index date was three and this increased at a slower rate to 4 at 20 years. Another rapid progress had less than one chronic condition at index date, which then increased faster to four at the end of follow up date. (Figure 7.3-3) Even though the non-OA group had similar clusters to that of the OA group, the burden of multimorbidity was less in each cluster compared to the OA group. For example, the gradual progressing cluster in OA had an average of 5-7 chronic conditions, while in non-OA group it was 3-4. (Figure 7.3-3)

Figure 7.3-3 Clusters of multimorbidity trajectories over time in non-OA



Shaded area represents 95% confidence interval. MM- multimorbid at index date

Table 7.3-6 describes the characteristics of the variables reported at the index date across the identified clusters in the non-OA group. The proportion of women was highest in the non-multimorbid with gradual progressing group (59.55%) followed by the multimorbid with rapidly progressing group (56.21%). The multimorbid with very rapidly progressing cluster had the oldest population with a mean age of 68.5 (SD 13.0) years. The prevalence of ex-smokers (29.43%) and ex-drinkers (3.41%) was higher in the multimorbid with very rapidly progressing cluster compared to others. More than 20% of the population in the multimorbid with gradual progressing (22.78%) and multimorbid with very rapid progressing (20.16%) clusters were obese. The proportion in the most deprived category was higher in the multimorbid with very rapidly progressing (16.92%) and non-multimorbid with gradual progressing (14.23%) groups, than the other groups.

Table 7.3-6 Descriptive characteristics of baseline variables across the clusters in non-OA

Variables	Non-multimorbid with very slow progress n=20,008	Non-multimorbid with gradual progress n=50,645	Multimorbid with rapid progress n=15,375	Multimorbid with very rapid progress n=2406
Gender				
Men	8888(44.42)	20488(40.45)	6732(43.79)	1080(44.89)
Women	11120(55.58)	30157(59.55)	8643(56.21)	1326(55.11)
Age at index date	62.68(13.41)	63.78(13.56)	65.71(12.79)	68.51(12.97)
Smoking				
Never smoked	11959(59.77)	27573(54.44)	8727(56.76)	1284(53.37)
Current smoker	3356(16.77)	9208(18.18)	2430(15.80)	414(17.21)
Ex-smoker	4693(23.46)	13864(27.37)	4218(27.43)	708(29.43)
Alcohol use				
Never	3310(16.54)	9327(18.42)	2753(17.91)	507(21.07)
Ex-drinker	328(1.64)	1391(2.75)	369(2.40)	82(3.41)
Current (1-9)	7791(38.94)	17866(35.28)	5556(36.14)	824(34.25)
Current (>=10)	4409(22.04)	10240(20.22)	3249(21.13)	438(18.20)
Current (Unknown)	4170(20.84)	11821(23.34)	3448(22.43)	555(23.07)
Body mass index				
Normal	8897(44.47)	19140(37.79)	6112(39.75)	924(38.40)
Underweight	441(2.20)	1088(2.15)	348(2.26)	80(3.33)
Overweight	7319(36.58)	18881(37.28)	5918(38.49)	917(38.11)
Obese	3351(16.75)	11536(22.78)	2997(19.49)	485(20.16)
Multiple deprivation				
IMD1 (Lowest)	5109(25.53)	12612(24.90)	3862(25.12)	575(23.90)
IMD2	4728(23.63)	11722(23.15)	3644(23.70)	553(22.98)
IMD3	4302(21.50)	10504(20.74)	3306(21.50)	478(19.87)
IMD4	3366(16.82)	8599(16.98)	2597(16.89)	393(16.33)
IMD5 (Highest)	2503(12.51)	7208(14.23)	1966(12.79)	407(16.92)

IMD- Index of multiple deprivation

7.3.5 Factors associated with trajectory groups in the non-OA group

In the non-OA group, women were 1.2 times more likely to be in the 'non-multimorbid with gradually progressing' group compared to that of in very slow progress group and men. Age was consistently associated with all the clusters, compared to the relatively healthy cluster. Either being a smoker or ex-smoker increased the risk of being in any of the non-healthy clusters. Smokers (OR 1.31; 95% CI 1.15-1.48) and ex-smokers (OR 1.30; 95% CI 1.18-1.44) were at greater risk of being in the multimorbid with very rapidly progressing group. Similarly, ex-drinkers had nearly 1.3 times higher association with the non-multimorbid with gradually progressing and very rapidly progressing clusters compared to the non-multimorbid with very slow progress cluster. Both overweight and obesity were significantly associated with all the non-healthy clusters. Obesity had relative risk ratios of 1.6 and 1.5 for

being in the non-multimorbid with gradual progressing and multimorbid with very rapidly progressing group, respectively. Being underweight also increased the risk of being in the multimorbid with very rapidly progressing group (OR 1.43; 95% CI 1.12-1.84). The most deprived people were more likely to be in the multimorbid with very rapidly progressing cluster (OR 1.42; 95% CI 1.23-1.64) compared to the other clusters. (Table 7.3-7)

Table 7.3-7 Factors associated with clusters from LCGA in non-OA

Variables	Non-multimorbid with very slow progress OR (95% CI)	Non-multimorbid with gradual progress OR (95% CI)	Multimorbid with rapid progress OR (95% CI)	Multimorbid with very rapid progress OR (95% CI)
Gender				
Men	1	Reference	Reference	Reference
Women	1	1.23(1.19-1.27) *	1.06(1.01-1.10) *	0.97(0.89-1.06)
Age at index date	1	1.01(1.01-1.02) *	1.02(1.01-1.02) *	1.03(1.03-1.04) *
Smoking				
Never smoked	1	Reference	Reference	Reference
Current smoker	1	1.26(1.20-1.32) *	1.08(1.01-1.15) *	1.31(1.15-1.48) *
Ex-smoker	1	1.29(1.24-1.34) *	1.18(1.12-1.24) *	1.30(1.18-1.44) *
Alcohol use				
Never	1	Reference	Reference	Reference
Ex-drinker	1	1.46(1.28-1.66) *	1.31(1.12-1.53) *	1.49(1.15-1.94) *
Current (1-9)	1	0.85(0.81-0.89) *	0.90(0.84-0.96)	0.77(0.68-0.87) *
Current (>=10)	1	0.88(0.83-0.93) *	0.95(0.88-1.02)	0.73(0.64-0.85) *
Current (Unknown)	1	1.02(0.96-1.07)	1.02(0.95-1.09)	0.92(0.81-1.05)
BMI				
Normal	1	Reference	Reference	Reference
Underweight	1	1.05(0.93-1.17)	1.05(0.91-.121)	1.43(1.12-1.84) *
Overweight	1	1.22(1.18-1.27) *	1.18(1.12-1.23) *	1.21(1.10-1.33) *
Obese	1	1.61(1.53-1.68) *	1.34(1.26-1.42) *	1.47(1.31-1.66) *
Multiple deprivation				
IMD3	1	Reference	Reference	Reference
IMD1 (Lowest)	1	1.06(1.01-1.12) *	1.01(0.95-1.08)	1.07(0.94-1.22)
IMD2	1	1.03(0.98-1.08)	1.01(0.95-1.07)	1.07(0.94-1.21)
IMD4	1	1.01(0.96-1.07)	0.99(0.93-1.07)	1.02(0.89-1.18)
IMD5 (Highest)	1	1.11(1.05-1.18) *	1.01(0.94-1.10)	1.42(1.23-1.64) *

CI- Confidence interval; IMD- Index of multiple deprivation; OR- Odds ratio; *p value <0.05

7.3.6 Sensitivity analysis

Same LCGA method was used for detecting clusters of trajectories for multimorbidity among the OA and non-OA group without any comorbidities at the index date. Each group had 22,333 patients matched for age (±2), sex and practice. Four clusters in the non-OA and five in the OA group were found to give the best fit model according to the model statistics. The models were selected based on the change in likelihood ratio and each group in the clusters

should have a size of minimum 1%. Details of the model statistics for each group are given in Table 7.3-8.

Table 7.3-8. Statistical parameters of optimal number of clusters from LCGA in the OA group

Classes	Parameters	LL	AIC	BIC	aBIC	Entropy
OA						
1	4	-214039	428085.7	428118.1	428105.4	
2	9	-119739	239495.2	239568.1	239539.5	0.969
3	14	-108897	217821.2	217934.6	217890.1	0.925
4	19	-106181	212399.5	212553.4	212493	0.85
5	24	-104933	209913.2	210107.6	210031.3	0.803
6	29	-104274	208606.2	208841.1	208749	0.79
7	34	-103920	207908.8	208184.2	208076.2	0.783
8	39	-103762	207601.2	207917.1	207793.1	0.747
9	44	-96661.2	193410.4	193766.8	193627	0.736
10	49	-96513.3	193124.6	193521.5	193365.7	0.721
Non-OA						
1	4	-207493	414994.4	415026.8	415014.1	
2	9	-116016	232049.9	232122.8	232094.2	0.972
3	14	-105835	211698.2	211811.6	211767.1	0.931
4	19	-102747	205531.5	205685.4	205625	0.858
5	24	-101481	203010.4	203204.8	203128.6	0.828
6	29	-100790	201638.3	201873.2	201781.1	0.813
7	34	-100486	201040.6	201316	201207.9	0.817
8	39	-100281	200639.7	200955.6	200831.6	0.779
9	44	-93233.4	186554.8	186911.2	186771.4	0.77
10	49	-93048.5	186195	186591.9	186436.2	0.766

AIC- Akaike information criteria; BIC- Bayesian information criteria; aBIC- Sample size adjusted BIC; LL – Log Likelihood

Among people with OA, five clusters were identified to explain the trajectories. Most people (73.9%) continued to be in the healthy group with nearly zero comorbidities. Only 2.3% of the study population developed fewer than two comorbidities after 10 years from the index date. Another group constituting 11.6% subjects developed multimorbidity slowly after the index date, but the mean number of conditions was always less than two. Only two groups distinctively showed a multimorbidity trajectory with rapid onset (2.7%) or gradual onset (9.5%). The mean number of conditions after 20 years of the index date was 7 in the rapidly multimorbidity developing group and 4 in the group with gradual onset. *(*Figure 7.3-4)

In the non-OA group, the four cluster model was found to give the best fit for trajectory. Of these only one group had a very distinct path of developing multimorbidity named as gradual onset (4.6%). Nearly two thirds of the subjects were relatively healthy and 14.3% developed

multimorbidity after the index date but at a slower pace. Another group (5.1%) started developing comorbidities after 8 years of follow-up from the index date. (Figure 7.3-5)

Figure 7.3-4. Trajectories clusters of multimorbidity in people with OA without any comorbidities at index date

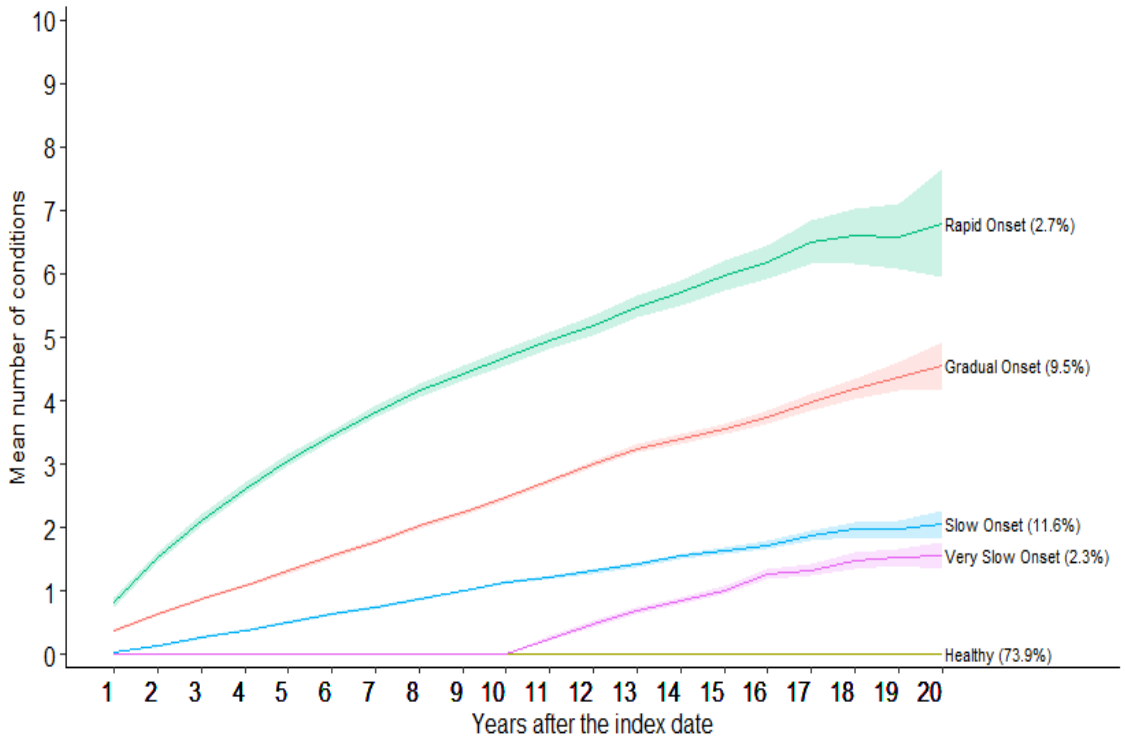
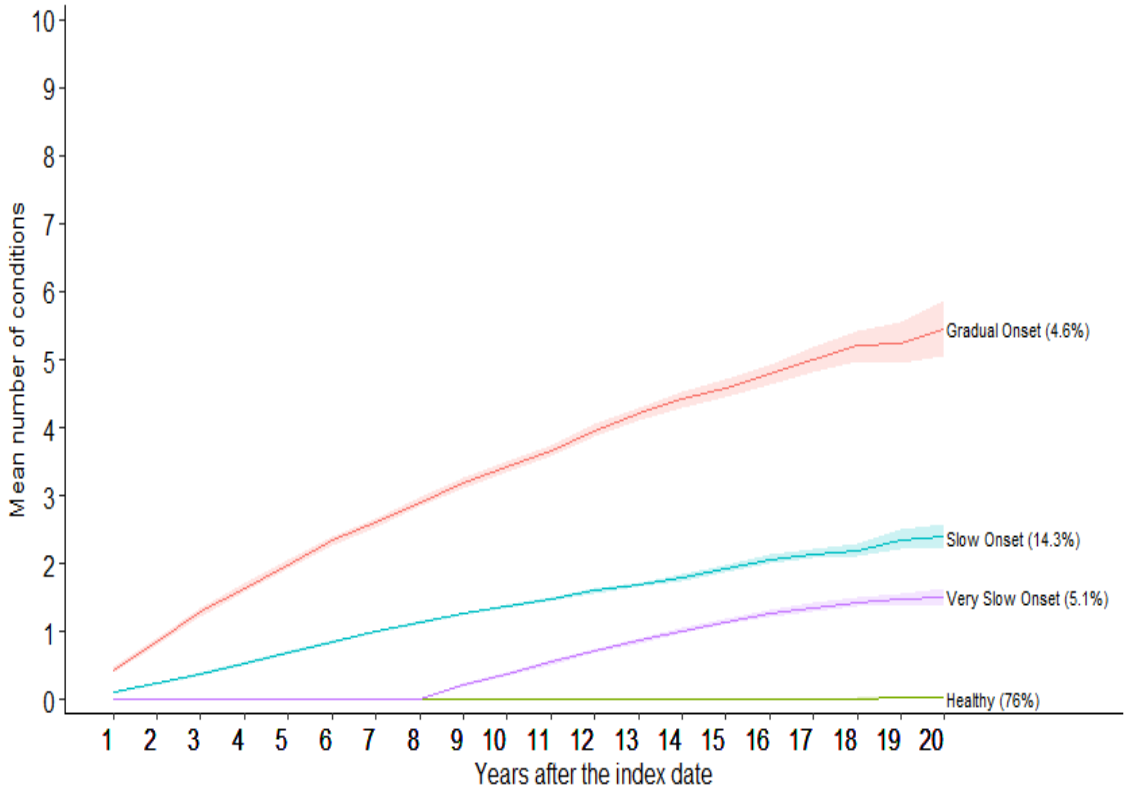


Figure 7.3-5 Trajectories clusters of multimorbidity in non- OA without any comorbidities at index date



7.4 Discussion

This study examined the multimorbidity trajectories of chronic diseases in the OA and non-OA groups over a time span of 20 years. Trajectories of developing multimorbidity has not been studied before in people with OA. Five clusters within the OA group and four within the non-OA group were identified using the LCGA method. Key findings from this study are: (1) the burden of multimorbidity was high in the OA group compared to the non-OA group; (2) a group of nearly 17.5% of people with OA accumulated multimorbidity rapidly and 28% had a low risk of rising multimorbidity over 20 years; (3) people who were obese, smokers and ex-drinkers at baseline had higher risks of rapidly developing multimorbidity compared to the relatively healthy group; and (4) within the non-OA group nearly 3% developed multimorbidity rapidly after the index date.

7.4.1 Trajectories in OA

To my knowledge this is the first study to explore the trajectories of multimorbidity within people with OA. The accumulation of multimorbidity depends on multiple factors such as age, chronic conditions present at the index date, lifestyle (e.g., smoking, diet and physical activity) and pharmacological effects. Very few studies have explored the trajectories of multimorbidity without any index date. However, these studies varied in their methods, study population, follow-up time span, nature of database and number of conditions studied. Jackson et al examined the factors associated with multimorbidity trajectory among middle-aged women in Australia and identified five clusters (Jackson *et al.*, 2015). Similarly, Strauss et al, using a UK primary care database (CiPCA) to explore the trajectories, also reported five clusters (Strauss *et al.*, 2014). Though these studies are not comparable with population of the current study, they provide some information on general population cohorts. Compared to 40% in the Strauss et al study, this study had only 28% in the very slow progress group in OA population. This group of people had the lowest percentage of obesity/overweight compared to other cohorts. As obesity is one of the biggest risk factors

for many comorbidities (Dhalwani N.N. *et al.*, 2017), the lower BMI in this group of people might have been keeping them healthy. Also, the prevalence of smoking was low in this group compared to others, which further supports the explanation of healthy lifestyle practice keeping them less prone to accumulating multimorbidity over time (Singer *et al.*, 2019).

Nearly 50% of the cohorts in both the groups had high multimorbidity count at the index date and continued to grow gradually. A high multimorbidity burden at the index date indirectly suggests the high GP consultations reflecting better health literacy and awareness (Cassell A. *et al.*, 2018). Patient education, counselling and supportive therapies might have moderated the high-risk factors at the index date allowing for slow building up of multimorbidity. Women had nearly 1.3 times higher risk compared to men to be in gradual progress group . People from this group (nearly 2/3rd were women) may have had better adherence to management of chronic conditions and drugs as they had higher consultation rates compared to men (Mukhtar *et al.*, 2018). Another crucial factor affecting the trajectory is the nature, rather than just the count of conditions at the index date. These gradual progress group had higher burden of all the conditions and nearly 50% of them had back pain at the index date and 40% had hypertension. Despite a high burden of chronic conditions in this gradual progress group, further study is needed to understand the further slow development of multimorbidity, needed for successful chronic are model.

Nearly 12% of people with OA (very rapid progress group) had less than two comorbidities at the index date which increased sharply to four after five years and to nearly 10 after 20 years. These represent a high risk group developing multimorbidity more rapidly than others. The distribution of population characteristics shows that the people in this group are older and have a higher prevalence of smoking, alcohol use and obesity compared to other groups. All these are well documented risk factors for multimorbidity (Dhalwani N.N. *et al.*, 2017). Jackson et al reported the association of obesity with longitudinal changes in multimorbidity (Jackson *et al.*, 2015). Another study from the UK among primary care patients also mentioned the strong association of all the above risk factors with

multimorbidity (Booth, 1994). Having fewer than two comorbidities at the age of 60s suggests two possibilities - either these are a comparatively healthy group, or they had fewer consultations leading to less diagnosis of conditions. However, the trajectory does not favour the first assumption of being healthy. Possibly, these group of people despite having high risk behaviours did not visit GPs often, but after the diagnosis of OA, there might have been increase in consultations leading to more diagnosis of comorbidities. Previously reported high consultation rates in the most deprived areas of UK echoes the findings from this study (Mukhtar *et al.*, 2018). Also, in this study people from this very rapid progress group had nearly equal proportions of CVD and MSK, which might have influenced the trajectories (Lappenschaar *et al.*, 2013). This provides with the opportunity to explore other associated factors on the illness course in people with OA. Another interesting finding from this study was the trajectory cluster of 5.7% people with OA developing multimorbidity rapidly. Again, the association with the least deprived group could have influenced the health seeking behaviour. No significant association was seen with gender and the most deprived group.

7.4.2 Trajectories in non-OA

In the non-OA group, four trajectory clusters were found. However, the burden of multimorbidity in each cluster in terms of mean number of comorbidities was less than in respective groups in the OA group. This supports the previous findings of a high burden of multimorbidity in people with OA. Only 3% of people without OA developed multimorbidity at a faster rate. The mean age in this group among non-OA was 68.5 years at the index date. Also, a strong association was found with other risk factors similar that to found in OA. The four trajectory groups in the non-OA group and the slower progression of multimorbidity in a similar population structure suggests a possible contribution of OA towards multimorbidity. The additional presence of OA might have accelerated the accumulation of multimorbidity due to the pathophysiology, shared risk factors, increased health care visits or due to the drugs used to manage chronic pain. Two studies that looked at the trajectory of

multimorbidity reported that the presence of chronic conditions at the index date led to evolution of multiple other related and non-related comorbidities during the life course (Hsu, 2015; Vos *et al.*, 2015).

For sensitivity analyses, people without any comorbidities at the index date in both the OA and non-OA groups were used. Even though there was the possibility of selecting a relative healthy group, a linear trajectory path was identified. Five clusters within the OA group and four within the non-OA group were estimated, with the burden of multimorbidity being higher in the OA group. The sensitivity analyses suggest that after the diagnosis of OA the chances of being diagnosed with other comorbidities increases. This could be because of the mentioned shared risk factors or differences in health care utilisation.

There is strong evidence from this study that in people with OA there are five distinct trajectory paths for multimorbidity. Longitudinal studies on multimorbidity report strong associations with poorer health outcomes, poor prognosis and early mortality (Wang *et al.*, 2009; Aarts *et al.*, 2012). This study is consistent with other findings in identifying different subgroups, despite the change in study population, methods and diseases included. Similar to this study, a birth cohort study has reported higher trajectories of multimorbidity among women and obese people (Canizares *et al.*, 2018). Another study used the LCGA method to explore trajectories of pain in a knee pain population reported worse outcomes in the group with comorbidities (Dowsey, Smith and Choong, 2015). Along with disease specific approaches, a broad system theory and non-specific approaches should be used to understand the different trajectories. Chronic health problems such as OA may affect multiple sites in the body due to a wide range of pathophysiological and mechanical factors.

Thus, clinicians need to recognise that people with OA experience different morbidities and accrual of comorbidities over time. It is not only the number of comorbidities but also the type of condition diagnosed that determines the trajectory, together with other socio-economic factors. Subgrouping the people with OA based on the trajectories provides methods for differentiating level of risks and designing different intervention approaches.

7.4.3 Strengths and limitations

Our study benefits from several strengths. To my knowledge this is first study to explore the multimorbidity trajectory in people with OA. The LCGA model overcomes the problems of cross-sectional analysis of multimorbidity, as reported before. Age was included as a random factor accounting for the variation in the age during the life course, which has not been considered in previous studies. A long follow-up of 20 years and consideration of 49 conditions adds to the strength of the study. Because of the longer follow-up more distinct patterns could be studied. Even though there were high attrition rates at each year, the 'lcmm' package of R could handle the missingness on the assumption of missing at random. Inclusion of baseline risk factors helped to establish the association with the trajectory groups, which can help with early diagnosis. The analysis examined non-linear patterns with age (linear, quadratic, and cubic) and the number of classes ranged from 2 to 10. Thus, the obtained model is the best model possible. The sensitivity analysis also reflected the same number of trajectory groups, reinforcing the validity of the findings.

There are some limitations to consider. The severity of the chronic conditions could not be included in the modelling, which might have influenced the trajectory groups. The other inherent limitations of database research, such as recording, or diagnosis biases of chronic conditions, could also have been influenced the results. Even though 49 chronic conditions were included it would be ideal to study an even larger number of conditions. Inclusion of the deprivation index limited the risk factor association analysis to England only, which could be expanded for other regions. Other caveats could be the healthcare utilisation pattern and health behaviour of individuals which might have delayed the diagnosis of chronic conditions. The attrition and deaths that occurred between the follow-up periods represent the impact of multimorbidity, however a complete case analysis would provide more strength.

7.4.4 Conclusion

The trajectories identified represent the high burden and diversity of multimorbidity after the index date in people with OA. This provides better understanding for care and illness pathways, which can be used for designing an appropriate care model. More detailed studies can be done among gradual onset multimorbidity cohort to recognize the success stories behind slowing down the further development of multimorbidity. Early identification of chronic conditions in the population with associated modifiable risk factors such as obesity and lifestyle may be able to prevent the future occurrence and worsening of multimorbidity. The time point of likely growth of multimorbidity can be set differently for different trajectory cohorts, in terms of screening and follow-up based on the rate of growth. Identifying the high-risk population can help the health system for effective resource allocation. Finally, a prediction model could be developed to predict the future risk of multimorbidity.

Summary of Chapter 7

Chapter 7 used LCGA to group people based on the trajectories of accumulation of number of chronic conditions with time in both people with OA and non-OA.

The key findings from the study are:

- Five groups in people with OA and four in people with non-OA were identified. Groups are based on the number of chronic conditions at the time of diagnosis and the rate of accumulation of additional comorbidities.
- A group of nearly 17.5% of people with OA with low multimorbid status at index date, accumulated multimorbidity rapidly and 28% had a low risk of rising multimorbidity over 20 years,
- People with obesity, smokers, and ex-drinkers at baseline had higher risks of rapidly developing multimorbidity compared to the relatively healthy group
- In non-OA group nearly 3% developed multimorbidity rapidly after the index date.

Chapter 7 tells us the rapid accumulation of chronic conditions in people with OA compared to non-OA. The burden of multiple chronic conditions is more in people with OA.

I was further interested to understand the outcome of being diagnosed with OA and the comorbidity clusters in both the groups. The outcomes of study interest are number of GP consultations recorded, hospital admission, all-cause mortality, and the loss in-terms of DALYs. Chapter 8 explores the association with these outcomes with OA and identified clusters in both the groups.

8 Chapter 8

Outcomes in OA and associated comorbidities

8.1 Introduction

Association of all-cause mortality with OA is nonconclusive (Hochberg, 2008). Studies have shown significant associations with cause-specific mortality such as CVD, and with all-cause mortality (Cleveland, Nelson and Callahan, 2019). The causes of death in people with OA depend on various associated factors and are less likely to be attributed to structural OA itself. Thus, understanding the mortality risk in OA would provide information on the burden or the risk of mortality rather than the association.

Even though it is assumed, that the people with OA would have increased health utilisation, this has not been studied in detail especially in the primary care population of the UK. Healthcare utilisation in primary care depends on a wide range of factors such as socioeconomic, demographic, accessibility, and availability. There are various ways to measure healthcare utilisation from a health system perspective. Two commonly used indicators are the number of hospital visits and number of inpatient admissions per person (Andersen and Newman, 1973).

According to the Global Burden of Diseases (GBD), hip and knee OA are the 11th highest contributor to global disability and the 38th highest in disability adjusted life years (DALYs). Increasing life expectancy and the ageing population are expected to make OA the fourth leading cause of disability by the year 2020 (Woolf and Pfleger, 2003). Years of life with disability (YLDs) for hip and knee OA increased by 6.6 million over the period 1990 to 2010 (10.5 million in 1990 to 17.1 million in 2010). The demerits of the GBD burden of disease evaluation that OA itself is not regarded as a valid clinical cause of death. The burden of disease measured so far is disease specific, rather than person-specific. In the context of

multimorbidity, it is important to evaluate the burden within a person as a whole considering all the diseases the person has.

Another important understudied aspect is the variation of disease pattern within people with OA. As seen in previous chapters, there are various groups within OA according to the multimorbidity clusters and they follow different illness pathways. So, exploring the health outcomes in these subgroups would provide more information towards person-centred care. Therefore, the current study explored all-cause mortality, GP visits, inpatient admission, and the burden of multimorbidity in people with OA and matched controls. Also, the distribution of above-mentioned health outcomes was studied within OA and matched controls subgroups.

8.2 Methods

8.2.1 Participants

The same identified group of OA cases and matched non-OA controls described in the previous Chapter 4, Section 4.3.2 for this analysis were used for the analysis.

8.2.2 Outcomes

- Average GP consultations per year

  This is defined as the average number of consultations per year recorded in the database for each person within the period of 1st Jan 1997 until the last record available for the person or the 31st of Dec 2017. The average was calculated by dividing the total number of GP consultations recorded by the number of years of registration in the database. The consultation includes visit to a GP or nurse or any other healthcare practitioner which has been recorded in the CPRD GOLD database for any purposes such as diagnosis, test or follow up.

  For example, if a person had a total of 15 years of registration with the database after 1997, and had 120 consultations recorded during that period, then the average number of consultations for that person is:

  = 120 /15 = 8 consultations per year

- Inpatient admission

  Information on inpatient admissions was obtained from the HES linkage data. The number of hospitalisations irrespective of any cause was used for the calculation of average hospitalisations per year. A similar formula to that used for calculation of GP consultations was used to estimate average inpatient admissions.

- Disability adjusted life years (DALY)

  For estimation of the burden of the comorbidities, the WHO proposed disability adjusted life years (DALY) method was used. However, interpretation of the burden of disease in terms of DALYs can be complicated if multiple conditions co-exist within individuals. A multiplicative methods was used for the estimation DALY in multiple conditions (Mathers, Iburg and Begg, 2006). This method has been used previously for calculating disease burden in comorbidity and multimorbidity (Hilderink *et al.*, 2016).

- All-cause mortality

  All-cause mortality data were obtained from the HES linkage data. The death date recorded in the database was used in the model to estimate the mortality risk in the OA group compared to that in the non-OA group.

8.2.3 Calculation of DALY for multiple chronic conditions

DALY is calculated by YLD+YLL

YLD (Years of living with disability)- Years lived with conditions x disability weight (Dw)

YLL (Years of life loss) – Years lost because of the condition (Life expectancy – age of death)

For independent comorbidities

For independent comorbidities, the probability of having two (comorbid) conditions is assumed to equal the product of the probabilities for having each of the diseases. (Mathers, Iburg and Begg, 2006)

Disability weight for multiple chronic conditions can be calculated by a multiplicative method

$$DW_{1+2} = 1 - (1-DW_1) \times (1-DW_2) \qquad \text{(equation 1)}$$

Where DW_1 is the Disability weight of the first chronic condition, and DW_2 is the disability weight for the second chronic condition

For Disability Adjusted Life Years (DALYs)-

$$DALY_{1+2} = 1- \{(1-DALY_1) \times (1-DALY_2)\} \qquad - \qquad \text{(equation 1)}$$

$DALY_1$ is the DALY for 1st condition

$DALY_2$ is the DALY for 2nd condition

In the presence of multiple conditions:

$$DALY_{total} = 1- \pi_i(1-DALY_i)$$

π is the product operator and i= 1 to n^{th} chronic condition

For dependent comorbidities (Mathers, Iburg and Begg, 2006)

In the presence of dependent comorbidities, the severity of conditions is shared with each other rather than being additive. For estimating the DALY for **dependent comorbidities**, the shared factor f_{1+2} can be calculated as below

$$f_{1+2} = DALY_{1+2} / (DALY_1 \times DALY_2) \qquad \text{- (equation 2)}$$

$DALY_{1+2}$ - calculated from equation 1

$DALY_1$- DALY for 1st condition

$DALY_2$ – DALY for 2nd condition

DALY for dependent comorbidities:

$$DALY_{1+2} = DALY_1 + DALY_2 - (f_{1+2} \times DALY_1 \times DALY_2)$$

f_{1+2} – as per equation 2

$DALY_1$- DALY for 1st condition

$DALY_2$ – DALY for 2nd condition

The disability weight was adopted from WHO, European version. (Appendix Table 40, page 350) For this study, dependent comorbidities assumption was made while estimating the DALY.

8.2.4 Covariates

All information available at the index date such as gender, age at index date, smoking, alcohol, and BMI was used in the analysis.

The Elixhauser comorbidity index (ECI) at baseline was calculated to estimate the burden of comorbidities. The ECI groups comorbidities of patients based on the international classification of diseases (ICD) diagnosis. Each comorbidity is categorised dichotomously as either present or not. The original index contained 30 comprehensive categories of comorbidity based on ICD-9-CM coding found in hospital abstracts data (Elixhauser et al., 1998), but later the list was expanded to 31 conditions and the scoring system was modified to reflect "the strength of each comorbidity group's independent association with hospital death." (van Walraven et al., 2009) It is reported that the Elixhauser comorbidity system can be condensed to a single numeric score that summarizes disease burden and is adequately discriminative for death in hospital. Details of the list of condition are given in Appendix Table 41 (page 351).

8.2.5 Statistical analysis

All the analyses were performed in two group of samples. Firstly, the outcomes were compared for the OA and the non-OA group. Secondly, the association of each outcome was explored within each group across the identified cluster .

Descriptive statistics of each outcome are reported as both mean (standard deviation) and median (inter quartile range). Normality distribution of outcomes was tested using histogram and shaprio-wilk test.

Firstly, the association of GP consultation rate with OA was assessed by linear regression and the residuals were checked for normality assumption. Because the outcome was continuous and skewed with non-zero distribution, a ‘gamma regression’ model with log link function was used. For inpatient admissions, because of the excessive ‘zeros’ the assumptions for a linear regression method could not be met. So, to account for excessive zeros and continuous data a two-part model was used. In the two-part model, a binary choice model is fit for probability of observing a positive-versus-zero outcome. Then, conditional on a positive outcome, an appropriate regression model is fit for the positive outcome. Logistic regression was used for the first part to compute the association for positive outcome, and linear regression in the second part to predict the association with increased hospitalisation. After the model, post model margin effects were estimated and compared with the observed estimates. The association with DALY was explored using a linear regression method. In the adjusted model covariates such as age, gender, smoking, alcohol, BMI and ECI were included.

For all-cause mortality, a cohort study design was used. The death after the index date was assessed. Both the OA and matched non-OA cohorts were followed for up to 20 years after the index date. For people with non-OA the start date was the assigned index date that of corresponding matched OA case. The follow-up period was until the earliest date of death, transfer out or end of the study (31st Dec 2017). The Kaplan-Meir method was used to

display the cumulative probability of death in people with incident OA and matched controls. HRs) and 95% CI were calculated adjusting for age, gender, BMI, smoking, alcohol use, and ECI at the index date. For association of death within each group identified clusters at the index date, at year 5, 10, 15 and 20 were considered as time varying covariates. In the adjusted model, gender, age, BMI, smoking, alcohol use and multimorbidity count at the index date were included. The interaction of OA with identified clusters for all-cause mortality was examined in cox model. Proportionality assumption for each comorbidity was examined with Schoenfeld residual tests. The statistical analyses were performed using STATA statistical software V.15 (STATA corp, Texas) and R software V3.5.

8.2.6 Sensitivity analysis

As a sensitivity analysis for the association with all-cause mortality, the analysis was re-run for people with OA and matched controls without any comorbidities before or on the index date. Time varying covariate Cox proportional hazard models were used to estimate the HR for all-cause mortality adjusted for, smoking, alcohol use, BMI and identified clusters at each five years interval.

8.3 Results

A total of 221,807 OA cases and 221,807 age, sex and practice matched non-OA controls were included in the analysis. Details of the description of the baseline characteristics in the two groups are provided in Chapter 4, Table 4.3-1(page 115).

8.3.1 Outcomes in OA

8.3.1.1 GP consultations for any purposes per year

The mean number of GP consultations in the OA group per year after 1st January 1997 was 19 compared to 15 in the non-OA group. However, the median difference between the groups was more with OA group had four more visits than non-OA. (Table 8.3-1)

Table 8.3-1. Summary of the GP consultations per year in the OA and non-OA groups

	OA (n=221,807)	**Non-OA (n=221,807)**
Mean, SD	19.40 (13.14)	15.31 (11.27)
Median, IQR	16.28 (10.27-25.05)	12.53 (7.42-20.11)

IQR-Inter quartile range; SD- Standard deviation

People with OA had 1.27 times more GP consultations without adjusting for other covariates. However, in the adjusted model that decreased to 1.16 (95%CI 1.15-1.17) times compared to the non-OA group. (Table 8.3-2)

Table 8.3-2. Gamma regression for association of GP consultations for any reasons with OA

	Unadjusted Incidence rate ratio 95% CI	Adjusted[#] Incidence rate ratio (IRR) 95% CI
Non-OA	Reference	Reference
OA	1.27(1.26-1.28) *	1.16(1.15-1.17) *

[#]Adjusted for smoking, alcohol, BMI, ECI and number of comorbidities at baseline

*P-value <0.05; IRR- Incidence rate ratio; CI-confidence interval

8.3.1.2 Inpatient admission per year

The mean number of hospitalisations per year was higher in the OA (0.25) compared to the non-OA group (0.15). Hospitalisation data was highly skewed towards the right side with a large proportion of zeros. Nearly 63% of all people with OA were not hospitalised at all compared to 66% in the non-OA group. (Table 8.3-3)

Table 8.3-3. Summary for the number of hospitalizations per year in the OA and non-OA groups

	OA (n=221807)	**Non-OA (n=221807)**
Mean, SD	0.25 (1.25)	0.15 (0.80)
Median, IQR	0 (0-0.25)	0 (0-0.13)

SD- Standard deviation; IQR- Inter quartile range

The two-part regression model showed a significant association of OA with hospitalisation. In the first part model, the odds of being hospitalised was 1.09 (95% CI 1.07-1.14) in OA compared with non-OA after adjusting for the covariates. In the second part of the model the

association of OA with numbers of hospitalisations was 1.16 (95% CI 1.14-1.17) times more compared to non-OA controls. (Table 8.3-4)

Table 8.3-4. Two-part regression model for association of number of hospitalisations per year in the OA and non-OA groups

	Unadjusted Exp (B) 95% CI	Adjusted Exp (B) 95% CI
First part model		
Non-OA	Reference	Reference
OA	1.11(1.10-1.12) *	1.09 (1.07-1.14) *
Second part model		
Non-OA	Reference	Reference
OA	1.16 (1.14-1.17) *	1.13 (1.11-1.15) *

#Adjusted for smoking, alcohol, BMI, ECI and number of comorbidities at baseline; *P-value <0.05

8.3.1.3 DALYs

DALY was calculated for each group accounting for the multiple comorbidities. On average a person with OA loses 13.23 years of life due to the comorbidities compared to 11.51 years in non-OA controls. The median number of years lost was 2 years more in people with OA than in non-OA controls. (Table 8.3-5)

Table 8.3-5. Disability adjusted life years (DALYs) in the OA and non-OA groups

DALY (in years)	**OA (n=221807)**	**Non-OA (n=221807)**
Mean, SD	13.23 (12.76)	11.51 (12.01)
Median, IQR	10.01 (3.43-19.39)	8.09 (2.13-17.04)

IQR- Interquartile range; SD- Standard deviation

Linear regression model showed that having OA increased the DALY by 3.25 (95% CI 3.02-3.49) years compared to non-OA after adjusting for other covariates. (Table 8.3-6)

Table 8.3-6. Linear regression for DALY for association with OA

	Unadjusted Exp (B) 95% CI	Adjusted Exp (B) 95% CI
Non-OA	Reference	Reference
OA	5.62(5.22-6.04)	3.25(3.02-3.49)

#Adjusted for smoking, alcohol, BMI, ECI and number of comorbidities at baseline; *P-value <0.05

8.3.1.4 All-cause mortality

Of those with OA, 20,617 (9.3%) died during the study period, compared with 13,087 (5.9%) in those who did not have OA. The mortality rate was nearly two times higher in the OA group (13.52 per 1000 person-years compared to 7.14 per 1000 person-years in the non-OA group). The unadjusted HR was 2.02 (95% CI 1.98 - 2.06) which reduced to 1.89 (95% CI 1.85-1.93) after adjustment for other covariates including multimorbidity and ECI. (Table 8.3-7)

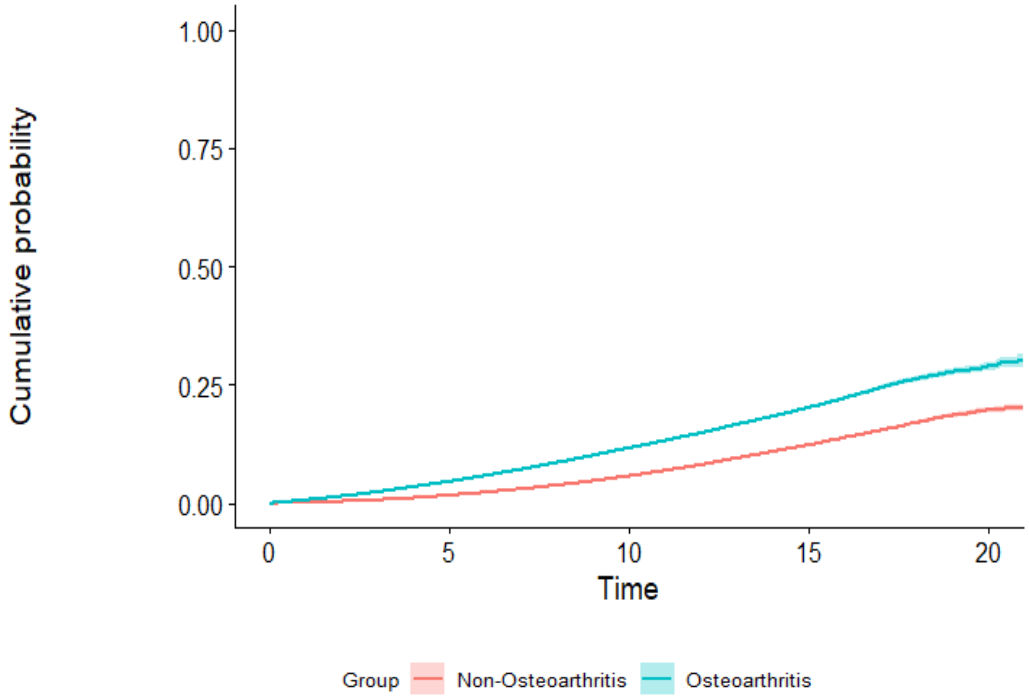
Table 8.3-7. All-cause mortality in the OA and non-OA

	Total deaths	Incidence per 1000 person-years (95% CI)	Unadjusted HR (95% CI)	Adjusted HR[#] (95% CI)
Non-OA (n=221807)	13087	7.14 (7.02-7.27)	Reference	Reference
OA (n=221807)	20617	13.52 (13.34-13.70)	2.02 (1.98-2.06) *	1.89 (1.85-1.93) *

[#]Adjusted for smoking, alcohol, BMI, ECI and number of comorbidities at baseline; *P-value <0.05; HR- Hazard ratio

The cumulative probability of death at 5 years was 6% in OA compared to 2% in the non-OA group compared to. This increased to 30% in the OA group after 20 years of follow-up compared to 20% in the non-OA group. (Figure 8.3-1)

Figure 8.3-1. Cumulative probabilities of all-cause mortality in the OA and non-OA groups



8.3.1.4.1 Sensitivity analysis

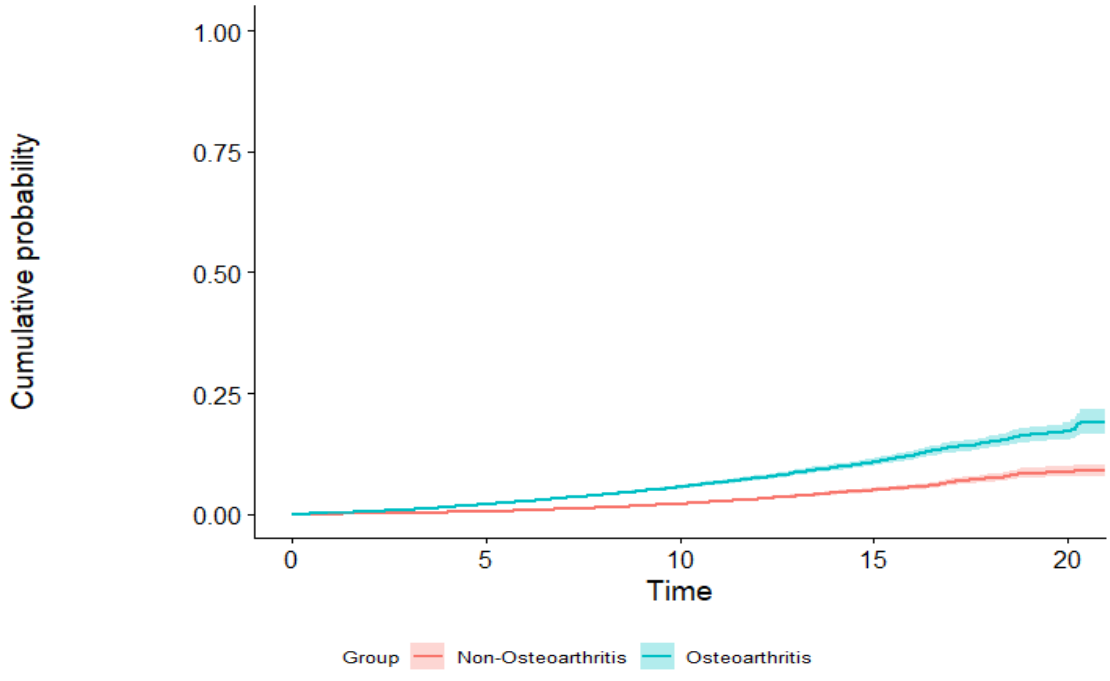
In the cohorts without any comorbidities at the index date, the mortality rate among people with OA was 6.26 per 1000 person-years compared to 2.99 in non-OA controls. The unadjusted HR was 2.22 (95% CI 2.02 - 2.44) which reduced to 2.15 (95% CI 2.00-2.43) after adjustment for other covariates including multimorbidity and ECI. (Table 8.3-8) The cumulative probability of death at year 5 was 1% in the non-OA group compared to 2% in the OA group. This increased to 9% after 20 years of follow-up in the non-OA group compared to 16% in the OA group. (Figure 8.3-2)

Table 8.3-8. All-cause mortality in people with OA and non-OA controls without any comorbidities at index date (sensitivity analysis)

	Total deaths	Incidence per 1000 person-years (95% CI)	Unadjusted HR (95% CI)	Adjusted HR[#] (95% CI)
Non-OA(n=22333)	653	2.99 (2.77-3.23)	Reference	Reference
OA (n=22333)	1197	6.26 (5.92-6.63)	2.22 (2.02-2.44) *	2.15 (2.00-2.43) *

[#]Adjusted for smoking, alcohol, and BMI; *P-value <0.05

Figure 8.3-2. Cumulative probability of all-cause mortality in people with OA and non-OA control without any comorbidities at index date



8.3.2 Outcomes within OA across the clusters

8.3.2.1 GP Consultations per year

Within OA clusters the mean number of GP consultations for any reasons varied from 15.35 per year in relative healthy cluster to 34.36 per year in CV-MSK cluster. The same pattern was observed with the median number of visits across the clusters. In the non-OA group clusters, the lowest mean number of outpatient visits was reported in relative healthy group with 11.99 per year and the highest was 29.50 per year in CV-MSK cluster. The same pattern was seen for median number of visits. The Kruskal-Wallis test detected statistically significant differences in median number of visits across the clusters in both OA and non-OA with p value <0.05. (Table 8.3-9)

Table 8.3-9. GP consultations per year within OA and non-OA clusters

	Mean (SD)	Median (IQR)
OA		
Relatively Healthy (n=118220)	15.35 (10.72)	12.68 (8.11-19.53)
CV-MSK (n=11234)	34.36 (17.09)	31.22 (22.47-42.26)
Thyroid (n=7049)	23.09 (14.24)	19.87 (13.56-28.96)
CV (n=47976)	23.37 (13.03)	20.59 (14.32-29.41)
MSK-MH (n=37311)	21.98 (12.99)	19.13 (13.08-27.65)
Non-OA		
Relatively Healthy (n=129891)	11.99 (9.14)	9.68 (5.77-15.50)
CV-MSK (n=9007)	29.50 (14.57)	26.73 (19.19-36.50)
Thyroid (n=6183)	18.87 (11.97)	16.06 (10.94-23.67)
CV (n=42793)	20.50 (11.72)	18.11 (12.29-26.02)
MSK-MH (n=31553)	17.20 (10.95)	14.69 (9.77-21.70)

IQR- Interquartile range; SD- Standard deviation; CV- Cardiovascular; MH- Mental health; MSK- Musculoskeletal

8.3.2.2 Inpatient admission

The mean number of inpatient admissions per year was highest for CV-MSK in the OA group, followed by CV cluster. On average a person from CV-MSK cluster was hospitalised 0.43 times in a year compared to 0.31 times in CV cluster. In the non-OA group, the mean

number of hospitalisations was lower compared with the OA group, the mean number of hospitalisations in CV-MSK being 0.34 per year followed by 0.20 in CV cluster. The Kruskal-Wallis test detected statistically significant differences in the median number of hospitalisations across the clusters in both OA and non-OA with p value <0.05. (Table 8.3-10)

Table 8.3-10 Mean inpatient admission within OA and non-OA clusters

	Mean (SD)	Median (IQR)
OA		
Relatively Healthy (n=118220)	0.20 (0.74)	0 (0-0.19)
CV-MSK (n=11234)	0.43 (2.57)	0 (0-0.30)
Thyroid (n=7049)	0.27 (0.79)	0 (0-0.55)
CV (n=47976)	0.31 (1.96)	0 (0-0.33)
MSK-MH (n=37311)	0.27 (0.72)	0 (0-0.31)
Non-OA		
Relatively Healthy (n=132,183)	0.12 (0.59)	0 (0-0.10)
CV-MSK (n=9019)	0.34 (1.81)	0 (0-0.40)
Thyroid (n=6190)	0.18 (0.45)	0 (0-0.19)
CV (n=42841)	0.20 (1.14)	0 (0-0.20)
MSK-MH (n=31574)	0.17 (0.53)	0 (0-0.18)

IQR- Interquartile range; SD- Standard deviation; CV- Cardiovascular; MH- Mental health; MSK- Musculoskeletal

8.3.2.3 DALYs across the clusters

The mean DALY was calculated across the clusters in the OA and non-OA groups. As the number and type of comorbidity clusters are associated with DALY, no further analysis was done to examine the association.

In the OA group, CV-MSK cluster had a mean DALY of 33.53 years compared to 21.42 years in metabolic cluster and 20.60 years in MSK-MH. This means that people in CV-MSK on average lose 33.53 years of their lives due to the multiple comorbidities. Nearly 20 years of life is lost due to comorbidities in the remaining clusters. (Table 8.3-11)

Table 8.3-11 DALYs across the clusters in OA and non-OA

	Mean (SD)	Median (IQR)
OA		
Relatively Healthy (n=118,279)	6.41 (7.24)	4.32 (0.57-9.60)
CV-MSK (n=11,235)	33.53 (16.01)	31.02 (21.87-42.66)
Thyroid (n=7049)	21.42 (15.20)	17.91 (10.27-28.74)
CV (n=47,976)	18.38 (11.14)	16.35 (10.23-24.29)
MSK-MH (n=37,312)	20.60 (12.21)	18.30 (11.73-26.97)
Non-OA		
Relatively Healthy (n=132,183)	5.73 (6.87)	3.64 (0.05-8.45)
CV-MSK (n=6190)	20.68 (15.53)	16.75 (9.15-28.40)
Thyroid (n=9019)	32.83 (15.73)	30.54 (21.32-41.63)
CV (n=42,841)	18.02 (10.88)	16.07 (10.18-23.84)
MSK-MH (n=31,574)	18.95 (11.78)	16.69 (10.42-24.99)

IQR- Interquartile range; SD- Standard deviation; CV- Cardiovascular; MH- Mental health; MSK- Musculoskeletal

A similar pattern was seen in the non-OA group. CV-MSK had the highest DALY compared to other clusters. Clusters in the OA group had slightly higher DALYs compared to the non-OA group for each cluster. The Kruskal-Wallis test detected a statistically significant difference between the median number of DALYs across the groups in both OA and non-OA with p value <0.05.

8.3.2.4 All-cause mortality

People with OA and non-OA were grouped into five clusters based on the LCA (Chapter 5, Figure 5.3-3, and Figure 5.3-6). Reported mortality across the OA group showed CV-MSK cluster had 19.63% of deaths followed by 14.96% in CV cluster. The all-cause mortality rate was 42.2 per 1000 person-years in CV-MSK followed by 24.9 in CV. (Table 8.3-12)

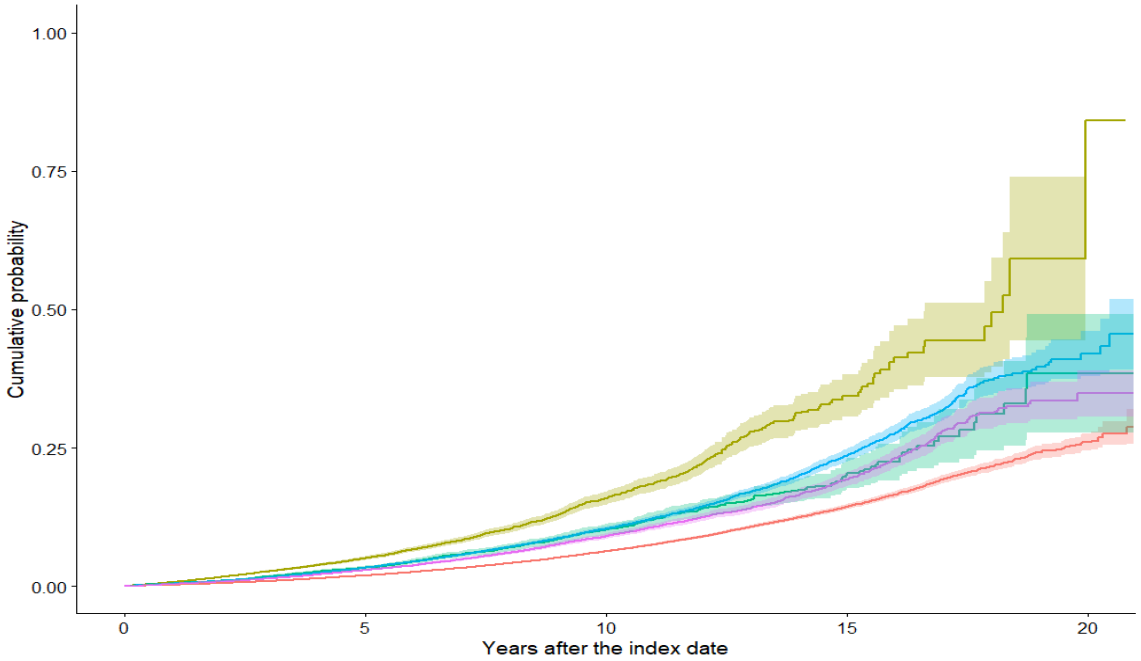
Table 8.3-12. All-cause mortality across the clusters within the OA group

	Total death N (%)	Mortality rate per 1000 person-years	Crude HR (95% CI)	Adjusted[#] HR (95% CI)
Healthy (n=118279)	8100 (6.84)	8.9(8.7-9.1)	Reference	Reference
CV-MSK (n=11235)	2206 (19.63)	42.2(40.5-44.0)	2.08(1.88-2.30) *	1.82(1.65-2.02) *
Thyroid (n=7049)	674 (9.56)	15.8(14.7-17.1)	1.53(1.35-1.74) *	1.52(1.33-1.72) *
CV (n=47976)	7181 (14.96)	24.9(24.3-25.5)	1.71(1.56-1.89) *	1.53(1.39-1.68) *
MSK-MH (n=37312)	2456 (6.58)	10.4(9.90-10.8)	1.27(1.13-1.42) *	1.23(1.09-1.38) *

[#]Adjusted for smoking, alcohol, BMI, ECI and number of comorbidities at baseline; HR- Hazard ratio; CI-Confidence interval; *P-value <0.05

The mortality rate in OA was highest for CV-MSK cluster aHR 1.82 (95% CI 1.65-2.02), followed by 1.53 (95% CI 1.39-1.68) and 1.52 (95% CI 1.33-1.72) in CV and metabolic clusters respectively, after adjusting for other covariates and time varying aspect of the clusters change. Figure 8.3-3 depicts the cumulative probabilities of all-cause mortality across the clusters in OA and shows that CV-MSK had consistently higher probabilities of death over the time compared to other clusters.

Figure 8.3-3. Cumulative probability of death across the clusters in the OA group



Red- Relative healthy; Yellow- CV-MSK cluster; Green- Metabolic cluster; Blue- CV cluster, Purple- MSK-MH cluster

Among non-OA people five clusters were identified as described in Chapter 5, Figure 5.3-6. The highest proportion of deaths was observed in CV-MSK cluster (14.58%) and CV cluster (10.39%). The mortality rate was 25.8 per 1000 person-years in CV-MSK which was two times higher than the mortality rate in CV cluster (13.6 per 1000 person-years). The lowest rate was seen in relatively healthy at 4.8 per 1000 person-years. (Table 8.3-13)

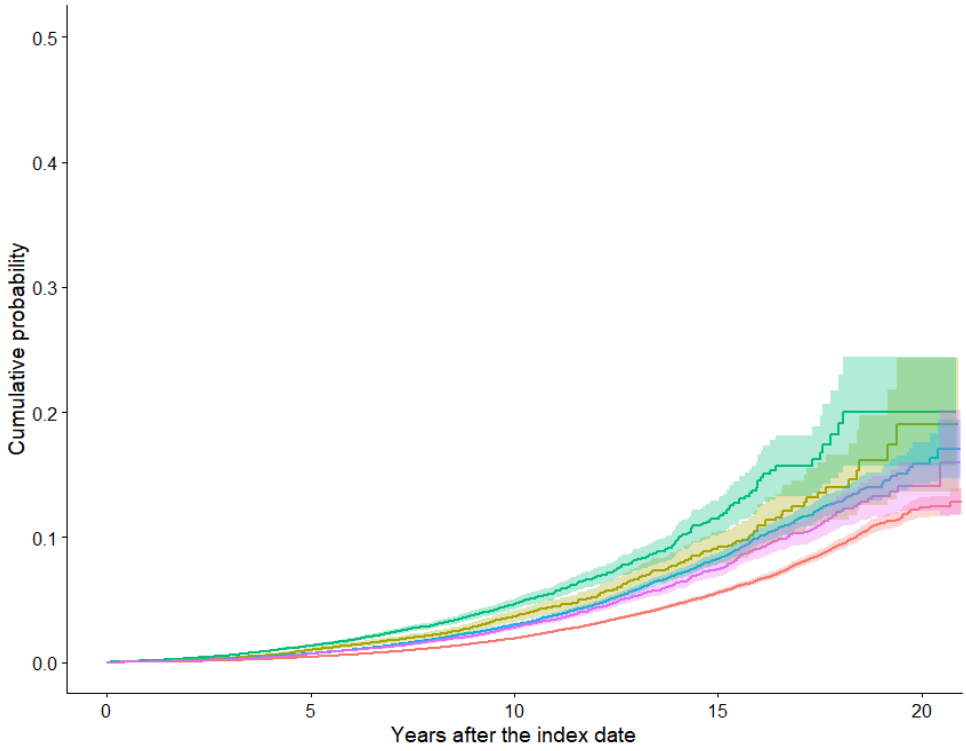
Table 8.3-13. All-cause mortality across the clusters in the non-OA group

	Total death	Mortality rate per 1000 person-years	Crude HR (95% CI)	Adjusted[#] HR (95% CI)
Healthy (n=132,183)	5739 (4.34)	4.8(4.7-4.9)	Reference	Reference
Thyroid (n=6190)	425 (6.86)	9.4(8.5-10.3)	1.78(1.45-2.19) *	1.69(1.37-2.08) *
CV-MSK (n=9019)	1315 (14.58)	25.8(24.5-27.3)	3.14(2.60-3.77) *	2.51(2.08-3.03) *
CV (n=41841)	4348 (10.39)	13.6(13.2-14.1)	2.49(2.07-2.99) *	2.05(1.70-2.47) *
MSK (n=31574)	1260 (3.99)	5.4(5.1-5.7)	1.85(1.48-2.32) *	1.63(1.29-2.06) *

[#]Adjusted for smoking, alcohol, BMI, ECI and number of comorbidities at baseline; HR- Hazard ratio; CI-Confidence interval; *P-value <0.05

The adjusted hazard ratio for all-cause mortality in CV-MSK cluster compared with the healthy cluster was 2.51 (95% CI 2.08-3.03), followed by 2.05 (95% CI 1.70-2.47) in CV cluster, after adjusting for other covariates and change in clusters over time. (Table 8.3-13) Figure 8.3-4 shows that the cumulative probability of death was highest for CV-MSK followed by metabolic clusters.

Figure 8.3-4. Cumulative probability of death across clusters in the non-OA group



Red- Relative healthy; Green- CV-MSK cluster; Yellow- Metabolic cluster; Blue- CV cluster, Purple- MSK-MH cluster. CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

8.3.2.4.1 Interaction between OA and clusters

The interaction between OA and the identified clusters were tested and used in the cox hazard model to identify the association with all-cause mortality. (Table 8.3-14)

Combination of OA with identified clusters had higher mortality probabilities compared to non-OA cluster groups. People with OA belong to CV cluster had four times higher chances of death (aHR 4.09 95% CI 3.91-4.26), followed by people with OA and CV-MSK cluster with hazard ratio of 3.29 (95% CI 3.03-3.56). The observed mortality risk was high in combination with OA compared to non-OA clusters. Non-OA group with MSK cluster had 20% less risk of death (aHR 0.80 95% CI 0.75-0.88) compared to non-OA healthy group. (Table 8.3-14)

Table 8.3-14. Contribution of interaction between OA and identified cluster on mortality

	Crude HR (95% CI)	Adjusted[#] HR (95% CI)
Non-OA*Healthy (n=132,183)	Reference	Reference
Non-OA*CV-MSK (n=9019)	6.76 (6.36-7.17)*	1.98 (1.82-2.15)*
Non-OA*thyroid (n=6190)	2.14 (1.94-2.36)*	1.40 (1.26-1.56)*
Non-OA*CV (n=42841)	3.14 (3.01-3.26)*	2.10 (2.01-2.20)*
Non-OA*MSK-MH (n=31574)	1.23 (1.16-1.31)*	0.80 (0.75-0.86)*
OA*Healthy (n=118,279)	1.98 (1.92-2.05)*	2.05 (1.98-2.12)*
OA*CV-MSK (n=11235)	11.91 (11.33-12.52)*	3.29 (3.03-3.56)*
OA*thyroid (n=7049)	3.91 (3.61-4.24)*	2.45 (2.23-2.67)*
OA*CV (n=47976)	6.24 (6.03-6.46)*	4.09 (3.91-4.26)*
OA*MSK-MH (n=37312)	2.49 (2.39-2.61)*	1.59 (1.50-1.68)*

#Adjusted for smoking, alcohol, BMI, ECI and number of comorbidities at baseline; HR- hazard ratio, CI-confidence interval. CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

8.4 Discussion

Outcomes in OA are variable and have been well studied with respect to pain progression, activity limitations, and functional decline, participation restriction and quality of life. However, health utilisation outcomes and the burden of the disease have not been well studied, especially in the presence of multimorbidity. Four outcomes were studied namely: GP consultations, hospitalisations (inpatient admission), disease burden measured as DALY and all cause morality. Firstly, these outcomes were explored for OA cases compared to controls, and then across the clusters identified within the separate OA and non-OA control groups. In the OA group, all the studied outcomes were higher than in non-OA controls, as were the studied associations. People with OA had: (1) 1.2 times more GP consultations; (2) 1.1 times higher risk of hospitalisation (3) nearly 1.9 times higher risk of all-cause mortality; (4) 3.2 years increased DALY compared to non-OA. Within the OA group, people with the cluster of conditions led by back pain and hypertension had the highest number of outpatient visits, hospital admissions and DALYs and the risk of mortality was nearly 1.82 times higher compared to the relatively healthy group; (5) clusters of CV and CV-MSK in OA had higher risk mortality compared to non-OA group. A similar pattern

was seen in the non-OA control population, with the cluster of people having hypertension and back pain having the worst outcomes.

8.4.1 GP consultations

One of the major outcomes in OA is the number of consultations in primary care. An increased consultations per year in the OA group was seen compared to the non-OA controls, as well as the higher associations. Multiple factors could influence the number of GP consultations in people with OA. Bedson et al reported that joint pain and the severity of pain are the strongest factors for increased consultations (Bedson *et al.*, 2007). However, they did not find any difference in the median number of consultations for comorbidities between knee pain consulters and knee non-consulters. In the analysis, adjustment was done for the number of comorbidities and the severity of comorbidities measured as ECI. There are not enough studies on the consultation rates of different conditions in the UK primary care setting. According to one report, nearly one third of people consulted GPs for musculoskeletal (MSK) problems (Versus Arthritis, 2009). The results of this study confirm the higher consultation rates in people with OA after the diagnosis. It is possible to have multiple consultations after the first recording in the GP database to confirm the diagnosis through tests such as radiographs, but it seems likely that most subsequent visits are for management. The increased consultations support the high burden of MSK conditions in primary care, which merits more detailed investigation and comparison with other chronic conditions.

Within OA, increased consultation rate was seen for CV-MSK cluster followed by the CV cluster only. Even though comorbidity count and severity were adjusted, the increased consultation rates in these groups suggests the compounding effect of comorbidity. For example, the adjustment for number of comorbidities is not sufficient to address the severity of pain related diseases, one of the factors for frequent GP visits. Zhu et al reported similar findings with higher consultation rates among multimorbidity clusters with pain related conditions (Zhu *et al.*, 2020). The increased consultation rates found in clusters with CVD

and depression could be due to the emphasised targeting of these conditions in the QOF guidelines for UK National Health Services (Forbes *et al.*, 2017). The cluster specific consultation rate pattern in the OA group was similar to that found in the control group, which suggests that the combination of both MSK, and CVD increases the number of healthcare visits and that this becomes even more frequent in people with OA. The causality association between the number of conditions and GP visits is difficult to establish in this study. However, the identified clusters with multiple conditions clearly increase the burden to primary care.

8.4.2 Hospitalisations

Another important outcome was hospitalisations for any cause in OA cases compared to non-OA controls. Using HES data the average number of hospitalisations per year was higher in people with OA. After adjusting for other covariates, comorbidities and ECI, firstly the risk of being hospitalised remained high and then the risk of hospitalisation was nearly 1.1 times higher than in controls. The severity of the conditions was adjusted using the ECI index, thus the excess hospitalisation could be related to OA related conditions. Morgan et al found an increased trend of hospitalisations due to OA in the UK (Morgan *et al.*, 2019). The excess hospitalisation in OA could be due to the increased risk of falls and injury (Dore A *et al.*, 2013) and to requirement for joint replacement, especially of the knee (Culliford D *et al.*, 2015; Ackerman *et al.*, 2019, Dixon, 2004; Ibrahim *et al.*, 2010). Another important associated factor could be the use of NSAIDs in OA (Tramèr *et al.*, 2000). Studies have shown that long-term use of NSAIDs causes increased hospitalisations especially due to gastrointestinal bleeding (Henry *et al.*, 1996; MacDonald *et al.*, 1997) and CVD (Jüni *et al.*, 2004; Hippisley-Cox and Coupland, 2005).

Within both the OA and non-OA group, the mean number of hospitalisations was highest for the clusters with hypertension and back pain followed by the cluster led by hypertension only. This is comparable to the results of a previous study (Zhu *et al.*, 2020). The presence of CVD increases the chances of inpatient admission due to known outcomes such as

myocardial infarction, arrhythmias, and heart failure. In the OA group, the average admission rate was higher than in the non-OA controls. The association of OA with CVD might in part explain these increased hospitalisations (British Heart Foundation, 2019). Also, the mean number of conditions in these clusters was higher compared to other clusters. Cassell et reported in the UK that nearly 56% of the hospitalisations were due to multimorbidity (Cassell A. *et al.*, 2018). Increased inpatient admissions in these clusters indicates the severity and risk of the complex multimorbidity. In both the OA and non-OA groups the average hospitalisation for pain related clusters and depression and/or thyroid clusters were analogous. Further studies are required to understand the pattern of hospitalisation after adjusting for medication use and healthcare access.

8.4.3 All-cause mortality

In this large and matched cohort study, people with OA had an excess all-cause mortality compared with non-OA controls. The association of all-cause mortality with OA is inconclusive (Cleveland, Nelson and Callahan, 2019). Both significant and non-significant associations have been reported previously. A systematic review in 2008 reported a moderate increase in mortality among people with OA (Hochberg, 2008), but a primary care database study from Sweden found no significant association of excess all-cause mortality in OA (Turkiewicz *et al.*, 2016). However, findings from this study is similar to that reported by Nuesch et al in a UK GP database (Nuesch *et al.*, 2011). Reasons for the discordant findings may be due to methodological differences, including the definition of OA, age range, study design, length of follow-up, and whether variables that can change and develop over time, such as measures of OA,BMI and comorbidities were accounted for. In this study, comorbidity counts and the severity of comorbidities at the index date along with other socio-economic variables at the index date were adjusted in the models. The risk of mortality in people with OA was also higher in the sensitivity analysis. Because in the sensitivity analysis people at the index date did not have any comorbidities at baseline, the increased mortality

risk could be due to the subsequent higher comorbidity incidence in the OA group after the index date.

Within the OA and non-OA groups, the mortality rate in this study was highest for the complex cluster led by hypertension and back pain, followed by the cluster led by hypertension only, and clusters led by depression. Previous studies have reported excess mortality related to CVD in people with OA, which accords with results of this study (Cleveland, Nelson and Callahan, 2019). There is a lack of studies on the association of depression or psychological disease related mortality in OA. This study found that the all-cause mortality rate in people with OA and depression and/or thyroid conditions is quite high, and a similar pattern of mortality was seen across clusters in the non-OA group. The reported higher risks across clusters within non-OA controls compared to that of corresponding clusters within the OA group could be due to the relatively healthy control population. Because the estimated HR is a relative risk compared to the reference group, the people within that group in the non-OA population appear to be healthier compared to those in the OA group. Cause-specific mortality in OA was not investigated, but the excess mortality in each identified cluster can be a surrogate indicator of this. The lowest mortality rates were seen among people with chronic pain after relative healthy group, even though the association was significant in both the groups. A recently published systematic review on chronic pain and mortality found no significant association and reported wide heterogeneity across the studies (Smith *et al.*, 2014).

Studies on multimorbidity clusters and mortality have consistently reported a higher association with CVD clusters (Haug *et al.*, 2020). A 15-year longitudinal study reported higher mortality (nearly two times) in clusters led by MSK and CVD (Willadsen *et al.*, 2018) and another study in an elderly population reported higher mortality among people with arthritis and depression (Teh *et al.*, 2018). Similar findings have been reported by other studies, including Zhu et al who found an increased proportion of all-cause mortality among clusters with painful conditions and depression (Zhu *et al.*, 2020).

Various reasons, apart from comorbidities, have been shown to help explain the higher mortality in people with OA. For example, obesity, pain, and disability or functional limitations. These factors limit physical activity levels in people with OA, thus reducing the protection against CVD. Nuesch et al reported that walking disability was associated with increased mortality in people with OA (Nuesch *et al.*, 2011). Another possible explanation could be chronic subclinical inflammation which may predispose to an increase in comorbidities, especially CVD and degenerative diseases (Couzin-Frankel, 2010). Also, the use of analgesics such as NSAIDs in OA increases the risk of CVD comorbidities (Trelle *et al.*, 2011). Even though cause-specific mortality was not part of the current research, further studies to explore the pattern of mortality within the clusters, especially with depression and chronic pain appear warranted.

8.4.4 DALYs

Measuring DALYs in multimorbidity is an active medical research field (Hilderink *et al.*, 2016; Boshuizen *et al.*, 2017). The latest available method was used for estimation of DALYs accounting for the comorbidities in the OA and non-OA groups. People with OA lose an average of 13 years of productive life due to comorbidities. The burden of comorbidities was higher in the OA compared to non-OA group. Losina et al reported a loss of 12 years productive life due to OA (without accounting for mortality) (Losina E *et al.*, 2011).

In both the groups the DALYs was highest for the cluster led by hypertension and back pain, followed by pain clusters and depression. The increased DALY in the complex cluster of MSK and CVD could be due to the high death rate among people in this cluster. Whereas the high burden in pain clusters, despite low death rates, suggests that the burden of chronic pain in an individual in the presence of other conditions is a key reason for high DALYs. In these clusters chronic pain played a major role as the leading chronic condition, especially back pain. The Global burden of disease study reported the high burden of low back pain and highest DALY in Western Europe (Hoy *et al.*, 2014; Blyth *et al.*, 2019).. People with depression have been found to lose nearly 28 good quality years of their life (including loss

due to suicide) (Jia *et al.*, 2015). There was higher clustering of depression at a younger age and the DALY was nearly 16 years in both groups. Such loss of productive live due to multiple chronic conditions suggest the importance of prevention and management of multiple conditions in people with OA and controls. A paradigm shift is needed towards pain comorbidity, especially with coexisting conditions. Data from this study indicate that public health and policy measures aimed at decreasing the multimorbidity and disability in OA may have the potential to produce large health gains.

8.4.5 Strengths and limitations

This study has several strengths. This is the first study to provide information on health utilisation and burden in people with OA compared to age, sex matched controls in the UK. Also, the association of these outcomes within OA and controls were examined with respect to multimorbidity clusters. One of the major strengths is the inclusion of multimorbidity for estimation of all-cause mortality and the burden of diseases in OA. The study population had a long follow-up time of maximum 20 years and considered 49 chronic conditions during the modelling. The time varying analysis accounted for the change in clusters and incidence of new conditions after the index disease, making the model more powerful. Another important strength is consideration of multimorbidity in the estimation of DALYs, which gives a true picture of the burden in a person rather than the disease.

There are several limitations to this study. Because focus of the study was on the burden of OA and the identified clusters, perform joint-specific OA analysis were not done. Findings are likely to differ in joint specific analysis, as discussed before. The conditions and the clusters identified are linked to the recording pattern in the GP database. Therefore, ascertainment biases due to misdiagnosis, miscoding and delayed recording may all be present. Only all-cause mortality was estimated, whereas cause-specific mortality might provide better information about the clusters. The inpatient admission data was gathered from the HES linkage, which was not available for all people. For DALY, the used disability weights can vary as per the UK population and assumed the comorbidities are dependent on

each other. However, in the absence of these estimates for the UK, the best possible matched available information was used as per the WHO guidelines. The hospital admissions and GP visits were calculated irrespective of any specific cause, which could have been influenced by other incidence of other conditions. The consultation definition in this may not be accurate as it includes recording of visit to the primary care for any purposes. Possibly, people who visited more frequent had more chance of being diagnosed with multiple chronic conditions and vice versa. This is difficult to establish through the study. However, both the number and severity of the chronic conditions were adjusted at the index date. The change in lifestyle health behaviour pattern and drug use for the reported outcomes were not considered, which might have confounded the association with OA and/or the clusters within them.

8.4.6 Conclusion

The burden of health utilisation, disease severity and mortality were high in the people with OA. Within them, people with clusters of conditions led by hypertension and back pain, and by depression, had a higher burden compared to others. The findings from this study strengthen the importance of identified clusters in people both with and without OA. Coexistence of MSK and CVD leads to the worst health outcomes and incurs an increased burden on the health system. It was found that clusters with pain and depression also have high healthcare utilisation and more loss of years of productive life, which needs further investigation. Further research is warranted to better understand the causes and patterns in detail. Further studies could also be done to calculate the economic losses incurred among individuals in each cluster.

Summary of Chapter 8

Chapter 8 explored the outcomes associated with OA and the identified disease clusters in both OA and non-OA population. Key findings from the study are

- All the outcomes are worse in people with OA compared to non-OA with nearly 2 times higher risk of all-cause mortality and 3.2 years increased DALY compared to non-OA.
- Within the OA group, people with cardiovascular only or both cardiovascular and musculoskeletal conditions had highest risk of mortality compared to other groups (1.82 times more) and higher number of hospital visits, inpatient admissions, and high DALY.
- A similar pattern was seen in the non-OA control population, with the cluster of people having hypertension and back pain having the worst outcomes.

Chapter 8 reveals OA has significant burden on the primary care and increased risk of mortality. Within OA people with cardiovascular and musculoskeletal clusters have worse outcomes and increase burden to the GP practice.

Chapter 9 summarises the findings together and describes the possible clinical implications and future work to be done.

9 Chapter 9

Discussion

9.1 Key findings and interpretation

This PhD thesis documents the epidemiology of OA and associated comorbidities in the UK using a large primary care database. Six different studies were conducted to investigate various aspects of OA and its comorbidities, including the epidemiology and trends of OA prevalence and incidence, comorbidities associated with OA prior to and post the diagnosis, common clusters of comorbidities, transitions between clusters over time, multimorbidity trajectories over time, and outcomes such as GP visits, hospitalisations, DALYs and mortality, which are relevant from both clinical and health policy perspectives.

Key findings are:

1. The prevalence of OA in UK primary care in 2017 was 10.7% and the incidence was 6.8 per 1000 person-years in people aged 20 and over, more common in women than men, and increased with age.
2. The prevalence has increased at a rate of 1.4% per year since 1998, whereas the incidence has been declining at a rate of 1.6% per year.
3. People with OA are more likely to have comorbidities and multimorbidity, both prior to and following the diagnosis of OA, than people without OA.
4. Musculoskeletal (MSK), gastrointestinal (GI), cardiovascular (CV) and psychological conditions were associated both before and after the diagnosis of OA, whereas dementia and SLE were only associated with OA after its diagnosis. Other conditions that showed significant prior and after associations with OA, were anaemia, inflammatory bowel disease, benign prostatic hypertrophy , gallstones, liver disease, cancer, and hearing impairment.

5. Five multimorbidity clusters were identified in people with OA, namely, relatively healthy, MSK, CV, CV-MSK, and mental health clusters.
6. After the diagnosis of OA, both the type and size of comorbidity clusters changed over time.
7. Nearly 30% of people moved from MSK or CV clusters to CV-MSK clusters in 10 years since the OA diagnosis.
8. Five progressive trajectories of multimorbidity (i.e., number of comorbidities) in people with OA were identified, namely, very slow, slow, gradual, rapid, and very rapid progression, of which 17.5% of people developed multimorbidity very rapidly over time. Obesity and smoking in people with OA are strongly associated with faster development of multimorbidity.
9. OA is associated with an increased number of GP visits and hospitalisations, and with increased mortality and DALYs
10. Within OA, people with CV-MSK, metabolic and CV comorbidity clusters have increased mortality, DALYs, and health utilisations

There was an increasing trend to record joint pain and a change in recording patterns for OA and joint pain after 2005. Regardless of the coding pattern, there is a clear indication of a high burden of OA or joint pain in the primary care. A high recording of 'unspecified' OA was seen, which needs to be investigated further to avoid misclassification. The knee was the most reported site of OA, but the increasing trend of reporting hand and ankle/foot OA suggests increasing recognition of the clinical importance of these common forms of OA.

There is a relative paucity of evidence about comorbidity and multimorbidity in people with OA. Available evidence has limitations, including study of just one or a small number of conditions, a sole focus on the elderly population, small sample sizes and differing methodology. This thesis shows that comorbidities in OA are common. Besides the well-known CV and MSK conditions, the newly found conditions that associate with OA merit furthers investigation. For example, the associations of IBD or gallstone with OA have not

been studied before, and the reason for these associations is unknown. Also, this epidemiological study can help to identify the burden of multimorbidity and specific comorbidities in OA and inform the development of improved person-centred care.

This study found OA to be more likely to be clustered with CVDs rather than with MSK disorders. Even though there are some shared risk factors for both conditions, the affinity towards CVDs might provide more insights into the pathophysiology of the conditions. Identification of multimorbidity clusters in OA has been studied for the first time in this study. This study found a divergence in clustering pattern, across gender and different age groups. The identified clusters were centred around hypertension, back pain, and depression. These clusters might be identified due to the study population, methodology, and high prevalence of these conditions in the study population. The identified clusters in OA had higher burden from the controls. Complex multimorbidity clusters, such as presence of both painful conditions and CVDs, had the worst outcomes studied here. Findings from this study support the proposal that interventions to improve outcomes in multimorbidity may be more appropriately targeted on distinct types, and future management guidelines for multimorbidity including in OA should consider these patterns and lead conditions.

The analysis of distinct trajectories and transitions of clusters in the OA and control populations also is the first of its kind. This helps to fill the gaps in understanding of the evolution of a patient's health journey rather than the disease itself. People with high risk lifestyles smoking and obesity develop multimorbidity faster, which provides clues for early and effective interventions to prevent or reduce the risk factors and to minimise the future multimorbidity burden. Another interesting group of people had a high burden of multimorbidity at the index date but subsequently had slower progression of multimorbidity over time. The transition of people from one cluster to another over time indicates the increasing complexity of multimorbidity with increasing age. People with only back pain or hypertension at the index date develop more complex multimorbidity in later life and people

that initially belonged to the relatively healthy group move to clusters led by back pain or hypertension only.

Finally, the burden of the OA and the clusters within people with OA indicates the severity of the condition. The OA group in this study had uniformly worst outcomes and higher mortality compared to the controls. Within the OA group, people with complex multimorbidity had the worst health outcomes and highest mortality rates. The varied section of people with different patterns of comorbidity had different burden of the conditions and need from the health care system. For example, people with both CVD and MSK lose more years of their productive life compared to others and have early mortality and increased health visits.

9.2 Strengths and limitations

The strength of the study is the large database and consistent methodology. Furthermore, data in the CPRD are collected prospectively, reducing recall bias. Latent class approaches were used to examine the clusters and used the extension version of the same for analysing the trajectories and transitions to have uniform results. In the database the conditions were recorded by GPs reflecting the real-world medical practice in the UK. The findings can be adopted for clinical policies or recommendations review. In addition, the results from this thesis can be used for other developed countries, with a similar population structure and health system in place. The statistical methods used here can be used for other diseases for studying the trends and comorbidity patterns.

The CPRD, has many linked datasets and a more robust coding system through the HES data. Use of such linked data can provide more representative in both primary and secondary care and give more accuracy about the coding. A recently published study which used HES data to estimate the trends of OA had similar results to this study (Morgan *et al.*, 2019). The inclusion of CPRD primary care data only restricts the disease burden seen in primary care and raises questions on the quality of disease coding and confirmatory diagnosis. One of the limitations is, I did not include HES or secondary data to confirm the

diagnosis which uses ICD-10 code. The CPRD is limited by the mobility of the participating patients who can move in and out of registered practices thus limiting available data to periods of active registration. The use of joint pain definition in the chapter one has its limitation, which cannot be relied in the younger population, which needs to be interpreted carefully. However, joint pain definition in the older population could be an alternative approach to investigate the OA epidemiology. The estimates of prevalence and incidence of CPRD GOLD database therefore are “period” estimates, rather than lifetime estimates. More detailed research on the associations between OA and cancer or myocardial infarction (MI) can be performed using the cancer and MI linkages. Though most of the code groups have been validated before, validation of all the code groups would be helpful.

Throughout the study of comorbidities and the clusters the same matched cohorts were used. The purpose was to compare the results in people with OA with age and sex matched controls. Because often OA is studied as controls for many outcomes. Thus, the restricted matched sample might have different results of clustering of multimorbidity, because the latter is more reliant on the size of the study population. However, the sampled population here is a true representation of the whole OA population. Although with access to linkage data of IMD, it could not be used for all the analyses because IMD is restricted to the practices in the England, which could have affected the analyses. Obesity was included as a covariate in this study rather than a morbidity. Recent studies are debating on the inclusion of obesity as a disease entity rather than a risk factor. Though there is no correct answer to this, I considered BMI categories as a covariate in the analysis to understand the role of it towards each clusters and comorbidities, which could be one of the limitations. Similarly, metabolic syndrome which is group of symptoms such as central obesity, high cholesterol, hypertension, and uncontrolled diabetes could not be derived due to lack of information on lipid levels, which adds to the limitation of the study. The association of health utilisation and burden is dependent on the number of comorbidities reported. As it is difficult to ascertain the relationship between the number of conditions reported and consultations, the analysis

was limited within OA and non-OA to descriptive only. Multimorbidity clusters with respect to joint specific OA were not studied due to the small number of people in each group with exclusive single joint specific OA.

9.3 Clinical implications

The work in this thesis has very significant clinical implications as well as implications for policy makers. Some of them have been discussed in each chapter. The first important clinical message is from the burden of OA in the primary care. The prevalence of GP diagnosed OA is nearly 10% which increases further after including consultations for joint pain. Primary care needs to be prepared to handle the burden and the future trends of OA. As OA was associated with most of the chronic conditions, people with OA might be considered for screening for early diagnosis of the linked comorbidities and the associated risk factors. The OA management plan should consider the presence of known comorbidities. Another important implication is the clustering of OA. People with OA should be examined for potential CVDs and vice versa after the age of 40 years. People with OA can further be grouped into different subgroups based on the clinical comorbidities present. For example, a person with both CVD and MSK should be given some priority due to the worse outcomes. An integrated person-centred care package can be developed for each cluster within OA, which is more feasible than developing one for all instances of multimorbidity. Each person with OA and comorbidity should be monitored for early detection of further comorbidities so that the trajectory of the multimorbidity can be slowed.

Due to the lack of appropriate management protocol, multimorbidity constitutes a challenge for the organisation of health and social. As mentioned, there is need to provide person-centred integrated care as vis-a-vis fragmented and single-disease focused care. The evidence toward success of integrated model is convincing, however there lies the complexity in development of such care model. For instance, integrated care model in multimorbid people with OA has not been developed yet. Few researchers are focusing on

OA and CVDs or pain comorbidities. It is in real a tough task to design care models based on the diseases as multiple risk factors, medications prescribed and the health system's readiness each contributes to its complexity. A systematic review looked at the integrated care models in the Europe for multimorbidity, concluded the complexities of the diseases was not accounted in different models (Struckmann *et al.*, 2018). A cost-utility analysis on integrated model discarded the suitability of implementing for all age groups and not efficient (Lanzeta, Mar and Arrospide, 2016). One integrated care model trail for people with OA focused in improving the OA symptoms but not on comorbidities (Østerås *et al.*, 2019). Various MSK model of care was found to be successful in improving the joint pain symptoms, however it's impact on other diseases were not studied (Dziedzic *et al.*, 2016).

9.4 Future work

Some important questions have been addressed in this thesis, but others remain unanswered.

The burden of unspecified OA site in primary care, and the reasons for such recording, needs to be explored further. Although regional diversity was seen in the prevalence and incidence of OA, this can be better studied using a more complete primary care database of the UK. This study documented broadly that people with OA have higher risk of comorbidities at diagnosis and risk of some comorbidities also increased after the diagnosis, but more detailed studies to confirm and to further explore these associations are required. For example, the association between OA and hearing problems is interesting but hard to explain, and merits further research. Both epidemiological and biological studies need to be conducted to support the identified associations. Especially, the role of the disease and the associated drug treatments need to be differentiated and quantified to understand the major contributors towards the comorbidities. A different methodology in a different database may also be used to validate the findings.

For clustering of comorbidities, a different set of analytical approaches could be used to verify the clusters. Though results of alternative machine learning method is provided in the Appendix Method 2 (Page 347). More joint specific analysis should be helpful to understand the underlying potential pathophysiology. Even though age and BMI are thought to be a significant shared risk factor for comorbidities in people with OA, the new evidence of inflammatory substances needs further attention. The details of the possible pathophysiology of the comorbidities in people with OA is given in Figure 4.41, page 141 . As OA is a slow progressing disease and often diagnosed in advanced stage of joint degeneration, the inflammatory changes start appearing much before whose association with other comorbidities need to be explored. The emergence of association with new comorbidities in this study makes worth wondering the causal association with some of the conditions such as hearing loss, gall bladder stone and BPH. Both biological and genetic studies can be done to understand the shared risk factors or any hidden causal factor for the disease and their temporal associations. New methods such as mendelian randomisation and genome wise study can help in understanding the underlying pathophysiology for the disease.

Further research should be done within the identified clusters for shared risk factors. The cluster with high multimorbidity at the index date but slower progress afterwards might be studied in detail to determine the most effective chronic care model. More dynamic modelling can be used to find the factors associated with the transition of people between the clusters. More robust methodological research must be done into methods for analysing clustering of binary data. The evolving nature of multimorbidity should be modelled with dynamic changes in associated risk factors. More risk factors such as medicines use, diet, physical activity, and ethnicity can be included.

Further research should be carried out to understand the care complexities in identified clusters from both patient and physician's perspectives. Health seeking behaviour and access to health care factors should be considered in the analysis of comorbidities, particularly while using the electronic health records. Alternatively, large scale community

studies should be encouraged among representative samples. Qualitative research among people within each cluster can help with identifying the facilitators and barriers towards chronic care for OA. The complexity of chronic conditions, drugs and management protocols can be studied in detail for designing effective care packages. The association of drugs used for other conditions such as CVDs or GI on OA should not be ignored. A life course epidemiology design can be used to understand the evolving nature of OA and other multimorbidity. The associated risk factors must be studied long before the clinical presentation or time of diagnosis of OA. The contribution of biological markers should be explored in detail, especially for inflammation and depression towards developing multimorbidity.

Economic analysis of loss due to OA and its multimorbidity can provide great insight for better resource allocation. The mortality accounted for by drug use in OA (NSAIDs, opioids) need to be investigated further. A risk prediction model could be developed for illness pathways of developing multimorbidity and the health utilisation and mortality in people with OA.

9.5 Conclusion

In conclusion, CPRD GOLD data have been used successfully to address the epidemiological and clinical research questions in OA. The thesis provides evidence for the current epidemiology of OA, associated comorbidities, comorbidity clusters, their evolution and trajectory of development, and the severity of each cluster for mortality and health utilisation. All these address clinically relevant questions for primary care. The methodologies established using this thesis lay the foundation for future research in multimorbidity using electronic health records.

10 References

Aarts, S. *et al.* (2012) 'The effect of multimorbidity on health related functioning: Temporary or persistent? Results from a longitudinal cohort study', *Journal of Psychosomatic Research*, 73(3), pp. 211–217. doi: 10.1016/j.jpsychores.2012.05.014.

Abhishek A and Doherty M (2013) *Diagnosis and Clinical Presentation of Osteoarthritis*, *Rheumatic Disease Clinics of North America*.

Ackerman, I. N. *et al.* (2019) 'The projected burden of primary total knee and hip replacement for osteoarthritis in Australia to the year 2030', *BMC Musculoskeletal Disorders*, 20(1). doi: 10.1186/s12891-019-2411-9.

Akker, M. van den, Buntinx, F. and Knottnerus, J. A. (1996) 'Comorbidity or multimorbidity', *European Journal of General Practice*, 2(2), pp. 65–70. doi: 10.3109/13814789609162146.

Al-Jarallah, K. *et al.* (2016) 'Knee Osteoarthritis in Type 2 Diabetes Mellitus: Does Insulin Therapy Retard Osteophyte Formation?', *Medical Principles and Practice*, 25(1), pp. 12–17. doi: 10.1159/000441418.

Allen, K. D. (2010) 'Racial and ethnic disparities in osteoarthritis phenotypes':, *Current Opinion in Rheumatology*, 22(5), pp. 528–532. doi: 10.1097/BOR.0b013e32833b1b6f.

Altman, R. *et al.* (1986) 'Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association', *Arthritis and Rheumatism*, 29(8), pp. 1039–1049.

Anandacoomarasamy, A. and March, L. (2010) 'Current evidence for osteoarthritis treatments', *Therapeutic Advances in Musculoskeletal Disease*, 2(1), pp. 17–28. doi: 10.1177/1759720X09359889.

Andersen, R. and Newman, J. F. (1973) 'Societal and individual determinants of medical care utilization in the United States', *The Milbank Memorial Fund Quarterly. Health and Society*, 51(1), pp. 95–124.

Argoff, C. E. and Gloth, F. M. (2011) 'Topical nonsteroidal anti-inflammatory drugs for management of osteoarthritis in long-term care patients', *Therapeutics and Clinical Risk Management*, 7, pp. 393–399. doi: 10.2147/TCRM.S24458.

Arora, A. *et al.* (2016) 'Cirrhosis-related musculoskeletal disease: radiological review', *The British Journal of Radiology*, 89(1066), p. 20150450. doi: 10.1259/bjr.20150450.

Ashraf, A. *et al.* (2014) 'Correlation between Degree of Radiologic Signs of Osteoarthritis and Functional Status in Patients with Chronic Mechanical Low Back Pain', *The Malaysian Journal of Medical Sciences : MJMS*, 21(2), pp. 28–33.

Bähler, C. *et al.* (2015) 'Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study', *BMC health services research*, 15, p. 23. doi: 10.1186/s12913-015-0698-2.

Bair, M. J. *et al.* (2003) 'Depression and Pain Comorbidity: A Literature Review', *Archives of Internal Medicine*, 163(20), p. 2433. doi: 10.1001/archinte.163.20.2433.

Barbour, K. E. *et al.* (2015) 'Hip Osteoarthritis and the Risk of All-Cause and Disease-Specific Mortality in Older Women: A Population-Based Cohort Study: HIP OA AND Mortality in Older Women', *Arthritis & Rheumatology*, 67(7), pp. 1798–1805. doi: 10.1002/art.39113.

Barnett, K. *et al.* (2012) 'Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study', *Lancet*, 380(9836), pp. 37–43. doi: 10.1016/S0140-6736(12)60240-2.

Bartholomew, D. J. (ed.) (2008) *Analysis of multivariate social science data*. 2nd ed. Boca Raton: CRC Press (Chapman & Hall/CRC statistics in the social and behavioral sciences series).

Bedson, J. *et al.* (2007) 'Knee pain and osteoarthritis in the general population: what influences patients to consult?', *Family Practice*, 24(5), pp. 443–453. doi: 10.1093/fampra/cmm036.

Bedson, J. and Croft, P. R. (2008) 'The discordance between clinical and radiographic knee osteoarthritis: A systematic search and summary of the literature', *BMC Musculoskeletal Disorders*, 9, p. 116. doi: 10.1186/1471-2474-9-116.

Benjamini, Y. and Yekutieli, D. (2001) 'The Control of the False Discovery Rate in Multiple Testing under Dependency', *The Annals of Statistics*, 29(4), pp. 1165–1188.

Berenbaum, F. (2013) 'Osteoarthritis as an inflammatory disease (osteoarthritis is not osteoarthrosis!)', *Osteoarthritis and Cartilage*, 21(1), pp. 16–21. doi: 10.1016/j.joca.2012.11.012.

Bhaskaran, K. *et al.* (2013) 'Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD)', *BMJ Open*, 3(9), p. e003389. doi: 10.1136/bmjopen-2013-003389.

Blyth, F. M. *et al.* (2019) 'The Global Burden of Musculoskeletal Pain—Where to From Here?', *American Journal of Public Health*, 109(1), pp. 35–40. doi: 10.2105/AJPH.2018.304747.

Bollegala, D., Perruccio, A. V. and Badley, E. M. (2011) 'Combined impact of concomitant arthritis and back problems on health status: Results from a nationally representative health survey', *Arthritis Care & Research*, 63(11), pp. 1584–1591. doi: 10.1002/acr.20595.

Booth, N. (1994) 'What are the Read Codes?', (II), pp. 177–182.

Boshuizen, H. C. *et al.* (2017) 'Taking multi-morbidity into account when attributing DALYs to risk factors: comparing dynamic modeling with the GBD2010 calculation method', *BMC Public Health*, 17(1). doi: 10.1186/s12889-017-4024-2.

Bravatà, V. *et al.* (2015) 'DVWA gene polymorphisms and osteoarthritis', *BMC Research Notes*, 8(1), p. 30. doi: 10.1186/s13104-015-0987-1.

Breedveld, F. C. (2004) 'Osteoarthritis--the impact of a serious disease', *Rheumatology (Oxford, England)*, 43 Suppl 1, pp. i4-8. doi: 10.1093/rheumatology/keh102.

Brennan-Olsen, S. L. *et al.* (2017) 'Prevalence of arthritis according to age, sex and socioeconomic status in six low and middle income countries: analysis of data from the World Health Organization study on global AGEing and adult health (SAGE) Wave 1', *BMC Musculoskeletal Disorders*, 18(1), p. 271. doi: 10.1186/s12891-017-1624-z.

Bridges, C. C. (1966) 'Hierarchical Cluster Analysis', *Psychological Reports*, 18(3), pp. 851–854. doi: 10.2466/pr0.1966.18.3.851.

Brilleman, S. L. and Salisbury, C. (2013) 'Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study', *Family Practice*, 30(2), pp. 172–178. doi: 10.1093/fampra/cms060.

British Heart Foundation (2019) 'Heart & Circulatory Disease Statistics 2019'. British Heart Foundation. Available at: https://www.bhf.org.uk/what-we-do/our-research/heart-statistics/heart-statistics-publications/cardiovascular-disease-statistics-2019.

Calderón-Larrañaga, A. *et al.* (2019) 'Multimorbidity and functional impairment-bidirectional interplay, synergistic effects and common pathways', *Journal of Internal Medicine*, 285(3), pp. 255–271. doi: 10.1111/joim.12843.

Canizares, M. *et al.* (2018) 'Increasing Trajectories of Multimorbidity Over Time: Birth Cohort Differences and the Role of Changes in Obesity and Income', *The Journals of Gerontology: Series B*, 73(7), pp. 1303–1314. doi: 10.1093/geronb/gbx004.

Carstensen, B. (2005) 'Demography and epidemiology: Age-Period-Cohort models in the computer age'. Steno Diabetes Center, Gentofte, Denmark & Department of Biostatistics, University of Copenhagen. Available at: www.pubhealth.ku.dk/~bxc (Accessed: 6 January 2019).

Carstensen, B. *et al.* (2019) *Epi package for epidemiological analysis in R*. Available at: http://bendixcarstensen.com/Epi/ (Accessed: 11 June 2019).

Cassell A. *et al.* (2018) 'The epidemiology of multimorbidity in primary care: A retrospective cohort study', *British Journal of General Practice*, 68(669), pp. e245–e251. doi: 10.3399/bjgp18X695465.

Charlier, E. *et al.* (2016) 'Insights on Molecular Mechanisms of Chondrocytes Death in Osteoarthritis', *International Journal of Molecular Sciences*, 17(12). doi: 10.3390/ijms17122146.

Charlson, M. *et al.* (1994) 'Validation of a combined comorbidity index', *Journal of Clinical Epidemiology*, 47(11), pp. 1245–1251. doi: 10.1016/0895-4356(94)90129-5.

Charlson, M. E. *et al.* (1987) 'A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation', *Journal of Chronic Diseases*, 40(5), pp. 373–383. doi: 10.1016/0021-9681(87)90171-8.

Chen, A. *et al.* (2012) 'The Global Economic Cost of Osteoarthritis: How the UK Compares', *Arthritis*, 2012. doi: 10.1155/2012/698709.

Chiu, D. S. and Talhouk, A. (2018) 'diceR: an R package for class discovery using an ensemble driven approach', *BMC Bioinformatics*, 19(1). doi: 10.1186/s12859-017-1996-y.

Choong, P. F. and Dowsey, M. M. (2014) 'The Grand Challenge – Managing End-Staged Joint Osteoarthritis', *Frontiers in Surgery*, 1. doi: 10.3389/fsurg.2014.00009.

Chudasama, Y. V. *et al.* (2019) 'Physical activity, multimorbidity, and life expectancy: a UK Biobank longitudinal study', *BMC Medicine*, 17(1). doi: 10.1186/s12916-019-1339-0.

Chughtai, B. *et al.* (2011) 'Role of inflammation in benign prostatic hyperplasia', *Reviews in Urology*, 13(3), pp. 147–150.

Clark, D. O. *et al.* (1995) 'A chronic disease score with empirically derived weights', *Medical Care*, 33(8), pp. 783–795.

Cleveland, R. J., Nelson, A. E. and Callahan, L. F. (2019) 'Knee and hip osteoarthritis as predictors of premature death: a review of the evidence', *Clinical and Experimental Rheumatology*, 37 Suppl 120(5), pp. 24–30.

Clinical Practice Research Datalink | CPRD (no date). Available at: https://www.cprd.com/ (Accessed: 13 February 2019).

Clynes, M. A. *et al.* (2014) 'Further evidence of the developmental origins of osteoarthritis: results from the Hertfordshire Cohort Study', *Journal of Developmental Origins of Health and Disease*, 5(6), pp. 453–458. doi: 10.1017/S2040174414000373.

Coderre, T. J. *et al.* (1993) 'Contribution of central neuroplasticity to pathological pain: review of clinical and experimental evidence', *Pain*, 52(3), pp. 259–285. doi: 10.1016/0304-3959(93)90161-h.

Collerton, J. *et al.* (2016) 'Deconstructing Complex Multimorbidity in the Very Old: Findings from the Newcastle 85+ Study', *BioMed Research International*, 2016, pp. 1–15. doi: 10.1155/2016/8745670.

Collins, L. M. and Lanza, S. T. (2009) *Latent Class and Latent Transition Analysis*. Hoboken, NJ, USA: John Wiley & Sons, Inc. (Wiley Series in Probability and Statistics). doi: 10.1002/9780470567333.

Conaghan, P. G. *et al.* (2008) 'Care and management of osteoarthritis in adults: summary of NICE guidance', *BMJ (Clinical research ed.)*, 336(7642), pp. 502–503. doi: 10.1136/bmj.39490.608009.AD.

da Costa, B. R. *et al.* (2017) 'Effectiveness of non-steroidal anti-inflammatory drugs for the treatment of pain in knee and hip osteoarthritis: a network meta-analysis', *The Lancet*, 390(10090), pp. e21–e33. doi: 10.1016/S0140-6736(17)31744-0.

Couzin-Frankel, J. (2010) 'Inflammation Bares a Dark Side', *Science*, 330(6011), pp. 1621–1621. doi: 10.1126/science.330.6011.1621.

Culliford D *et al.* (2015) *Future projections of total hip and knee arthroplasty in the UK: Results from the UK Clinical Practice Research Datalink*, *Osteoarthritis and Cartilage*. Available at: http://www.elsevier.com/inca/publications/store/6/2/3/0/5/5/index.htt.

Das, P., Naylor, C. and Majeed, A. (2016) 'Bringing together physical and mental health within primary care: a new frontier for integrated care', *Journal of the Royal Society of Medicine*, 109(10), pp. 364–366. doi: 10.1177/0141076816665270.

Dell'isola, A. *et al.* (2018) 'Knee extensor muscle weakness and radiographic knee osteoarthritis progression', *Acta Orthopaedica*, 89(4), pp. 406–411. doi: 10.1080/17453674.2018.1464314.

Dequeker, J., Aerssens, J. and Luyten, F. P. (2003) 'Osteoarthritis and osteoporosis: clinical and research evidence of inverse relationship', *Aging Clinical and Experimental Research*, 15(5), pp. 426–439.

Dequeker, J. and Luyten, F. P. (2008) 'The history of osteoarthritis-osteoarthrosis', *Annals of the Rheumatic Diseases*, 67(1), pp. 5–10. doi: 10.1136/ard.2007.079764.

Derry, S. *et al.* (2016) 'Topical NSAIDs for chronic musculoskeletal pain in adults', *Cochrane Database of Systematic Reviews*. Edited by Cochrane Pain, Palliative and Supportive Care Group. doi: 10.1002/14651858.CD007400.pub3.

Deruaz-Luyet A. *et al.* (2017) 'Multimorbidity and patterns of chronic conditions in a primary care population in Switzerland: A cross-sectional study', *BMJ Open*, 7(6), p. e013664. doi: 10.1136/bmjopen-2016-013664.

Deyo, R. (1992) 'Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases', *Journal of Clinical Epidemiology*, 45(6), pp. 613–619. doi: 10.1016/0895-4356(92)90133-8.

Dhalwani N.N. *et al.* (2017) 'Association Between Lifestyle Factors and the Incidence of Multimorbidity in an Older English Population', *The journals of gerontology. Series A, Biological sciences and medical sciences*, 72(4), pp. 528–534. doi: 10.1093/gerona/glw146.

Dillon, C. F. *et al.* (2006) 'Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94', *The Journal of Rheumatology*, 33(11), pp. 2271–2279.

Dixon, T. (2004) 'Trends in hip and knee joint replacement: socioeconomic inequalities and projections of need', *Annals of the Rheumatic Diseases*, 63(7), pp. 825–830. doi: 10.1136/ard.2003.012724.

Doherty, M. (2001) 'Risk factors for progression of knee osteoarthritis', *The Lancet*, 358(9284), pp. 775–776. doi: 10.1016/S0140-6736(01)06006-8.

Doherty, M. *et al.* (2008) 'Nonspherical femoral head shape (pistol grip deformity), neck shaft angle, and risk of hip osteoarthritis: A case-control study', *Arthritis & Rheumatism*, 58(10), pp. 3172–3182. doi: 10.1002/art.23939.

Doherty, M. *et al.* (eds) (2016) *Oxford Textbook of Osteoarthritis and Crystal Arthropathy*. Oxford University Press. doi: 10.1093/med/9780199668847.001.0001.

Doherty, M. and Dougados, M. (2001) 'Evidence-based management of osteoarthritis: practical issues relating to the data', *Best Practice & Research. Clinical Rheumatology*, 15(4), pp. 517–525. doi: 10.1053/berh.2001.0170.

Dore A *et al.* (2013) *Risk of falls increases with additional symptomatic osteoarthritic joints: The Johnston county osteoarthritis project*, *Arthritis and Rheumatism*.

Dowsey, M. M., Smith, A. J. and Choong, P. F. M. (2015) 'Latent Class Growth Analysis predicts long term pain and function trajectories in total knee arthroplasty: a study of 689 patients', *Osteoarthritis and Cartilage*, 23(12), pp. 2141–2149. doi: 10.1016/j.joca.2015.07.005.

Dua A. B *et al.* (2012) *Central sensitization is associated with spontaneous pain in knee osteoarthritis*, *Arthritis and Rheumatism*.

Duffield, S. J. *et al.* (2017) 'The contribution of musculoskeletal disorders in multimorbidity: Implications for practice and policy', *Best Practice and Research: Clinical Rheumatology*, 31(2), pp. 129–144. doi: 10.1016/j.berh.2017.09.004.

Dziedzic, K. S. *et al.* (2016) 'Implementation of musculoskeletal Models of Care in primary care settings: Theory, practice, evaluation and outcomes for musculoskeletal health in high-income economies', *Best Practice & Research Clinical Rheumatology*, 30(3), pp. 375–397. doi: 10.1016/j.berh.2016.08.004.

Elixhauser, A. *et al.* (1998) 'Comorbidity measures for use with administrative data', *Medical Care*, 36(1), pp. 8–27.

Enomoto, H. *et al.* (2003) 'Vascular Endothelial Growth Factor Isoforms and Their Receptors Are Expressed in Human Osteoarthritic Cartilage', *The American Journal of Pathology*, 162(1), pp. 171–181. doi: 10.1016/S0002-9440(10)63808-4.

Ettinger, W. H. *et al.* (1997) 'A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. The Fitness Arthritis and Seniors Trial (FAST)', *JAMA*, 277(1), pp. 25–31.

Feinstein, A. R. (1970) 'The pre-therapeutic classification of co-morbidity in chronic disease', *Journal of Chronic Diseases*, 23(7), pp. 455–468.

Felson, D. T. (2000) 'Osteoarthritis: New Insights. Part 1: The Disease and Its Risk Factors', *Annals of Internal Medicine*, 133(8), p. 635. doi: 10.7326/0003-4819-133-8-200010170-00016.

Felson, D. T. (2013) 'Osteoarthritis as a disease of mechanics', *Osteoarthritis and Cartilage*, 21(1), pp. 10–15. doi: 10.1016/j.joca.2012.09.012.

Ferguson, R. J. *et al.* (2019) 'Validation of hip osteoarthritis diagnosis recording in the UK Clinical Practice Research Datalink', *Pharmacoepidemiology and Drug Safety*, 28(2), pp. 187–193. doi: 10.1002/pds.4673.

Fernández López, J. C. and Ruano-Ravina, A. (2006) 'Efficacy and safety of intraarticular hyaluronic acid in the treatment of hip osteoarthritis: a systematic review', *Osteoarthritis and Cartilage*, 14(12), pp. 1306–1311. doi: 10.1016/j.joca.2006.08.003.

Feuillet, F. *et al.* (2015) 'On Comparison of Clustering Methods for Pharmacoepidemiological Data', *Journal of Biopharmaceutical Statistics*, 25(4), pp. 843–856. doi: 10.1080/10543406.2014.920855.

Findlay, D. M. (2007) 'Vascular pathology and osteoarthritis', *Rheumatology*, 46(12), pp. 1763–1768. doi: 10.1093/rheumatology/kem191.

Finney, A. *et al.* (2017) 'Multisite peripheral joint pain: a cross-sectional study of prevalence and impact on general health, quality of life, pain intensity and consultation behaviour', *BMC Musculoskeletal Disorders*, 18(1). doi: 10.1186/s12891-017-1896-3.

Fontana, L. *et al.* (2007) 'Osteoarthritis of the Thumb Carpometacarpal Joint in Women and Occupational Risk Factors: A Case–Control Study', *The Journal of Hand Surgery*, 32(4), pp. 459–465. doi: 10.1016/j.jhsa.2007.01.014.

Forbes, L. J. *et al.* (2017) 'The role of the Quality and Outcomes Framework in the care of long-term conditions: a systematic review', *British Journal of General Practice*, 67(664), pp. e775–e784. doi: 10.3399/bjgp17X693077.

Fortin, M. *et al.* (2010) 'Prevalence estimates of multimorbidity: a comparative study of two sources', *BMC Health Services Research*, 10(1), p. 111. doi: 10.1186/1472-6963-10-111.

Fransen, M. *et al.* (2015) 'Exercise for osteoarthritis of the knee', *Cochrane Database of Systematic Reviews*. Edited by Cochrane Musculoskeletal Group. doi: 10.1002/14651858.CD004376.pub3.

Friedman E.M., Christ S.L., and Mroczek D.K. (2015) 'Inflammation Partially Mediates the Association of Multimorbidity and Functional Limitations in a National Sample of Middle-Aged and Older Adults: The MIDUS Study', *Journal of aging and health*, 27(5), pp. 843–863. doi: 10.1177/0898264315569453.

Garfinkel, R. J., Dilisio, M. F. and Agrawal, D. K. (2017) 'Vitamin D and Its Effects on Articular Cartilage and Osteoarthritis', *Orthopaedic Journal of Sports Medicine*, 5(6), p. 232596711771137. doi: 10.1177/2325967117711376.

Geusens, P. and Bergh, J. van den (2016) 'Osteoporosis and osteoarthritis: shared mechanisms and epidemiology', *Current Opinion in Rheumatology*, 28(2), pp. 97–103. doi: 10.1097/BOR.0000000000000256.

Gómez, R. *et al.* (2011) 'What's new in our understanding of the role of adipokines in rheumatic diseases?', *Nature Reviews Rheumatology*, 7, p. 528.

Goodman, R. A. *et al.* (2016) 'Multimorbidity Patterns in the United States: Implications for Research and Clinical Practice', *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 71(2), pp. 215–220. doi: 10.1093/gerona/glv199.

Greenfield, S. *et al.* (1993) 'The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement', *Medical Care*, 31(2), pp. 141–154.

Greenland, S. (2008) 'Multiple comparisons and association selection in general epidemiology', *International Journal of Epidemiology*, 37(3), pp. 430–434. doi: 10.1093/ije/dyn064.

Guisado-Clavero, M. *et al.* (2018) 'Multimorbidity patterns in the elderly: a prospective cohort study with cluster analysis', *BMC Geriatrics*, 18(1). doi: 10.1186/s12877-018-0705-7.

Guthrie, B. *et al.* (2012) 'Adapting clinical guidelines to take account of multimorbidity', *BMJ*, 345(oct04 1), pp. e6341–e6341. doi: 10.1136/bmj.e6341.

Haag, M. D. M. (2008) 'Cyclooxygenase Selectivity of Nonsteroidal Anti-inflammatory Drugs and Risk of Stroke', *Archives of Internal Medicine*, 168(11), p. 1219. doi: 10.1001/archinte.168.11.1219.

Hall, A. J. *et al.* (2016) 'Association between osteoarthritis and cardiovascular disease: Systematic review and meta-analysis', *European Journal of Preventive Cardiology*, 23(9), pp. 938–946. doi: 10.1177/2047487315610663.

Haug, N. *et al.* (2020) 'High-risk multimorbidity patterns on the road to cardiovascular mortality', *BMC Medicine*, 18(1). doi: 10.1186/s12916-020-1508-1.

Hawker, G. A. *et al.* (2008) 'Understanding the pain experience in hip and knee osteoarthritis--an OARSI/OMERACT initiative', *Osteoarthritis and Cartilage*, 16(4), pp. 415–422. doi: 10.1016/j.joca.2007.12.017.

Hawker, G. A. *et al.* (2014) 'All-Cause Mortality and Serious Cardiovascular Events in People with Hip and Knee Osteoarthritis: A Population Based Cohort Study', *PLoS ONE*. Edited by R. B. Sim, 9(3), p. e91286. doi: 10.1371/journal.pone.0091286.

Henry, D. *et al.* (1996) 'Variability in risk of gastrointestinal complications with individual non-steroidal anti-inflammatory drugs: results of a collaborative meta-analysis', *BMJ*, 312(7046), pp. 1563–1566. doi: 10.1136/bmj.312.7046.1563.

Herrett, E. *et al.* (2010) 'Validation and validity of diagnoses in the General Practice Research Database: a systematic review', *British Journal of Clinical Pharmacology*, 69(1), pp. 4–14. doi: 10.1111/j.1365-2125.2009.03537.x.

Herrett, E., Gallagher, Arlene M., *et al.* (2015) 'Data Resource Profile: Clinical Practice Research Datalink (CPRD)', *International Journal of Epidemiology*, p. dyv098. doi: 10.1093/ije/dyv098.

Herrett, E., Gallagher, Arlene M, *et al.* (2015) 'Data Resource Profile: Clinical Practice Research Datalink (CPRD)', *International Journal of Epidemiology*, 44(3), pp. 827–836. doi: 10.1093/ije/dyv098.

Hilderink, H. B. M. *et al.* (2016) 'Accounting for multimorbidity can affect the estimation of the Burden of Disease: a comparison of approaches', *Archives of Public Health*, 74(1). doi: 10.1186/s13690-016-0147-7.

Hippisley-Cox, J. and Coupland, C. (2005) 'Risk of myocardial infarction in patients taking cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs: population based nested case-control analysis', *BMJ*, 330(7504), p. 1366. doi: 10.1136/bmj.330.7504.1366.

Hochberg, M. C. (2008) 'Mortality in osteoarthritis', *Clinical and Experimental Rheumatology*, 26(5 Suppl 51), pp. S120-124.

Hochberg, M. C. *et al.* (2015) *Rheumatology*. Available at: https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=974367 (Accessed: 28 August 2020).

Hochberg, M. C., Lethbridge-Cejku, M. and Tobin, J. D. (2004) 'Bone mineral density and osteoarthritis: data from the Baltimore Longitudinal Study of Aging', *Osteoarthritis and Cartilage*, 12 Suppl A, pp. S45-48.

Hoeven, T. A. *et al.* (2015) 'Markers of atherosclerosis in relation to presence and progression of knee osteoarthritis: a population-based cohort study', *Rheumatology (Oxford, England)*, 54(9), pp. 1692–1698. doi: 10.1093/rheumatology/kev106.

Hoogeboom, T. J., Broeder, A. A. den, *et al.* (2012) 'Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: a cross-sectional study', *Arthritis Care & Research*, 64(1), pp. 54–58. doi: 10.1002/acr.20647.

Hoogeboom, T. J., den Broeder, A. A., *et al.* (2012) 'Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: a cross-sectional study', *Arthritis Care & Research*, 64(1), pp. 54–58. doi: 10.1002/acr.20647.

Hootman, J. M. *et al.* (2003) 'Physical activity levels among the general US adult population and in adults with and without arthritis', *Arthritis & Rheumatism*, 49(1), pp. 129–135. doi: 10.1002/art.10911.

Hoy, D. *et al.* (2014) 'The global burden of low back pain: estimates from the Global Burden of Disease 2010 study', *Annals of the Rheumatic Diseases*, 73(6), pp. 968–974. doi: 10.1136/annrheumdis-2013-204428.

Hsu, H.-C. (2015) 'Trajectories of multimorbidity and impacts on successful aging', *Experimental Gerontology*, 66, pp. 32–38. doi: 10.1016/j.exger.2015.04.005.

Hsu, P.-S. *et al.* (2017) 'Increased risk of stroke in patients with osteoarthritis: a population-based cohort study', *Osteoarthritis and Cartilage*, 25(7), pp. 1026–1031. doi: 10.1016/j.joca.2016.10.027.

Huang, S.-W. *et al.* (2015) 'Osteoarthritis Increases the Risk of Dementia: A Nationwide Cohort Study in Taiwan', *Scientific Reports*, 5(1). doi: 10.1038/srep10145.

Huang, Z. and Ng, M. K. (2003) 'A Note on K-modes Clustering', *Journal of Classification*, 20(2), pp. 257–261. doi: 10.1007/s00357-003-0014-4.

Hui, M., Doherty, M. and Zhang, W. (2011) 'Does smoking protect against osteoarthritis? Meta-analysis of observational studies', *Annals of the Rheumatic Diseases*, 70(7), pp. 1231–1237. doi: 10.1136/ard.2010.142323.

Hunter, D. J. *et al.* (2013) 'Structural correlates of pain in joints with osteoarthritis', *Osteoarthritis and Cartilage*, 21(9), pp. 1170–1178. doi: 10.1016/j.joca.2013.05.017.

Hunter, D. J. and Eckstein, F. (2009) 'Exercise and osteoarthritis', *Journal of Anatomy*, 214(2), pp. 197–207. doi: 10.1111/j.1469-7580.2008.01013.x.

Hussain, S. M. *et al.* (2018) 'Could low birth weight and preterm birth be associated with significant burden of hip osteoarthritis? A systematic review', *Arthritis Research & Therapy*, 20(1), p. 121. doi: 10.1186/s13075-018-1627-7.

Hwang, H. S. *et al.* (2015) 'Monosodium Urate Crystal-Induced Chondrocyte Death via Autophagic Process', *International Journal of Molecular Sciences*, 16(12), pp. 29265–29277. doi: 10.3390/ijms161226164.

Ibarra-Castillo, C. *et al.* (2018) 'Survival in relation to multimorbidity patterns in older adults in primary care in Barcelona, Spain (2010–2014): a longitudinal study based on electronic health records', *Journal of Epidemiology and Community Health*, 72(3), pp. 185–192. doi: 10.1136/jech-2017-209984.

Ibrahim, T. *et al.* (2010) 'Temporal trends in primary total hip and knee arthroplasty surgery: results from a UK regional joint register, 1991–2004', *The Annals of The Royal College of Surgeons of England*, 92(3), pp. 231–235. doi: 10.1308/003588410X12628812458572.

Im, G.-I. and Kim, M.-K. (2014) 'The relationship between osteoarthritis and osteoporosis', *Journal of Bone and Mineral Metabolism*, 32(2), pp. 101–109. doi: 10.1007/s00774-013-0531-0.

Imhof, H. *et al.* (2000) 'Subchondral Bone and Cartilage Disease: A Rediscovered Funct... : Investigative Radiology', 35(10), pp. 581–588.

Islam, M. M. *et al.* (2014) 'Multimorbidity and Comorbidity of Chronic Diseases among the Senior Australians: Prevalence and Patterns', *PLoS ONE*, 9(1). doi: 10.1371/journal.pone.0083783.

Jackson, C. A. *et al.* (2015) 'Body mass index and socioeconomic position are associated with 9-year trajectories of multimorbidity: A population-based study', *Preventive Medicine*, 81, pp. 92–98. doi: 10.1016/j.ypmed.2015.08.013.

Jensen, A. B. *et al.* (2014) 'Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients', *Nature Communications*, 5(1). doi: 10.1038/ncomms5022.

Jia, H. *et al.* (2015) 'Impact of depression on quality-adjusted life expectancy (QALE) directly as well as indirectly through suicide', *Social Psychiatry and Psychiatric Epidemiology*, 50(6), pp. 939–949. doi: 10.1007/s00127-015-1019-0.

Joinpoint Regression Program (2018). US: Statistical Methodology and Applications Branch, Surveillance Research Program, National Cancer Institute. Available at: https://surveillance.cancer.gov/joinpoint/.

Jonas, J. B. *et al.* (2018) 'Systemic inflammation and eye diseases. The Beijing Eye Study', *PLOS ONE*. Edited by G.-S. Liu, 13(10), p. e0204263. doi: 10.1371/journal.pone.0204263.

Jordan, J. M. *et al.* (2007) 'Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project', *The Journal of Rheumatology*, 34(1), pp. 172–180.

Jordan, K. *et al.* (2007) 'Measuring disease prevalence: a comparison of musculoskeletal disease using four general practice consultation databases', *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, 57(534), pp. 7–14.

Jordan, K. P. *et al.* (2010) 'Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study', *BMC musculoskeletal disorders*, 11, p. 144. doi: 10.1186/1471-2474-11-144.

Jordan, K. P. *et al.* (2014) 'International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden', *Annals of the Rheumatic Diseases*, 73(1), pp. 212–218. doi: 10.1136/annrheumdis-2012-202634.

Jüni, P. *et al.* (2004) 'Risk of cardiovascular events and rofecoxib: cumulative meta-analysis', *The Lancet*, 364(9450), pp. 2021–2029. doi: 10.1016/S0140-6736(04)17514-4.

Jüni, P. *et al.* (2015) 'Intra-articular corticosteroid for knee osteoarthritis', *Cochrane Database of Systematic Reviews*, (10). doi: 10.1002/14651858.CD005328.pub3.

Kadam, U., Jordan, K. and Croft, P. (2004) 'Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales', *Annals of the Rheumatic Diseases*, 63(4), pp. 408–414. doi: 10.1136/ard.2003.007526.

Kadam, U. T., Jordan, K. and Croft, P. R. (2004) 'Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales', *Annals of the Rheumatic Diseases*, 63(4), pp. 408–414. doi: 10.1136/ard.2003.007526.

Kaplan, M. H. and Feinstein, A. R. (1974) 'The importance of classifying initial co-morbidity in evaluating the outcome of diabetes mellitus', *Journal of Chronic Diseases*, 27(7–8), pp. 387–404.

Kellgren, J. H. and Lawrence, J. S. (1957) 'Radiological Assessment of Osteo-Arthrosis', *Annals of the Rheumatic Diseases*, 16(4), pp. 494–502. doi: 10.1136/ard.16.4.494.

Keyes, K. M. *et al.* (2010) 'What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971–2006', *Social Science & Medicine*, 70(7), pp. 1100–1108. doi: 10.1016/j.socscimed.2009.12.018.

Khan, Nada F, Harrison, S. E. and Rose, P. W. (2010) 'Validity of diagnostic coding within the General Practice Research Database: a systematic review', *The British Journal of General Practice*, 60(572), pp. e128–e136. doi: 10.3399/bjgp10X483562.

Khan, Nada F., Harrison, S. E. and Rose, P. W. (2010) 'Validity of diagnostic coding within the General Practice Research Database: a systematic review', *The British Journal of General Practice*, 60(572), pp. e128–e136. doi: 10.3399/bjgp10X483562.

Kiadaliri, A. A. *et al.* (2018) 'Temporal trend and regional disparity in osteoarthritis hospitalisations in Sweden 1998–2015', *Scandinavian Journal of Public Health*, p. 1403494818766785. doi: 10.1177/1403494818766785.

Kim, C. *et al.* (2014) 'Prevalence of Radiographic and Symptomatic Hip Osteoarthritis in an urban US Community: the Framingham Osteoarthritis Study', *Arthritis & rheumatology (Hoboken, N.J.)*, 66(11), pp. 3013–3017. doi: 10.1002/art.38795.

Kim, H. S. *et al.* (2016) 'Association between Knee Osteoarthritis, Cardiovascular Risk Factors, and the Framingham Risk Score in South Koreans: A Cross-Sectional Study', *PLOS ONE*, 11(10), p. e0165325. doi: 10.1371/journal.pone.0165325.

Kim, H.-J. *et al.* (2000) 'Permutation tests for joinpoint regression with applications to cancer rates', *Statistics in Medicine*, 19(3), pp. 335–351. doi: 10.1002/(SICI)1097-0258(20000215)19:3<335::AID-SIM336>3.0.CO;2-Z.

Kinds, M. B. *et al.* (2011) 'A systematic review of the association between radiographic and clinical osteoarthritis of hip and knee', *Osteoarthritis and Cartilage*, 19(7), pp. 768–778. doi: 10.1016/j.joca.2011.01.015.

King, K. B. and Rosenthal, A. K. (2015) 'The adverse effects of diabetes on osteoarthritis: update on clinical evidence and molecular mechanisms', *Osteoarthritis and Cartilage*, 23(6), pp. 841–850. doi: 10.1016/j.joca.2015.03.031.

Kingsbury, S. R. *et al.* (2014) 'Osteoarthritis in Europe: impact on health status, work productivity and use of pharmacotherapies in five European countries', *Rheumatology (Oxford, England)*, 53(5), pp. 937–947. doi: 10.1093/rheumatology/ket463.

Kirkness, C. S., Yu, J. and Asche, C. V. (2008) 'The effect on comorbidity and pain in patients with osteoarthritis', *Journal of Pain & Palliative Care Pharmacotherapy*, 22(4), pp. 336–348. doi: 10.1080/15360280802536649.

Klussmann, A. *et al.* (2010) 'Individual and occupational risk factors for knee osteoarthritis: results of a case-control study in Germany', *Arthritis Research & Therapy*, 12(3), p. R88. doi: 10.1186/ar3015.

Koenig, W. *et al.* (1997) 'Leisure-Time Physical Activity but Not Work-Related Physical Activity Is Associated With Decreased Plasma Viscosity: Results From a Large Population Sample', *Circulation*, 95(2), pp. 335–341. doi: 10.1161/01.CIR.95.2.335.

Kontopantelis, E. *et al.* (2018) 'Spatial distribution of clinical computer systems in primary care in England in 2016 and implications for primary care electronic medical record databases: a cross-sectional population study', *BMJ Open*, 8(2), p. e020738. doi: 10.1136/bmjopen-2017-020738.

Kousoulis, A. A., Rafi, I. and de Lusignan, S. (2015) 'The CPRD and the RCGP: building on research success by enhancing benefits for patients and practices', *British Journal of General Practice*, 65(631), pp. 54–55. doi: 10.3399/bjgp15X683353.

Kramer, S. E. *et al.* (2002) 'The Association of Hearing Impairment and Chronic Diseases with Psychosocial Health Status in Older Age', *Journal of Aging and Health*, 14(1), pp. 122–137. doi: 10.1177/089826430201400107.

Kraus, V. B. *et al.* (2015) 'Call for standardized definitions of osteoarthritis and risk stratification for clinical trials and clinical use', *Osteoarthritis and Cartilage*, 23(8), pp. 1233–1241. doi: 10.1016/j.joca.2015.03.036.

Krause, A. J. *et al.* (2019) 'The Pain of Sleep Loss: A Brain Characterization in Humans', *The Journal of Neuroscience*, 39(12), pp. 2291–2300. doi: 10.1523/JNEUROSCI.2408-18.2018.

Kuo, C.-F. *et al.* (2014a) 'Comorbidities in patients with gout prior to and following diagnosis: case-control study', *Annals of the Rheumatic Diseases*, p. annrheumdis-2014-206410. doi: 10.1136/annrheumdis-2014-206410.

Kuo, C.-F. *et al.* (2014b) 'Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study', *Annals of the Rheumatic Diseases*, p. annrheumdis-2013-204463. doi: 10.1136/annrheumdis-2013-204463.

Kyrkanides, S. *et al.* (2011) 'Osteoarthritis accelerates and exacerbates Alzheimer's disease pathology in mice', *Journal of Neuroinflammation*, 8(1), p. 112. doi: 10.1186/1742-2094-8-112.

Lanza, S. T. and Rhoades, B. L. (2013) 'Latent Class Analysis: An Alternative Perspective on Subgroup Analysis in Prevention and Treatment', *Prevention Science*, 14(2), pp. 157–168. doi: 10.1007/s11121-011-0201-1.

Lanzeta, I., Mar, J. and Arrospide, A. (2016) 'Cost-utility analysis of an integrated care model for multimorbid patients based on a clinical trial', *Gaceta Sanitaria*, 30(5), pp. 352–358. doi: 10.1016/j.gaceta.2016.05.002.

Lappenschaar, M. *et al.* (2013) 'Multilevel temporal Bayesian networks can model longitudinal change in multimorbidity', *Journal of Clinical Epidemiology*, 66(12), pp. 1405–1416. doi: 10.1016/j.jclinepi.2013.06.018.

Lawrence, R. C. *et al.* (2008) 'Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II', *Arthritis and Rheumatism*, 58(1), pp. 26–35. doi: 10.1002/art.23176.

Lee, S. and Kim, S.-J. (2017) 'Prevalence of knee osteoarthritis, risk factors, and quality of life: The Fifth Korean National Health And Nutrition Examination Survey', *International Journal of Rheumatic Diseases*, 20(7), pp. 809–817. doi: 10.1111/1756-185X.12795.

Lee, T. *et al.* (2016) 'Use of non-steroidal anti-inflammatory drugs correlates with the risk of venous thromboembolism in knee osteoarthritis patients: a UK population-based case-control study', *Rheumatology*, 55(6), pp. 1099–1105. doi: 10.1093/rheumatology/kew036.

Lespérance, F. *et al.* (2002) 'Five-Year Risk of Cardiac Mortality in Relation to Initial Severity and One-Year Changes in Depression Symptoms After Myocardial Infarction', *Circulation*, 105(9), pp. 1049–1053. doi: 10.1161/hc0902.104707.

Libby, P., Ridker, P. M. and Maseri, A. (2002) 'Inflammation and Atherosclerosis', *Circulation*, 105(9), pp. 1135–1143. doi: 10.1161/hc0902.104353.

Linn, B. S., Linn, M. W. and Gurel, L. (1968) 'Cumulative illness rating scale', *Journal of the American Geriatrics Society*, 16(5), pp. 622–626.

Linzer, D. A. and Lewis, J. B. (2011) 'Package for Polytomous Variable Latent Class Analysis', *Journal of Statistical Software*, 42(10). doi: 10.18637/jss.v042.i10.

Litwic, A. *et al.* (2013) 'Epidemiology and burden of osteoarthritis', *British Medical Bulletin*, p. lds038. doi: 10.1093/bmb/lds038.

Lluch, E. *et al.* (2014) 'Evidence for central sensitization in patients with osteoarthritis pain: a systematic literature review', *European Journal of Pain (London, England)*, 18(10), pp. 1367–1375. doi: 10.1002/j.1532-2149.2014.499.x.

Loftis, T., Ellis, B. and Margham, T. (2014) 'Musculoskeletal conditions and multimorbidity'. Arthritis Research UK. Available at: https://www.versusarthritis.org/media/2078/msk-conditions-and-multimorbidity-report.pdf.

Long-term trends in UK employment: 1861 to 2018 - Office for National Statistics (2019). Available at: https://www.ons.gov.uk/economy/nationalaccounts/uksectoraccounts/compendium/economicreview/april2019/longtermtrendsinukemployment1861to2018 (Accessed: 10 July 2019).

Losina E *et al.* (2011) *Impact of obesity and knee osteoarthritis on morbidity and mortality in older Americans*, *Annals of Internal Medicine*. Available at: http://www.annals.org/content/154/4/217.full.pdf+html.

Louati, K. *et al.* (2015) 'Association between diabetes mellitus and osteoarthritis: systematic literature review and meta-analysis', *RMD Open*, 1(1), p. e000077. doi: 10.1136/rmdopen-2015-000077.

Lu, M.-C. *et al.* (2015) 'Higher incidence of rheumatoid arthritis in patients with symptomatic osteoarthritis or osteoarthritis-related surgery: a nationwide, population-based, case–control study in Taiwan', *BMJ Open*, 5(12), p. e008513. doi: 10.1136/bmjopen-2015-008513.

Luong, M.-L. N. *et al.* (2012) 'Social determinants and osteoarthritis outcomes', *Aging health*, 8(4), pp. 413–437. doi: 10.2217/ahe.12.43.

Ma, C. A. and Leung, Y. Y. (2017) 'Exploring the Link between Uric Acid and Osteoarthritis', *Frontiers in Medicine*, 4. doi: 10.3389/fmed.2017.00225.

MacDonald, T. M. *et al.* (1997) 'Association of upper gastrointestinal toxicity of non-steroidal anti-inflammatory drugs with continued exposure: cohort study', *BMJ*, 315(7119), pp. 1333–1337. doi: 10.1136/bmj.315.7119.1333.

Mahmood, Z. and Malghooth, Z. (2019) 'Relationship Of Hips And Knees Osteoarthritis With Bronchial Asthma', *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, 10, p. 64.

Mandalia, V. *et al.* (2005) 'Bone bruising of the knee', *Clinical Radiology*, 60(6), pp. 627–636. doi: 10.1016/j.crad.2005.01.014.

Martel-Pelletier, J. *et al.* (2016) 'Osteoarthritis', *Nature Reviews Disease Primers*, 2, p. 16072. doi: 10.1038/nrdp.2016.72.

Mathers, C. D., Iburg, K. M. and Begg, S. (2006) 'Adjusting for dependent comorbidity in the calculation of healthy life expectancy', *Population Health Metrics*, 4(1). doi: 10.1186/1478-7954-4-4.

McAlindon, T. E. *et al.* (1996) 'Do antioxidant micronutrients protect against the development and progression of knee osteoarthritis?', *Arthritis and Rheumatism*, 39(4), pp. 648–656.

McCarthy, E. M., Sheane, B. J. and Cunnane, G. (2009) 'Greater focus on clinical rheumatology is required for training in internal medicine', *Clinical Rheumatology*, 28(2), pp. 139–143. doi: 10.1007/s10067-008-0997-7.

Medicare, C. for, Baltimore, M. S. 7500 S. B. and Usa, M. (2017) *CC_Main*. Available at: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Chronic-Conditions/CC_Main.html (Accessed: 24 October 2018).

Melzack, R. *et al.* (2001) 'Central neuroplasticity and pathological pain', *Annals of the New York Academy of Sciences*, 933, pp. 157–174. doi: 10.1111/j.1749-6632.2001.tb05822.x.

Menz, H. B. *et al.* (2010) 'Characteristics of primary care consultations for musculoskeletal foot and ankle problems in the UK', *Rheumatology*, 49(7), pp. 1391–1398. doi: 10.1093/rheumatology/keq092.

Miller, G. D. *et al.* (2006) 'Intensive weight loss program improves physical function in older obese adults with knee osteoarthritis', *Obesity (Silver Spring, Md.)*, 14(7), pp. 1219–1230. doi: 10.1038/oby.2006.139.

Morgan, O. J. *et al.* (2019) 'Osteoarthritis in England: Incidence Trends From National Health Service Hospital Episode Statistics', *ACR Open Rheumatology*, 1(8), pp. 493–498. doi: 10.1002/acr2.11071.

Moskowitz, R. W. (2007) *Osteoarthritis: Diagnosis and Medical/surgical Management*. Lippincott Williams & Wilkins.

Moss, P., Knight, E. and Wright, A. (2016) 'Subjects with Knee Osteoarthritis Exhibit Widespread Hyperalgesia to Pressure and Cold', *PLOS ONE*, 11(1), p. e0147526. doi: 10.1371/journal.pone.0147526.

Mukhtar, T. K. *et al.* (2018) 'Factors associated with consultation rates in general practice in England, 2013–2014: a cross-sectional study', *British Journal of General Practice*, 68(670), pp. e370–e377. doi: 10.3399/bjgp18X695981.

Murray, C. *et al.* (2018) 'Population prevalence and distribution of ankle pain and symptomatic radiographic ankle osteoarthritis in community dwelling older adults: A systematic review and cross-sectional study', *PLOS ONE*. Edited by M. J. Lammi, 13(4), p. e0193662. doi: 10.1371/journal.pone.0193662.

Muthén, B. and Asparouhov, T. (2009) *Growth mixture modeling: analysis with non-Gaussian random effects*. G. Fitzmaurice, M. Davidian, G. Verbeke, G. Molenberghs (Eds.). FL: Chapman & Hall/CRC Press, Boca Raton.

Muthen, B. and Muthen, L. K. (2000) 'Integrating Person-Centered and Variable-Centered Analyses: Growth Mixture Modeling With Latent Trajectory Classes', *Alcoholism: Clinical and Experimental Research*, 24(6), pp. 882–891. doi: 10.1111/j.1530-0277.2000.tb02070.x.

National Health Services, UK (no date). Available at: https://www.nhs.uk/ (Accessed: 13 February 2019).

Neogi, T. *et al.* (2006) 'Low vitamin K status is associated with osteoarthritis in the hand and knee', *Arthritis and Rheumatism*, 54(4), pp. 1255–1261. doi: 10.1002/art.21735.

Neogi, T. *et al.* (2009) 'Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies', *BMJ*, 339, p. b2844. doi: 10.1136/bmj.b2844.

Neogi, T. (2013) 'The epidemiology and impact of pain in osteoarthritis', *Osteoarthritis and Cartilage*, 21(9), pp. 1145–1153. doi: 10.1016/j.joca.2013.03.018.

Nevitt, M. C. *et al.* (2010) 'High systemic bone mineral density increases the risk of incident knee OA and joint space narrowing, but not radiographic progression of existing knee OA: the MOST study', *Annals of the Rheumatic Diseases*, 69(1), pp. 163–168. doi: 10.1136/ard.2008.099531.

Nguyen, H. *et al.* (2019) 'Prevalence of multimorbidity in community settings: A systematic review and meta-analysis of observational studies', *Journal of Comorbidity*, 9, p. 2235042X1987093. doi: 10.1177/2235042X19870934.

Nguyen, U.-S. D. T. *et al.* (2011) 'Increasing Prevalence of Knee Pain and Symptomatic Knee Osteoarthritis: Survey and Cohort Data', *Annals of Internal Medicine*, 155(11), p. 725. doi: 10.7326/0003-4819-155-11-201112060-00004.

NHS (2016) *Quality of Framework*, *QOF 2016/17*. Available at: https://qof.digital.nhs.uk/ (Accessed: 24 October 2018).

NHS (2018) *Body mass index (BMI) in UK population*, *nhs.uk*. Available at: https://www.nhs.uk/common-health-questions/lifestyle/what-is-the-body-mass-index-bmi/ (Accessed: 5 June 2019).

NICE (2014) *Overview | Osteoarthritis: care and management | Guidance | NICE*. NICE. Available at: https://www.nice.org.uk/guidance/cg177 (Accessed: 28 October 2020).

Nielsen, B. (2015) 'apc: An R Package for Age-Period-Cohort Analysis', 7(2). Available at: https://journal.r-project.org/archive/2015/RJ-2015-020/RJ-2015-020.pdf.

Nieves-Plaza, M. *et al.* (2013) 'Association of hand or knee osteoarthritis with diabetes mellitus in a population of Hispanics from Puerto Rico', *Journal of Clinical Rheumatology: Practical Reports on Rheumatic & Musculoskeletal Diseases*, 19(1), pp. 1–6. doi: 10.1097/RHU.0b013e31827cd578.

Njeze, G. E. (2013) 'Gallstones', *Nigerian Journal of Surgery: Official Publication of the Nigerian Surgical Research Society*, 19(2), pp. 49–55. doi: 10.4103/1117-6806.119236.

Noble, B. (2003) 'Bone microdamage and cell apoptosis', *European Cells and Materials*, 6, pp. 46–56. doi: 10.22203/eCM.v006a05.

Nuesch, E. *et al.* (2011) 'All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study', *BMJ*, 342(mar08 2), pp. d1165–d1165. doi: 10.1136/bmj.d1165.

Nygård, L. H. *et al.* (2017) 'The effect of non-steroidal anti-inflammatory drugs on risk of benign prostatic hyperplasia', *The Prostate*, 77(9), pp. 1029–1035. doi: 10.1002/pros.23359.

Nylund, K. *et al.* (2007) 'Subtypes, Severity, and Structural Stability of Peer Victimization: What Does Latent Class Analysis Say?', *Child Development*, 78(6), pp. 1706–1722. doi: 10.1111/j.1467-8624.2007.01097.x.

Nylund, K. L., Asparouhov, T. and Muthén, B. O. (2007) 'Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study', *Structural Equation Modeling: A Multidisciplinary Journal*, 14(4), pp. 535–569. doi: 10.1080/10705510701575396.

Obesity and overweight (no date) *World Health Organization*. Available at: http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (Accessed: 22 October 2018).

Office For National Statistics, National Records Of Scotland, and Northern Ireland Statistics And Research Agency (2016) '2011 Census aggegate data (Data downloaded: 1 June 2016)'. UK Data Service. doi: 10.5257/census/aggregate-2011-1.

Ogle, D., removal()), P. W. (Added method='Burnham' to and dunnTest()), A. D. (Provided base functionality of (2019) *FSA*. Available at: https://CRAN.R-project.org/package=FSA (Accessed: 5 June 2019).

Oliveria, S. A. *et al.* (1995) 'Incidence of symptomatic hand, hip, and knee osteoarthritis among patients in a health maintenance organization', *Arthritis and Rheumatism*, 38(8), pp. 1134–1141.

O'Neill, T. W. and Felson, D. T. (2018) 'Mechanisms of Osteoarthritis (OA) Pain', *Current Osteoporosis Reports*, 16(5), pp. 611–616. doi: 10.1007/s11914-018-0477-1.

Osteoarthritis - NICE CKS (no date). Available at: https://cks.nice.org.uk/osteoarthritis#!topicsummary (Accessed: 21 October 2018).

Østerås, N. *et al.* (2019) 'Implementing a structured model for osteoarthritis care in primary healthcare: A stepped-wedge cluster-randomised trial', *PLOS Medicine*. Edited by C. Nguyen, 16(10), p. e1002949. doi: 10.1371/journal.pmed.1002949.

Papatheodoridis, G. V., Sougioultzis, S. and Archimandritis, A. J. (2006) 'Effects of Helicobacter pylori and Nonsteroidal Anti-Inflammatory Drugs on Peptic Ulcer Disease: A Systematic Review', *Clinical Gastroenterology and Hepatology*, 4(2), pp. 130–142. doi: 10.1016/j.cgh.2005.10.006.

Park, B., Lee, H. A. and Park, H. (2019) 'Use of latent class analysis to identify multimorbidity patterns and associated factors in Korean adults aged 50 years and older', *PLOS ONE*. Edited by K. Latham-Mintus, 14(11), p. e0216259. doi: 10.1371/journal.pone.0216259.

Parkinson, L., Waters, D. L. and Franck, L. (2017) 'Systematic review of the impact of osteoarthritis on health outcomes for comorbid disease in older people', *Osteoarthritis and Cartilage*, 25(11), pp. 1751–1770. doi: 10.1016/j.joca.2017.07.008.

Parmelee, P. A., Tighe, C. A. and Dautovich, N. D. (2015) 'Sleep Disturbance in Osteoarthritis: Linkages with Pain, Disability and Depressive Symptoms', *Arthritis care & research*, 67(3), pp. 358–365. doi: 10.1002/acr.22459.

Parsons, S. *et al.* (2011) 'A HEAVY BURDEN: The occurrence and impact of musculoskeletal conditions in the United Kingdom today'. The University of Manchester. Available at: https://www.escholar.manchester.ac.uk/api/datastream?publicationPid=uk-ac-man-scw:123774&datastreamId=FULL-TEXT.PDF.

Payne, R. A. *et al.* (2013) 'The effect of physical multimorbidity, mental health conditions and socioeconomic deprivation on unplanned admissions to hospital: a retrospective cohort study', *Canadian Medical Association Journal*, 185(5), pp. E221–E228. doi: 10.1503/cmaj.121349.

Pazzi, P. *et al.* (1998) 'Nonsteroidal antiinflammatory drug use and gallstone disease prevalence: a case-control study', *The American Journal of Gastroenterology*, 93(9), pp. 1420–1424. doi: 10.1111/j.1572-0241.1998.00453.x.

Pearce, F. *et al.* (2013) 'Does Smoking Reduce the Progression of Osteoarthritis? Meta-Analysis of Observational Studies', *Arthritis Care & Research*, 65(7), pp. 1026–1033. doi: 10.1002/acr.21954.

Persson, M. S. M. *et al.* (2018) 'The relative efficacy of topical non-steroidal anti-inflammatory drugs and capsaicin in osteoarthritis: a network meta-analysis of randomised controlled trials', *Osteoarthritis and Cartilage*. doi: 10.1016/j.joca.2018.08.008.

Piette, J. D. and Kerr, E. A. (2006) 'The Impact of Comorbid Chronic Conditions on Diabetes Care', *Diabetes Care*, 29(3), pp. 725–731. doi: 10.2337/diacare.29.03.06.dc05-2078.

Plotnikoff, R. *et al.* (2015) 'Osteoarthritis prevalence and modifiable factors: a population study', *BMC Public Health*, 15. doi: 10.1186/s12889-015-2529-0.

Popescu, D., Andrenscou, D. and Babes, P. A. (2018) 'The Association Between Helicobacter Pylori Infection and Liver and Biliary Tract Disorders', *Current Health Sciences Journal*, (2), pp. 186–191. doi: 10.12865/CHSJ.44.02.16.

Prados-Torres, A. *et al.* (2014) 'Multimorbidity patterns: a systematic review', *Journal of Clinical Epidemiology*, 67(3), pp. 254–266. doi: 10.1016/j.jclinepi.2013.09.021.

Prieto-Alhambra, D. *et al.* (2014) 'Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints', *Annals of the Rheumatic Diseases*, 73(9), pp. 1659–1664. doi: 10.1136/annrheumdis-2013-203355.

Prior, J. A. *et al.* (2012) 'Comorbidity Cohort (2C) study: cardiovascular disease severity and comorbid osteoarthritis in primary care', *BMC health services research*, 12, p. 295. doi: 10.1186/1472-6963-12-295.

Proust-Lima, C. *et al.* (2020) *lcmm: Extended Mixed Models Using Latent Classes and Latent Processes. R package version: 1.9.2*. Available at: https://cran.r-project.org/package=lcmm.

QGIS Geographic Information System. Open Source Geospatial Foundation Project. (2016). QGIS Development Team. Available at: http://qgis.osgeo.org.

Quality and Outcome Framework (QOF) (no date). Available at: https://digital.nhs.uk/article/8910/Quality-and-Outcome-Framework-QOF-Indicators-No-Longer-In-

QOF-INLIQ-Enhanced-Services-ES-Vaccinations-and-Immunisations-V-I-and-GMS-Core-Contract-CC-extraction-specifications-business-rules- (Accessed: 21 March 2018).

Quan, H. *et al.* (2011) 'Updating and Validating the Charlson Comorbidity Index and Score for Risk Adjustment in Hospital Discharge Abstracts Using Data From 6 Countries', *American Journal of Epidemiology*, 173(6), pp. 676–682. doi: 10.1093/aje/kwq433.

Quinones A.R. *et al.* (2011) 'How does the trajectory of multimorbidity vary across Black, White, and Mexican Americans in middle and old age?', *The journals of gerontology. Series B, Psychological sciences and social sciences*, 66(6), pp. 739–749.

R: A language and evironment for statistical computing. (no date). Vienna, Austria.: R Development Core Team (2008) (R Foundation for Statistical Computing,). Available at: https://www.r-project.org/ (Accessed: 13 February 2019).

Rahman, M. M. *et al.* (2013) 'The relationship between osteoarthritis and cardiovascular disease in a population health survey: a cross-sectional study', *BMJ Open*, 3(5), p. e002624. doi: 10.1136/bmjopen-2013-002624.

Rahman, M. M. *et al.* (2014) 'Osteoarthritis Incidence and Trends in Administrative Health Records from British Columbia, Canada', *The Journal of Rheumatology*, 41(6), pp. 1147–1154. doi: 10.3899/jrheum.131011.

Rawool, V. W. and Harrington, B. T. (2007) 'Middle Ear Admittance and Hearing Abnormalities in Individuals with Osteoarthritis', *Audiology and Neurotology*, 12(2), pp. 127–136. doi: 10.1159/000097799.

Reeuwijk, K. G. *et al.* (2010) 'Osteoarthritis of the hip or knee: which coexisting disorders are disabling?', *Clinical Rheumatology*, 29(7), pp. 739–747. doi: 10.1007/s10067-010-1392-8.

Ren, K. *et al.* (2017) 'Association of *ADAM12* gene polymorphisms with knee osteoarthritis susceptibility', *Oncotarget*, 8(44). doi: 10.18632/oncotarget.20772.

Riccardo, C., Fabio, C. and Pietro, R. (2017) 'Knee Osteoarthritis after Reconstruction of Isolated Anterior Cruciate Ligament Injuries: A Systematic Literature Review', *Joints*, 5(1), pp. 39–43. doi: 10.1055/s-0037-1601409.

Ricci-Cabello, I. *et al.* (2015) 'Impact of the Prevalence of Concordant and Discordant Conditions on the Quality of Diabetes Care in Family Practices in England', *The Annals of Family Medicine*, 13(6), pp. 514–522. doi: 10.1370/afm.1848.

Roberts, E. *et al.* (2016) 'Paracetamol: not as safe as we thought? A systematic literature review of observational studies', *Annals of the Rheumatic Diseases*, 75(3), pp. 552–559. doi: 10.1136/annrheumdis-2014-206914.

Robinson, W. H. *et al.* (2016) 'Low-grade inflammation as a key mediator of the pathogenesis of osteoarthritis', *Nature Reviews Rheumatology*, 12(10), pp. 580–592. doi: 10.1038/nrrheum.2016.136.

Roddy, E. and Menz, H. B. (2018) 'Foot osteoarthritis: latest evidence and developments', *Therapeutic Advances in Musculoskeletal Disease*, 10(4), pp. 91–103. doi: 10.1177/1759720X17753337.

de Rooij, M. *et al.* (2014) 'Development of comorbidity-adapted exercise protocols for patients with knee osteoarthritis', *Clinical Interventions in Aging*, 9, pp. 829–842. doi: 10.2147/CIA.S55705.

Rousseeuw, P. J. (1987) 'Silhouettes: a graphical aid to the interpretation and validation of cluster analysis.', 20, pp. 53–65.

Ruiz-Medrano, E., Espinosa-Ortega, H. F. and Arce-Salinas, C. A. (2019) 'The effect of concomitant hand osteoarthritis on pain and disease activity in patients with rheumatoid arthritis', *Clinical Rheumatology*. doi: 10.1007/s10067-019-04574-6.

Runhaar, J. *et al.* (2011) 'A systematic review on changed biomechanics of lower extremities in obese individuals: a possible role in development of osteoarthritis', *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity*, 12(12), pp. 1071–1082. doi: 10.1111/j.1467-789X.2011.00916.x.

Ryoo, J. H. *et al.* (2018) 'Longitudinal Model Building Using Latent Transition Analysis: An Example Using School Bullying Data', *Frontiers in Psychology*, 9. doi: 10.3389/fpsyg.2018.00675.

van Saase, J. L. *et al.* (1989) 'Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations.', *Annals of the Rheumatic Diseases*, 48(4), pp. 271–280.

Salisbury, C. *et al.* (2018) 'Management of multimorbidity using a patient-centred care model: a pragmatic cluster-randomised trial of the 3D approach', *The Lancet*, 392(10141), pp. 41–50. doi: 10.1016/S0140-6736(18)31308-4.

Sarmanova, A. *et al.* (2017) 'Association between ultrasound-detected synovitis and knee pain: a population-based case–control study with both cross-sectional and follow-up data', *Arthritis Research & Therapy*, 19(1). doi: 10.1186/s13075-017-1486-7.

Sarmanova, A. *et al.* (2018) 'Contribution of central and peripheral risk factors to prevalence, incidence and progression of knee pain: a community-based cohort study', *Osteoarthritis and Cartilage*, 26(11), pp. 1461–1473. doi: 10.1016/j.joca.2018.07.013.

Saxena, A. *et al.* (2017) 'A review of clustering techniques and developments', *Neurocomputing*, 267, pp. 664–681. doi: 10.1016/j.neucom.2017.06.053.

Scherer, M. *et al.* (2016) 'Association between multimorbidity patterns and chronic pain in elderly primary care patients: a cross-sectional observational study', *BMC family practice*, 17, p. 68. doi: 10.1186/s12875-016-0468-1.

Schett, G. *et al.* (2013) 'Diabetes Is an Independent Predictor for Severe Osteoarthritis: Results from a longitudinal cohort study', *Diabetes Care*, 36(2), pp. 403–409. doi: 10.2337/dc12-0924.

Schiphof, D., Boers, M. and Bierma-Zeinstra, S. M. A. (2008) 'Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis', *Annals of the Rheumatic Diseases*, 67(7), pp. 1034–1036. doi: 10.1136/ard.2007.079020.

Schirmer, B. D., Winters, K. L. and Edlich, R. F. (2005) 'Cholelithiasis and cholecystitis', *Journal of Long-Term Effects of Medical Implants*, 15(3), pp. 329–338.

Schlenk, E. A. *et al.* (2011) 'Improving physical activity and function in overweight and obese older adults with osteoarthritis of the knee: a feasibility study', *Rehabilitation Nursing: The Official Journal of the Association of Rehabilitation Nurses*, 36(1), pp. 32–42.

van de Schoot, R. *et al.* (2017) 'The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies', *Structural Equation Modeling: A Multidisciplinary Journal*, 24(3), pp. 451–467. doi: 10.1080/10705511.2016.1247646.

Schram, B. *et al.* (2020) 'Risk factors for development of lower limb osteoarthritis in physically demanding occupations: A narrative umbrella review', *Journal of Occupational Health*, 62(1). doi: 10.1002/1348-9585.12103.

Shi, G. and Zhang, K. (2018) 'Proteomics analysis of osteoporosis with knee osteoarthritis', *Journal of Clinical Rehabilitative Tissue Engineering Research*, 22(36), pp. 5753–5759. doi: 10.3969/j.issn.2095-4344.0376.

Siemons, L. *et al.* (2013) 'Validity of summing painful joint sites to assess joint-pain comorbidity in hip or knee osteoarthritis', *BMC Musculoskeletal Disorders*, 14(1). doi: 10.1186/1471-2474-14-234.

Simoes D. *et al.* (2018) 'The population impact of rheumatic and musculoskeletal diseases in relation to other non-communicable disorders: comparing two estimation approaches', *Rheumatology International*, 38(5), pp. 905–915. doi: 10.1007/s00296-018-3990-8.

Simopoulou, T. *et al.* (2007) 'Differential expression of leptin and leptin's receptor isoform (Ob-Rb) mRNA between advanced and minimally affected osteoarthritic cartilage; effect on cartilage metabolism', *Osteoarthritis and Cartilage*, 15(8), pp. 872–883. doi: 10.1016/j.joca.2007.01.018.

Singer, L. *et al.* (2019) 'Social determinants of multimorbidity and multiple functional limitations among the ageing population of England, 2002–2015', *SSM - Population Health*, 8, p. 100413. doi: 10.1016/j.ssmph.2019.100413.

Skousgaard, S. *et al.* (2018) '308 Occupational risk factors for hip and knee osteoarthritis – evidence of gene-exposure interaction: a co-twin control study in danish twins', in *Musculoskeletal Disorders*. *32nd Triennial Congress of the International Commission on Occupational Health (ICOH), Dublin, Ireland, 29th April to 4th May 2018*, BMJ Publishing Group Ltd, p. A261.3-A262. doi: 10.1136/oemed-2018-ICOHabstracts.748.

Smith, D. *et al.* (2014) 'Chronic Pain and Mortality: A Systematic Review', *PLoS ONE*. Edited by D. Zaykin, 9(6), p. e99048. doi: 10.1371/journal.pone.0099048.

Smith, S. M. *et al.* (2012) 'Managing patients with multimorbidity: systematic review of interventions in primary care and community settings', *BMJ*, 345, p. e5205. doi: 10.1136/bmj.e5205.

Smith, S. M. *et al.* (2016) 'Interventions for improving outcomes in patients with multimorbidity in primary care and community settings', *The Cochrane Database of Systematic Reviews*, 3, p. CD006560. doi: 10.1002/14651858.CD006560.pub3.

Snijders, G. F. *et al.* (2011) 'Fatigue in knee and hip osteoarthritis: the role of pain and physical function', *Rheumatology*, 50(10), pp. 1894–1900. doi: 10.1093/rheumatology/ker201.

Spector, T. D. *et al.* (1996) 'Genetic influences on osteoarthritis in women: a twin study', *BMJ (Clinical research ed.)*, 312(7036), pp. 940–943.

Spector, T. D. and Campion, G. D. (1989) 'Generalised osteoarthritis: a hormonally mediated disease.', *Annals of the Rheumatic Diseases*, 48(6), pp. 523–527.

'StataCorp. 2011. Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.' (no date).

Statistics, c=AU; o=Commonwealth of A. ou=Australian B. of (2015) *Main Features - Arthritis and osteoporosis*. Available at: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/4364.0.55.001~2014-15~Main%20Features~Arthritis%20and%20osteoporosis~8 (Accessed: 21 October 2018).

'Statistics on Obesity, Physical Activity and Diet, England 2018' (2018). NHS England. Available at: https://files.digital.nhs.uk/publication/0/0/obes-phys-acti-diet-eng-2018-rep.pdf.

Staud, R. (2011) 'Evidence for Shared Pain Mechanisms in Osteoarthritis, Low Back Pain, and Fibromyalgia', *Current Rheumatology Reports*, 13(6), pp. 513–520. doi: 10.1007/s11926-011-0206-6.

Sterling, R. K. *et al.* (1995) 'Effect of NSAIDs on gallbladder bile composition', *Digestive Diseases and Sciences*, 40(10), pp. 2220–2226. doi: 10.1007/BF02209010.

Strauss, V. Y. *et al.* (2014) 'Distinct trajectories of multimorbidity in primary care were identified using latent class growth analysis', *Journal of Clinical Epidemiology*, 67(10), pp. 1163–1171. doi: 10.1016/j.jclinepi.2014.06.003.

Struckmann, V. *et al.* (2018) 'Relevant models and elements of integrated care for multi-morbidity: Results of a scoping review', *Health Policy (Amsterdam, Netherlands)*, 122(1), pp. 23–35. doi: 10.1016/j.healthpol.2017.08.008.

Stubbs, B. *et al.* (2016a) 'Prevalence of depressive symptoms and anxiety in osteoarthritis: a systematic review and meta-analysis', *Age and Ageing*, 45(2), pp. 228–235. doi: 10.1093/ageing/afw001.

Stubbs, B. *et al.* (2016b) 'Prevalence of depressive symptoms and anxiety in osteoarthritis: a systematic review and meta-analysis', *Age and Ageing*, 45(2), pp. 228–235. doi: 10.1093/ageing/afw001.

Suri, P. *et al.* (2010) 'Low back pain and other musculoskeletal pain comorbidities in individuals with symptomatic osteoarthritis of the knee: data from the osteoarthritis initiative', *Arthritis Care & Research*, 62(12), pp. 1715–1723. doi: 10.1002/acr.20324.

Swain, S. *et al.* (2019) 'Comorbidities in Osteoarthritis: A systematic review and meta-analysis of observational studies'. doi: 10.1002/acr.24008.

Swain, S. *et al.* (2020) 'Trends in incidence and prevalence of osteoarthritis in the United Kingdom: findings from the Clinical Practice Research Datalink (CPRD)', *Osteoarthritis and Cartilage*. doi: 10.1016/j.joca.2020.03.004.

Symmons, D., Mathers, C. and Pfleger, B. (2003) 'Global burden of osteoarthritis in the year 2000.' World Health Organization, Geneva. Available at: http://www3.who.int/whosis/menu.cfm?path=evidence,burden,burden_gbd2000docs&language=english.

Szoeke, C. E. I. *et al.* (2006) 'Factors affecting the prevalence of osteoarthritis in healthy middle-aged women: data from the longitudinal Melbourne Women's Midlife Health Project', *Bone*, 39(5), pp. 1149–1155. doi: 10.1016/j.bone.2006.05.016.

Takatsu, M. *et al.* (2005) 'Ear involvement in patients with rheumatoid arthritis', *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 26(4), pp. 755–761.

Tate, A. R. *et al.* (2017) '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*, 7(1), p. e012905. doi: 10.1136/bmjopen-2016-012905.

Teh, R. O. *et al.* (2018) 'Patterns of multi-morbidity and prediction of hospitalisation and all-cause mortality in advanced age', *Age and Ageing*, 47(2), pp. 261–268. doi: 10.1093/ageing/afx184.

Thomas, E., Peat, G. and Croft, P. (2014) 'Defining and mapping the person with osteoarthritis for population studies and public health', *Rheumatology*, 53(2), pp. 338–345. doi: 10.1093/rheumatology/ket346.

Timmins, K. A. *et al.* (2017) 'Running and Knee Osteoarthritis: A Systematic Review and Meta-analysis', *The American Journal of Sports Medicine*, 45(6), pp. 1447–1457. doi: 10.1177/0363546516657531.

Towheed, T. *et al.* (2006) 'Acetaminophen for osteoarthritis', *Cochrane Database of Systematic Reviews*. Edited by Cochrane Musculoskeletal Group. doi: 10.1002/14651858.CD004257.pub2.

Tramèr, M. R. *et al.* (2000) 'Quantitative estimation of rare adverse events which follow a biological progression: a new model applied to chronic NSAID use':, *Pain*, 85(1), pp. 169–182. doi: 10.1016/S0304-3959(99)00267-5.

Tran, G. *et al.* (2016) 'Does sports participation (including level of performance and previous injury) increase risk of osteoarthritis? A systematic review and meta-analysis', *Br J Sports Med*, p. bjsports-2016-096142. doi: 10.1136/bjsports-2016-096142.

Trelle, S. *et al.* (2011) 'Cardiovascular safety of non-steroidal anti-inflammatory drugs: network meta-analysis', *BMJ*, 342(jan11 1), pp. c7086–c7086. doi: 10.1136/bmj.c7086.

Turkiewicz, A. *et al.* (2015) 'Prevalence of knee pain and knee OA in southern Sweden and the proportion that seeks medical care', *Rheumatology (Oxford, England)*, 54(5), pp. 827–835. doi: 10.1093/rheumatology/keu409.

Turkiewicz, A. *et al.* (2016) 'All-cause Mortality in Knee and Hip Osteoarthritis and Rheumatoid Arthritis':, *Epidemiology*, 27(4), pp. 479–485. doi: 10.1097/EDE.0000000000000477.

Vaccarino, V. *et al.* (2007) 'Depression, Inflammation, and Incident Cardiovascular Disease in Women With Suspected Coronary Ischemia', *Journal of the American College of Cardiology*, 50(21), pp. 2044–2050. doi: 10.1016/j.jacc.2007.07.069.

Versus Arthritis (2009) 'Musculoskeletal matters bulletin 1: what do general practitioners see.' Keele University, Keele. Available at: https://www.keele.ac.uk/media/keeleuniversity/ri/primarycare/bulletins/MusculoskeletalMatters1.pdf (Accessed: 23 August 2020).

Versus Arthritis (2013) 'OSTEOARTHRITIS IN GENERAL PRACTICE; Data and perspectives'. Arthritis Research UK.

Versus Arthritis (2016) 'MSK conditions and multimorbidity'. Versus Arthritis. Available at: https://www.versusarthritis.org/media/2078/msk-conditions-and-multimorbidity-report.pdf.

Versus Arthritis (2019) *Musculoskeletal Calculator. A visual representation of the prevalence of MSK conditions across the UK.* Available at: https://www.arthritisresearchuk.org/arthritis-information/data-and-statistics/musculoskeletal-calculator.aspx (Accessed: 1 May 2018).

Vetrano, D. L. *et al.* (2020) 'Twelve-year clinical trajectories of multimorbidity in a population of older adults', *Nature Communications*, 11(1). doi: 10.1038/s41467-020-16780-x.

Vincent, H. K. *et al.* (2012) 'Obesity and weight loss in the treatment and prevention of osteoarthritis', *PM & R: the journal of injury, function, and rehabilitation*, 4(5 Suppl), pp. S59-67. doi: 10.1016/j.pmrj.2012.01.005.

Von Korff, M. *et al.* (1992) 'Grading the severity of chronic pain', *Pain*, 50(2), pp. 133–149.

Vos, R. *et al.* (2015) 'Trajectories of multimorbidity: exploring patterns of multimorbidity in patients with more than ten chronic health problems in life course', *BMC Family Practice*, 16(1), p. 2. doi: 10.1186/s12875-014-0213-6.

Vos, T. *et al.* (2012) '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*, 380(9859), pp. 2163–2196. doi: 10.1016/S0140-6736(12)61729-2.

Wandel, S. *et al.* (2010) 'Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis', *BMJ*, 341, p. c4675. doi: 10.1136/bmj.c4675.

Wang C. -C *et al.* (2011) *A highly organized three-dimensional alginate scaffold for cartilage tissue engineering prepared by microfluidic technology*, *Biomaterials*.

Wang, C. Y. *et al.* (2009) 'The Contribution of Longitudinal Comorbidity Measurements to Survival Analysis':, *Medical Care*, 47(7), pp. 813–821. doi: 10.1097/MLR.0b013e318197929c.

Wang, H. *et al.* (2016) 'Osteoarthritis and the risk of cardiovascular disease: a meta-analysis of observational studies', *Scientific Reports*, 6. doi: 10.1038/srep39672.

Wang, Q. *et al.* (2011) 'Identification of a central role for complement in osteoarthritis', *Nature Medicine*, 17(12), pp. 1674–1679. doi: 10.1038/nm.2543.

Wang Shunye (2013) 'An improved k-means clustering algorithm based on dissimilarity', in *Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)*. *Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)*, pp. 2629–2633. doi: 10.1109/MEC.2013.6885476.

Wang, W. *et al.* (2016) 'The low back pain in patients with hip osteoarthritis: current knowledge on the diagnosis, mechanism and treatment outcome', *Annals of Joint*, 1(4). Available at: http://aoj.amegroups.com/article/view/3544 (Accessed: 30 September 2019).

Weber, A. *et al.* (2019) 'Association between osteoarthritis and increased risk of dementia: A systemic review and meta-analysis', *Medicine*, 98(10), p. e14355. doi: 10.1097/MD.0000000000014355.

Weiner, J. P. *et al.* (1991) 'Development and application of a population-oriented measure of ambulatory care case-mix', *Medical Care*, 29(5), pp. 452–472.

Whitehead, W. E. *et al.* (2007) 'Comorbidity in Irritable Bowel Syndrome', *The American Journal of Gastroenterology*, 102(12), pp. 2767–2776. doi: 10.1111/j.1572-0241.2007.01540.x.

WHO report (2019) *Obesity and overweight*. Available at: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (Accessed: 28 October 2020).

WHO Scientific Group on the Burden of Musculoskeletal Conditions at the Start of the New Millennium (2003) 'The burden of musculoskeletal conditions at the start of the new millennium', *World Health Organization Technical Report Series*, 919, pp. i–x, 1–218, back cover.

Willadsen, T. *et al.* (2018) 'Multimorbidity and mortality: A 15-year longitudinal registry-based nationwide Danish population study', *Journal of Comorbidity*, 8(1), p. 2235042X1880406. doi: 10.1177/2235042X18804063.

Williams, M. F. *et al.* (2016) 'Type 2 diabetes and osteoarthritis: a systematic review and meta-analysis', *Journal of Diabetes and Its Complications*, 30(5), pp. 944–950. doi: 10.1016/j.jdiacomp.2016.02.016.

Williams, T. *et al.* (2012) 'Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource', *Therapeutic Advances in Drug Safety*, 3(2), pp. 89–99. doi: 10.1177/2042098611435911.

Wise, B. *et al.* (2010) 'Psychological factors and their relation to osteoarthritis pain', *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society*, 18(7), pp. 883–887. doi: 10.1016/j.joca.2009.11.016.

Wluka, A. E. *et al.* (2002) 'Supplementary vitamin E does not affect the loss of cartilage volume in knee osteoarthritis: a 2 year double blind randomized placebo controlled study', *The Journal of Rheumatology*, 29(12), pp. 2585–2591.

Wojdasiewicz, P., Poniatowski, Ł. A. and Szukiewicz, D. (2014) 'The Role of Inflammatory and Anti-Inflammatory Cytokines in the Pathogenesis of Osteoarthritis', *Mediators of Inflammation*, 2014, pp. 1–19. doi: 10.1155/2014/561459.

Wolfe, F. *et al.* (1996) 'Back pain in osteoarthritis of the knee', *Arthritis Care and Research: The Official Journal of the Arthritis Health Professions Association*, 9(5), pp. 376–383.

Woolf, A. D. and Pfleger, B. (2003) 'Burden of major musculoskeletal conditions', *Bulletin of the World Health Organization*, 81(9), pp. 646–656.

Woolf, C. J. and Salter, M. W. (2000) 'Neuronal plasticity: increasing the gain in pain', *Science (New York, N.Y.)*, 288(5472), pp. 1765–1769.

'World Population Prospects: The 2017 Revision' (2017). United Nations Economic and Social Affairs. Available at: https://esa.un.org/unpd/wpp/Publications/Files/WPP2017_KeyFindings.pdf.

Wshah, A. *et al.* (2018) 'Prevalence of osteoarthritis in individuals with COPD: a systematic review', *International Journal of Chronic Obstructive Pulmonary Disease*, 13, pp. 1207–1216. doi: 10.2147/COPD.S158614.

Wylde, V. *et al.* (2017) 'Post-operative patient-related risk factors for chronic pain after total knee replacement: a systematic review', *BMJ Open*, 7(11), p. e018105. doi: 10.1136/bmjopen-2017-018105.

Xu, X. *et al.* (2018) 'Progression of diabetes, heart disease, and stroke multimorbidity in middle-aged women: A 20-year cohort study', *PLOS Medicine*. Edited by S. R. Steinhubl, 15(3), p. e1002516. doi: 10.1371/journal.pmed.1002516.

Yu, D. *et al.* (2015) 'Annual consultation incidence of osteoarthritis estimated from population-based health care data in England', *Rheumatology (Oxford, England)*, 54(11), pp. 2051–2060. doi: 10.1093/rheumatology/kev231.

Yu, D. *et al.* (2017) 'Population trends in the incidence and initial management of osteoarthritis: age-period-cohort analysis of the Clinical Practice Research Datalink, 1992–2013', *Rheumatology*, 56(11), pp. 1902–1917. doi: 10.1093/rheumatology/kex270.

Yu, D., Jordan, K. P. and Peat, G. (2018) 'Underrecording of osteoarthritis in United Kingdom primary care electronic health record data', *Clinical Epidemiology*, 10, pp. 1195–1201. doi: 10.2147/CLEP.S160059.

Zak, M. and Pasiyeshvili, L. (2016) 'Chronic gastritis clinical features and stomach functional state during nonsteroidal anti-inflammatory drugs administration in patients with osteoarthritis', *EUREKA: Health Sciences*, 5, pp. 17–22. doi: 10.21303/2504-5679.2016.00178.

Zak, M. Y. *et al.* (2019) 'Medico-social value of osteoarthritis. secondary prevention and treatment of osteoarthritis in comorbidity with chronic gastritis', *Wiadomosci Lekarskie (Warsaw, Poland: 1960)*, 72(5 cz 2), pp. 1064–1067.

Zambon, S., Siviero, P., Denkinger, M., Limongi, F., Castell, M. V., van der Pas, S., Otero, Á., Edwards, M. H., Peter, R., Pedersen, N. L., Sánchez-Martinez, M., Dennison, E. M., Gesmundo, A., Schaap, L. A., Deeg, D. J. H., van Schoor, N. M., Maggi, S., and EPOSA Research Group (2015) 'Osteoarthritis, comorbidity and pain: Their role in determining functional limitations in older populations (European project on Osteoarthritis)', *Arthritis Care & Research*. doi: 10.1002/acr.22755.

Zambon, S., Siviero, P., Denkinger, M., Limongi, F., Castell, M. V., van der Pas, S., Otero, Á., Edwards, M. H., Peter, R., Pedersen, N. L., Sánchez-Martinez, M., Dennison, E. M., Gesmundo, A., Schaap, L. A., Deeg, D. J. H., van Schoor, N. M., Maggi, S. and EPOSA Research Group (2015) 'Osteoarthritis, comorbidity and pain: Their role in determining functional limitations in older populations (European project on Osteoarthritis)', *Arthritis Care & Research*. doi: 10.1002/acr.22755.

Zemedikun D.T. *et al.* (2018) 'Patterns of Multimorbidity in Middle-Aged and Older Adults: An Analysis of the UK Biobank Data', *Mayo Clinic Proceedings*, 93(7), pp. 857–866. doi: 10.1016/j.mayocp.2018.02.012.

Zhang, M. *et al.* (2019) 'Musculoskeletal Symptomatic Areas After Total Knee Replacement for Osteoarthritis', *ACR Open Rheumatology*. doi: 10.1002/acr2.11055.

Zhang, W. *et al.* (2008) 'OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines', *Osteoarthritis and Cartilage*, 16(2), pp. 137–162. doi: 10.1016/j.joca.2007.12.013.

Zhang, W., Jones, A. and Doherty, M. (2004) 'Does paracetamol (acetaminophen) reduce the pain of osteoarthritis? A meta-analysis of randomised controlled trials', *Annals of the Rheumatic Diseases*, 63(8), pp. 901–907. doi: 10.1136/ard.2003.018531.

Zhang, Y. and Jordan, J. M. (2010) 'Epidemiology of Osteoarthritis', *Clinics in Geriatric Medicine*, 26(3), pp. 355–369. doi: 10.1016/j.cger.2010.03.001.

Zheng, H. and Chen, C. (2015) 'Body mass index and risk of knee osteoarthritis: systematic review and meta-analysis of prospective studies', *BMJ open*, 5(12), p. e007568. doi: 10.1136/bmjopen-2014-007568.

Zhou, Z.-Y. *et al.* (2014) 'Body mass index and knee osteoarthritis risk: a dose-response meta-analysis', *Obesity (Silver Spring, Md.)*, 22(10), pp. 2180–2185. doi: 10.1002/oby.20835.

Zhu, Y. *et al.* (2020) 'Characteristics, service use and mortality of clusters of multimorbid patients in England: a population-based study', *BMC Medicine*, 18(1). doi: 10.1186/s12916-020-01543-8.

Ziade, N. *et al.* (2020) 'Prevalence and pattern of comorbidities in chronic rheumatic and musculoskeletal diseases: the COMORD study', *Scientific Reports*, 10(1). doi: 10.1038/s41598-020-64732-8.

Ziegler, J. (1998) 'Cancer and Arthritis Share Underlying Processes', 90(11). Available at: https://academic.oup.com/jnci/article-abstract/90/11/802/916061.

Zlateva, G. *et al.* (2010) 'Burden of anemia in patients with osteoarthritis and rheumatoid arthritis in French secondary care', *BMC Geriatrics*, 10(1). doi: 10.1186/1471-2318-10-59.

李红英 *et al.* (2006) '性教育在西部普通高校体育与健康教育课程中开设现状的调查研究', 武汉体育学院学报, 40(6), pp. 99–101. doi: 10.3969/j.issn.1000-520X.2006.06.026.

11 Appendices

Appendix Table 1 Read codes for Joint pain

Read Code	Name
N094000	Arthralgia of unspecified site
N094900	Arthralgia of multiple joints
N094600	Arthralgia of the lower leg
N094F00	Arthralgia of wrist
N094M00	Arthralgia of knee
N094400	Arthralgia of the hand
N094D00	Arthralgia of elbow
N094300	Arthralgia of the forearm
N094J00	Arthralgia of DIP joint of finger
N094700	Arthralgia of the ankle and foot
N094B00	Arthralgia of sternoclavicular joint
N094500	Arthralgia of the pelvic region and thigh
N094200	Arthralgia of the upper arm
N094G00	Arthralgia of MCP joint
N094H00	Arthralgia of PIP joint of finger
N094z00	Arthralgia NOS
N094K00	Arthralgia of hip
N094P00	Arthralgia of ankle
N094800	Arthralgia of other specified site
N094C00	Arthralgia of acromioclavicular joint
N094V00	Arthralgia of IP joint of toe
N094T00	Arthralgia of 1st MTP joint
N094R00	Arthralgia of talonavicular joint
N094U00	Arthralgia of lesser MTP joint
N094S00	Arthralgia of other tarsal joint
N094N00	Arthralgia of tibio-fibular joint
N094Q00	Arthralgia of subtalar joint
N094E00	Arthralgia of distal radio-ulnar joint
N094W	Anterior knee pain
N245012	Finger pain
N245000	Hand Pain
N245011	Thumb Pain
N245100	Foot pain
N245111	Toe Pain
1M10	Knee pain
1M11	Foot pain
1M13	Ankle pain

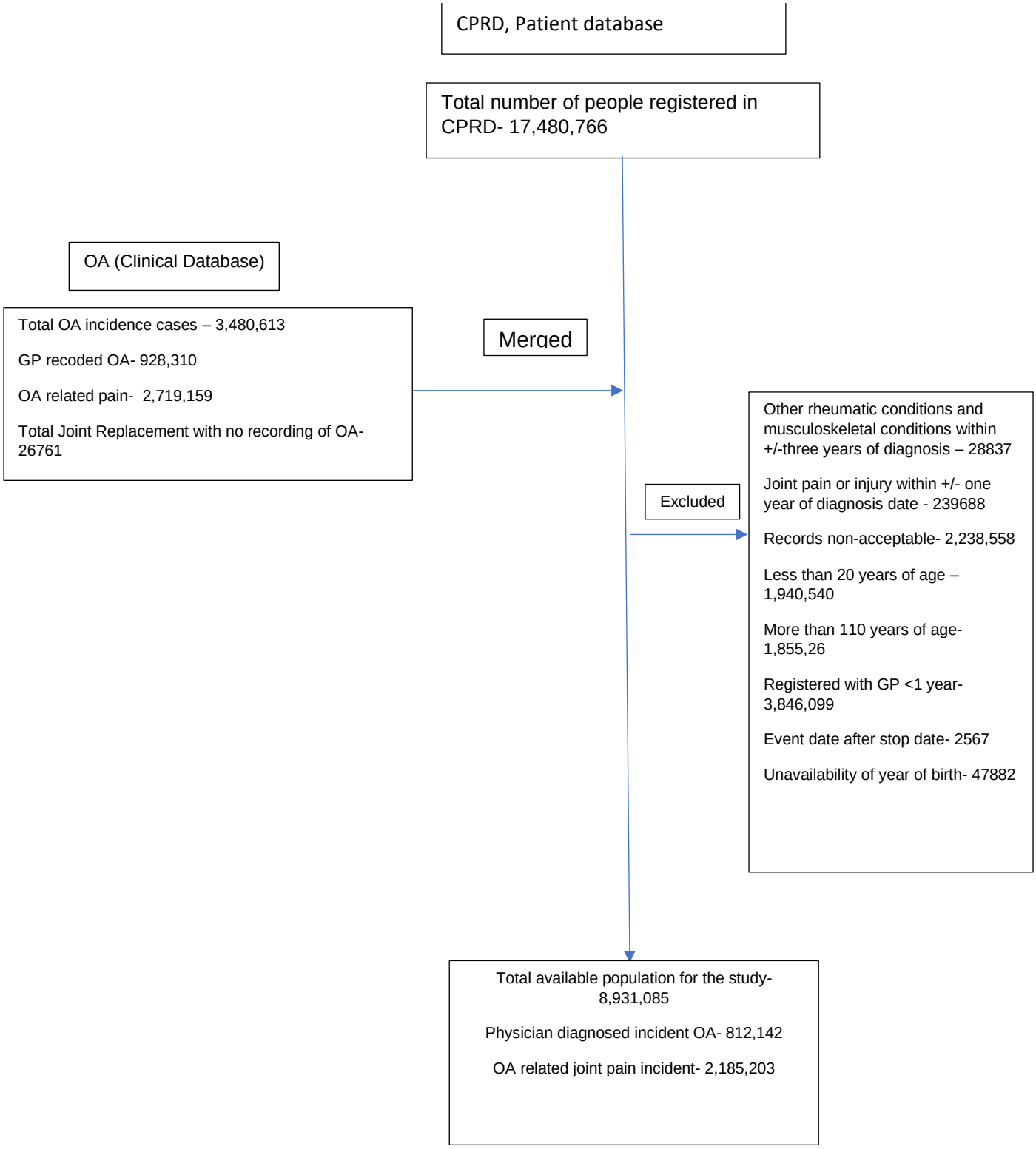
Appendix Table 2 Read codes for Osteoarthritis

Read Code	Name of the condition
N05zJ00	Osteoarthritis NOS, of hip
N053512	Hip osteoarthritis NOS
N05z511	Hip osteoarthritis NOS
N053500	Localised osteoarthritis, unspecified, pelvic region/thigh
N051500	Localised, primary osteoarthritis of the pelvic region/thigh
N052500	Localised, secondary osteoarthritis of pelvic region/thigh
N054500	Oligoarticular osteoarthritis, unspecified, of pelvis/thigh
N05z500	Osteoarthritis NOS, pelvic region/thigh
Nyu2E11	[X] Unilateral secondary coxarthrosis
Nyu2200	[X]Other dysplastic coxarthrosis
Nyu2300	[X]Other post-traumatic coxarthrosis
Nyu2100	[X]Other primary coxarthrosis
Nyu2E00	[X]Other secondary coxarthrosis
Nyu2400	[X]Other secondary coxarthrosis, bilateral
N051900	Primary coxarthrosis, bilateral
N05zL00	Osteoarthritis NOS, of knee
N05z611	Knee osteoarthritis NOS
N053600	Localised osteoarthritis, unspecified, of the lower leg
N05z600	Osteoarthritis NOS, of the lower leg
N051600	Localised, primary osteoarthritis of the lower leg
N052600	Localised, secondary osteoarthritis of the lower leg
N054600	Oligoarticular osteoarthritis, unspecified, of lower leg
N053611	Patellofemoral osteoarthritis
N05zM00	Osteoarthritis NOS, of tibio-fibular joint
Nyu2511	[X] Unilateral primary gonarthrosis
N051B00	Primary gonarthrosis, bilateral
Nyu2811	[X] Unilateral secondary gonarthrosis
Nyu2800	[X]Other secondary gonarthrosis
Nyu2700	[X]Other secondary gonarthrosis, bilateral
Nyu2500	[X]Other primary gonarthrosis
N052C00	Post-traumatic gonarthrosis, unilateral
N05zN00	Osteoarthritis NOS, of ankle
N05z700	Osteoarthritis NOS, of ankle and foot
N05zU00	Osteoarthritis NOS, of IP joint of toe
N05zT00	Osteoarthritis NOS, of lesser MTP joint
N05zS00	Osteoarthritis NOS, of 1st MTP joint
N05zR00	Osteoarthritis NOS, of other tarsal joint
N05zP00	Osteoarthritis NOS, of subtalar joint
N05z712	Foot Osteoarthritis NOS
N05zQ00	Osteoarthritis NOS, of talonavicular joint
N053700	Localised osteoarthritis, unspecified, of the ankle and foot
N051700	Localised, primary osteoarthritis of the ankle and foot
N051E00	Localised, primary osteoarthritis of toe
N052700	Localised, secondary osteoarthritis of the ankle and foot
N05z713	Toe osteoarthritis NOS
N05z711	Ankle osteoarthritis NOS
N054700	Oligoarticular osteoarthritis, unspecified, of ankle/foot
Nyu2900	[X]Other primary arthrosis of first carpometacarpal joint
N051C00	Primary arthrosis of first carpometacarpal joints, bilateral
Nyu2A00	[X]Other post-traumatic arthrosis/1st carpometacarpal joint
Nyu2B00	[X]Other 2ndry arthrosis/1st carpometacarpal joints, bilateral
N053400	Localised osteoarthritis, unspecified, of the hand

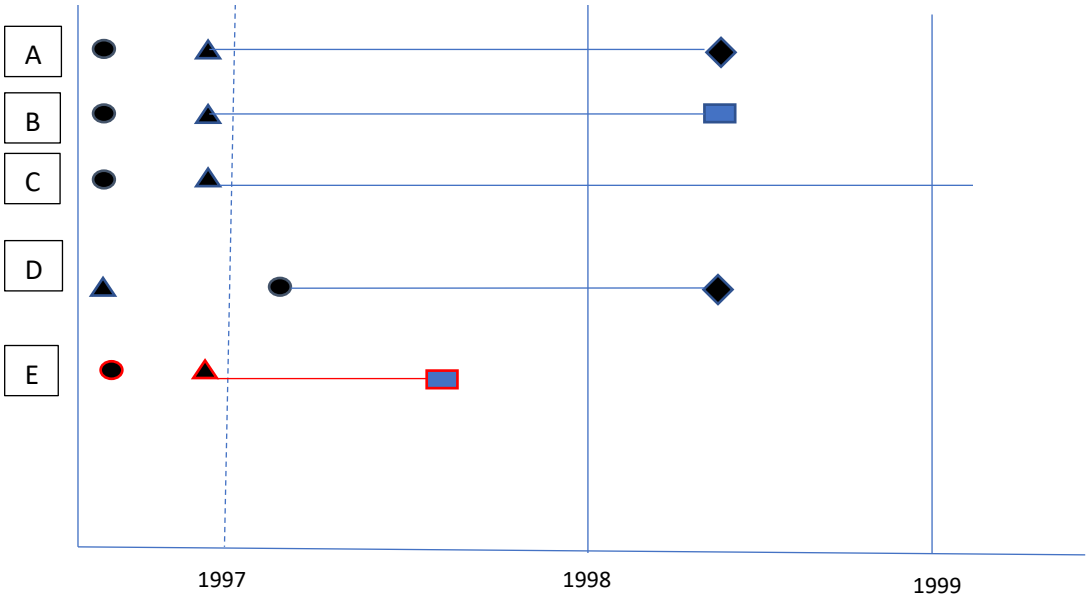
N051400	Localised, primary osteoarthritis of the hand
N05011	Heberden's node
N052400	Localised, secondary osteoarthritis of the hand
N05z412	Thumb osteoarthritis NOS
N050700	Heberden's node with arthropathy
N054400	Oligoarticular osteoarthritis, unspecified, of hand
N050112	Bouchard's node
N05zH00	Osteoarthritis NOS, of DIP joint of finger
N050300	Bouchard's node with arthropathy
N05zG00	Osteoarthritis NOS, of PIP joint of finger
N05z311	Wrist osteoarthritis NOS
N05z400	Osteoarthritis NOS, of the hand
N051D00	Localised, primary osteoarthritis of the wrist
N05z411	Finger osteoarthritis NOS
N05zE00	Osteoarthritis NOS, of wrist
N05zF00	Osteoarthritis NOS, of MCP joint
N050100	Generalized OA of hand
N06z311	Wrist arthritis NOS
N053100	Localised osteoarthritis, unspecified, of shoulder region
N051100	Localised, primary osteoarthritis of the shoulder region
N052200	Localised, secondary osteoarthritis of the upper arm
N052000	Localised, secondary osteoarthritis of unspecified site
N054100	Oligoarticular osteoarthritis, unspecified, of shoulder
N054200	Oligoarticular osteoarthritis, unspecified, of upper arm
N05z900	Osteoarthritis NOS, of shoulder
N05z100	Osteoarthritis NOS, of shoulder region
N052100	Localised, secondary osteoarthritis of the shoulder region
N05zC00	Osteoarthritis NOS, of elbow
N05zD00	Osteoarthritis NOS, of distal radio-ulnar joint
N06z211	Elbow arthritis NOS
N051300	Localised, primary osteoarthritis of the forearm
N051F00	Localised, primary osteoarthritis of elbow
N051800	Localised, primary osteoarthritis of other specified site
N051.00	Localised, primary osteoarthritis
N051z00	Localised, primary osteoarthritis NOS
N051000	Localised, primary osteoarthritis of unspecified site
N052.00	Localised, secondary osteoarthritis
N052z00	Localised, secondary osteoarthritis NOS
N052800	Localised, secondary osteoarthritis of other specified site
N050000	Osteoarthritis and allied disorders
N054.00	Oligoarticular osteoarthritis, unspecified
N054900	Oligoarticular osteoarthritis, unspecified, multiple sites
Nyu2.00	[X]Arthrosis
Nyu2000	[X]Other polyarthrosis
N054000	Oligoarticular osteoarthritis, unspec, of unspecified sites
N05z000	Osteoarthritis NOS, of unspecified site
N05..00	Osteoarthritis and allied disorders
N054800	Oligoarticular osteoarthritis, unspecified, other spec sites
N05z.00	Osteoarthritis NOS
N053z00	Localised osteoarthritis, unspecified, NOS
N053800	Localised osteoarthritis, unspecified, of other spec site
N05zz00	Osteoarthritis NOS
N053000	Localised osteoarthritis, unspecified, of unspecified site
N05..11	Osteoarthritis
N05z800	Osteoarthritis NOS, other specified site

N054z00	Osteoarthritis of more than one site, unspecified, NOS
N06z.11	Arthritis
N050500	Secondary multiple arthrosis
N050400	Primary generalized osteoarthrosis
N050Z00	Generalized OA NOS
N050200	Generalised OA Multiple sites
N050.00	Generalised OA

Appendix Figure 1 Flow chart of the eligible study participants



Appendix Figure 2 Selection of study population for incidence and prevalence



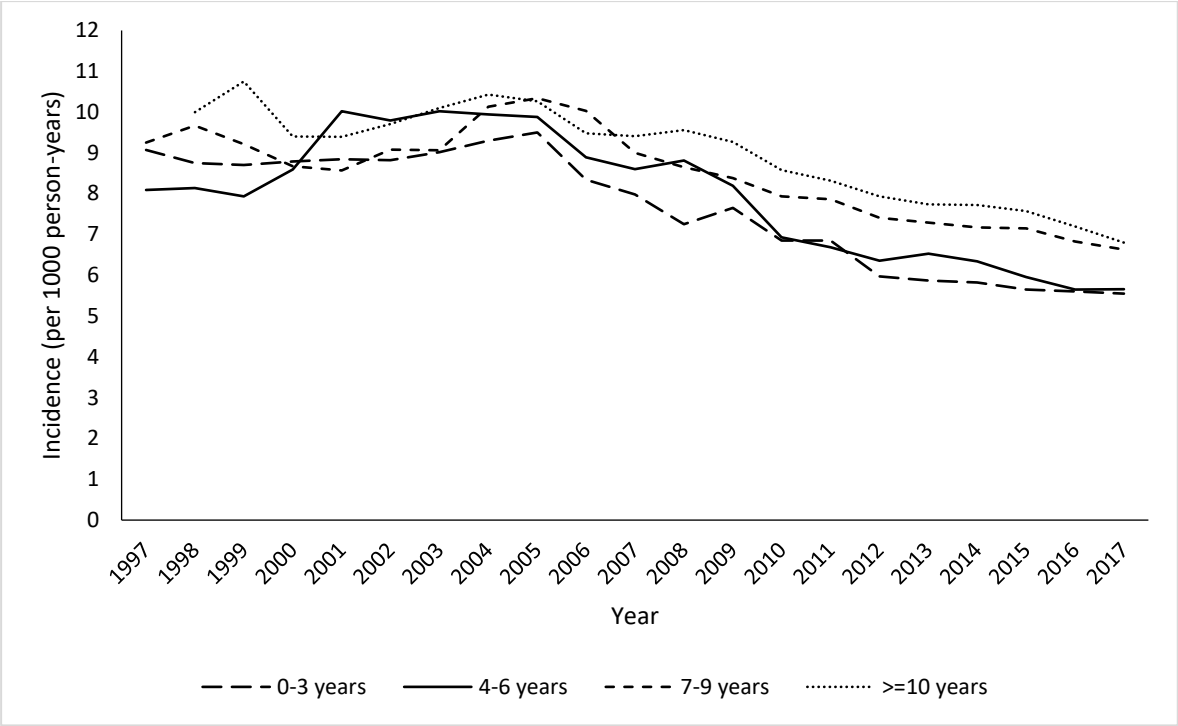
Circle: Up to standard; Triangle: Current Registration; Diamond: Diagnosis date; Square: Death/Transfer out/LCD

Appendix Table 3 Selection criteria based on the dates in the record

Scenario	UTS checking	CRD checking	Inclusion in Numerator	Inclusion in Denominator	Incidence
A	Yes	Yes	Yes	Yes	Yes
B	Yes	Yes	No	Yes	Yes
C	Yes	Yes	No	Yes	Yes
D	Fail	No	No	No	No
E	Yes	Yes	No	No	No

Appendix Figure 3 Crude incidence trends across length of data contribution for incidence (A) and prevalence (B)

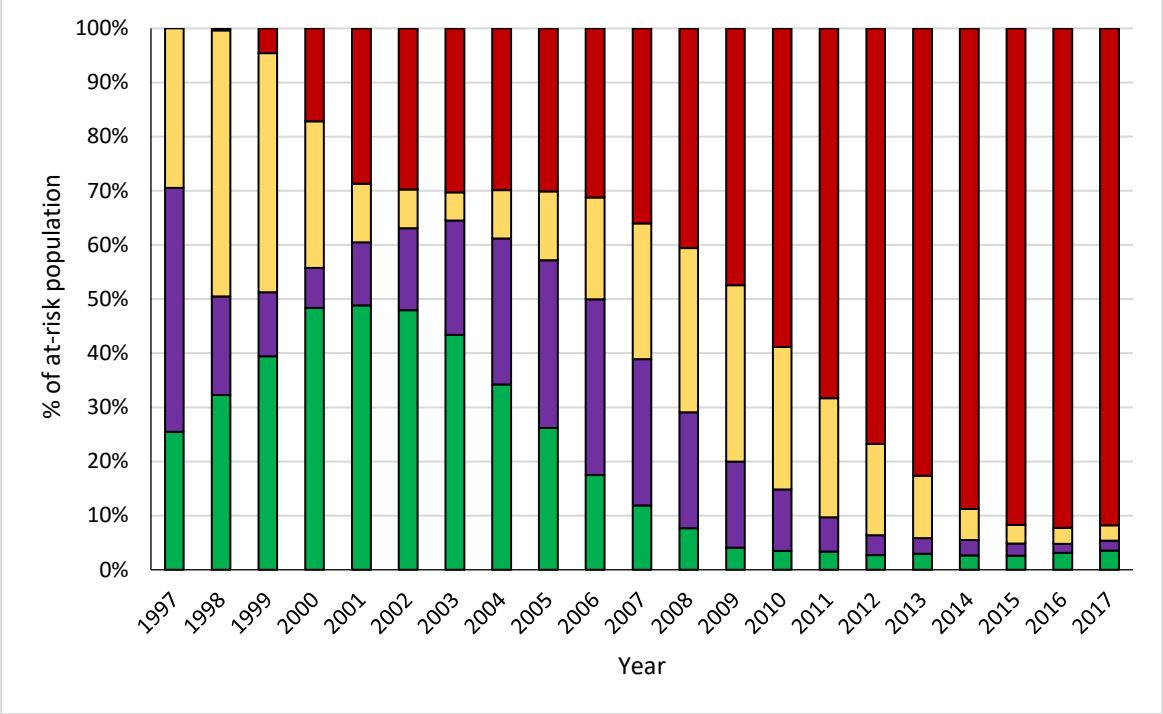
A. Incidence



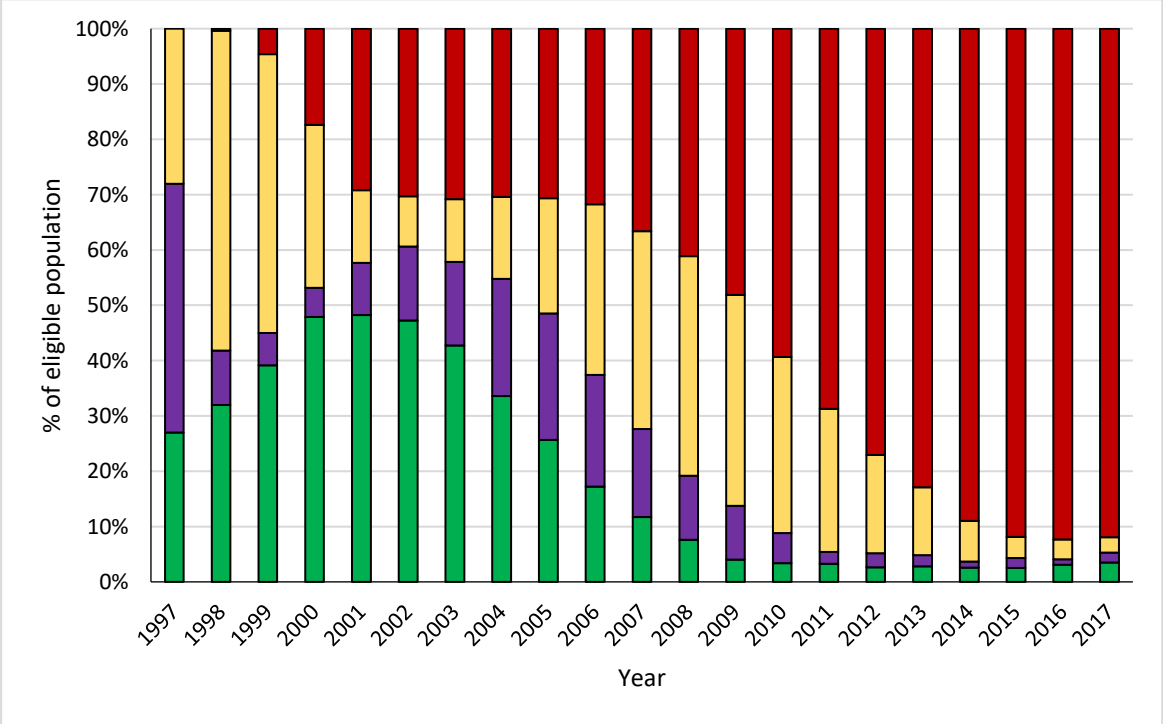
B. Prevalence

Appendix Figure 4 At-risk and eligible population at the study year for incidence (A) and prevalence (B) calculation across length of data contribution

A. Incidence

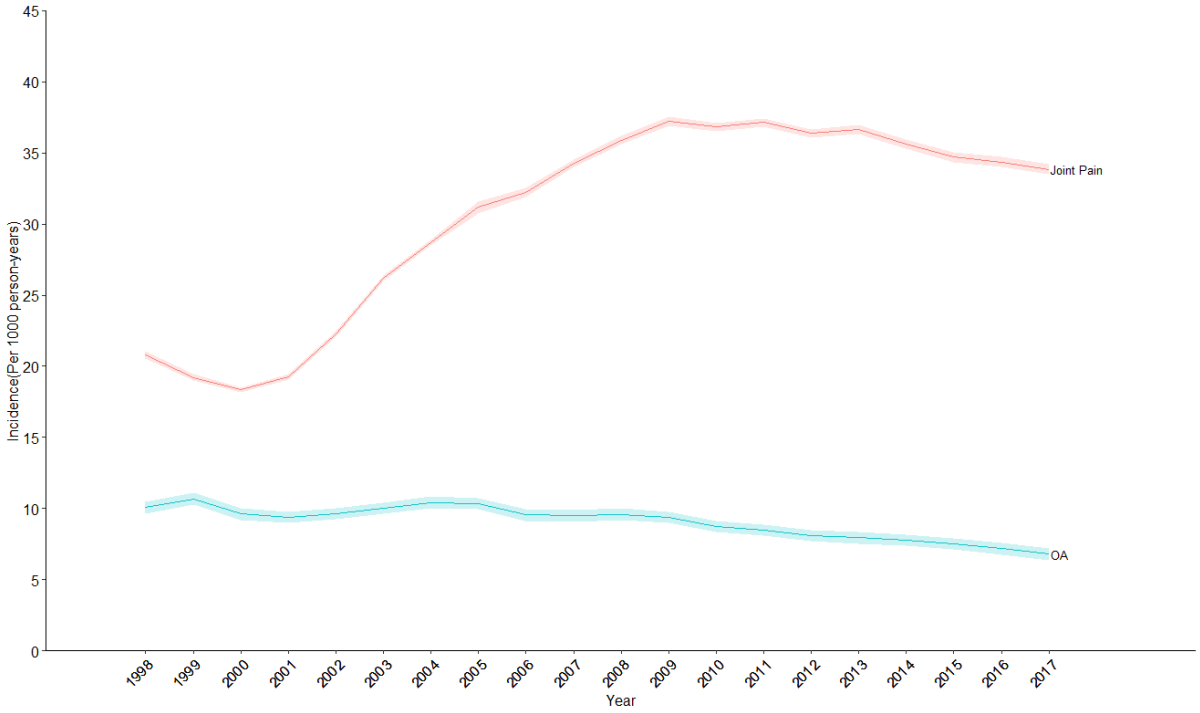


B. Prevalence

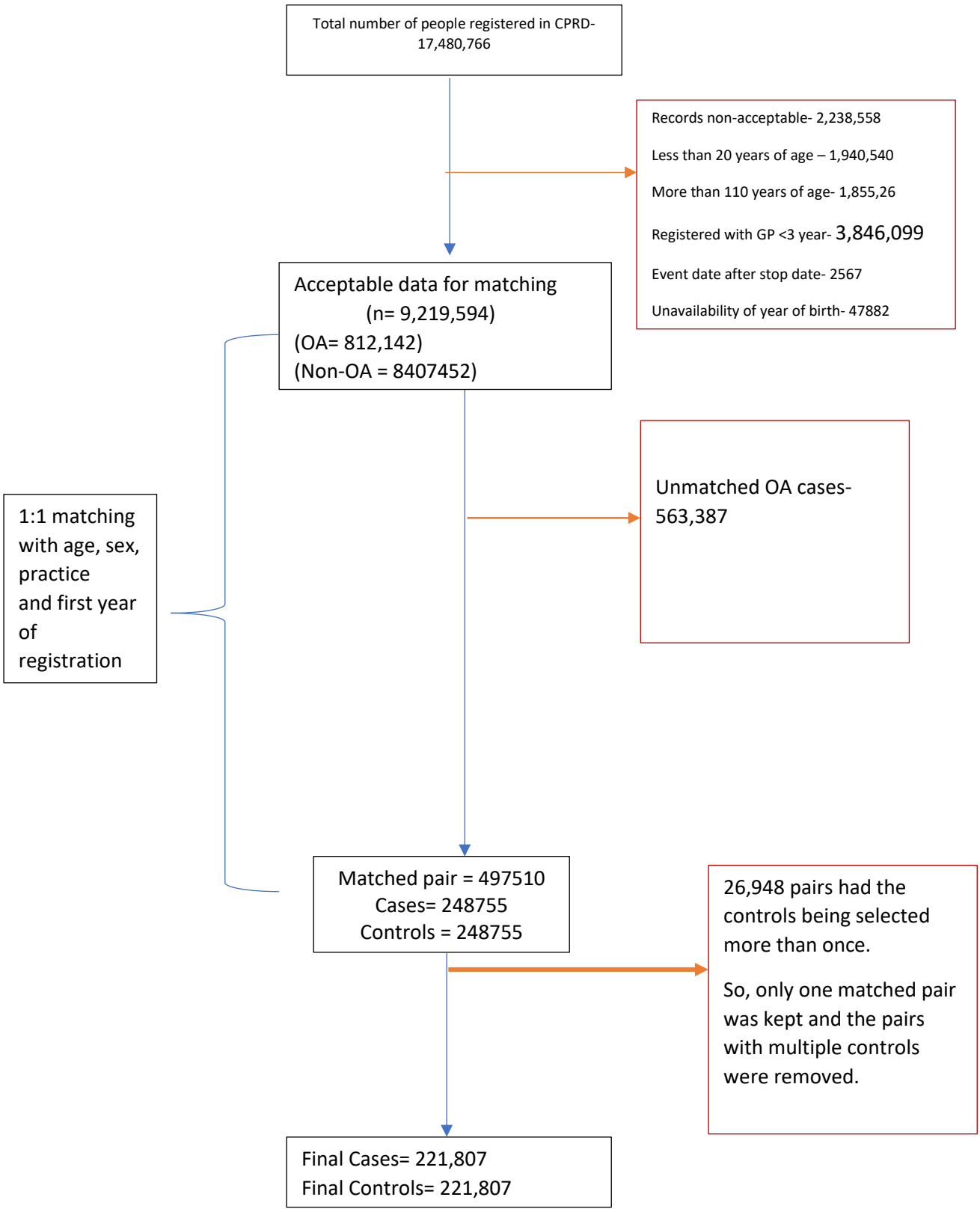


Legend: Green- 0-3 years; Purple- 4-6 years; Yellow- 7-9 years; Red- ≥10 years

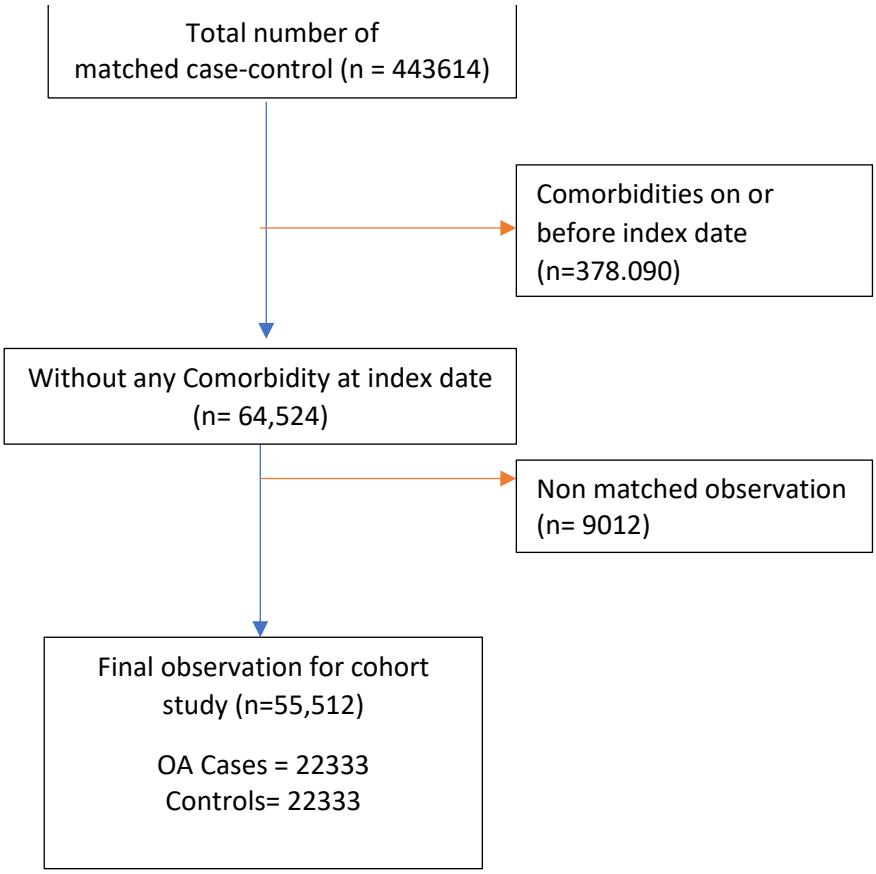
Appendix Figure 5 A comparison of trends of incidence of GP diagnosed OA and joint pain in the UK



Appendix Figure 6 Selection of matched case-controls for retrospective study



Appendix Figure 7 Selection of matched OA cases (exposed) with controls (unexposed) for cohort studies (No comorbidities on or before the index date)



Appendix Table 4 Code list of key comorbidities

Group	Diseases	Initials of major read codes
Musculoskeletal	Rheumatoid Arthritis	N040.0; N040..; N043… ;N04y…
	Back pain	N11….; N12…; N14….; N3…; Nyu…; S10.…; S49…; S57…;
	Crystal arthropathy	C34…; N02z…; N02y…; N022..; N021…
	Osteoporosis	N33…
	Fibromyalgia	N248.00; N239.00; F286…
Cardio-vascular	Coronary Heart Disease	G3…00; G30….; G30z….; G34…; G35…; G54….; G57….; Gyu…
	Arterial/Venous	G70…; G71…; G72….; G74…
	Heart failure	G580…; G81….; G232.00; G234.00
	Hypertension	G20…; G21…; G24…; G25…; G26…
Respiratory	Asthma	663…; 66Y…;
	COPD	H31…; H32…; H33…;
Genito-urinary	Chronic Kidney Disease	1Z12.00; 1Z13.00; 1Z14.00; 1Z15.00; 1Z16.00; K01…; K02…
	Renal stone	4G4…; 7B07…; K120….
Neurological	Stroke	9Om0.00; 9Om…; 8HBJ.oo; G60…; G61…; F22…; G64…; G63…
	Dementia	E00…; Eu0…; F11….
	Parkinson Disease	F12…
	Migraine	F26….
Psychiatric	Depression	Eu1…; Eu3…
	Psychosis	Eu2Z.11; Eu0z.11
	Schizophrenia	Eu2…
Metabolic/Endocrine	High Cholesterol	C324.00; C322.00; C328.00
	Diabetes Mellitus	9OL...00; 2BB…; 2G5…; 66A…; C10…; F42…
	Hyperthyroid	C02…
	Hypothyroidism	C03…; C04…
Digestive	Gastritis	J11…; J12…; J15…14C1…
	Gastrointestinal bleed	J110…; J111….; J120…; J121…; J13….; J14…. J681…
	Gall bladder stone	J65…; 4G2…
	Liver Disease	J61…; J63…; A70…;
Other	Hearing	F59….
	Sleep Disorder	Fy0…
	Anaemia	D00…

The codes are initials of the read codes representing the comorbidities.

Appendix Table 5 Comorbidities in the past years prior to the diagnosis of OA at any joint (Expanded version)

	1 year				5 years				10 years				15 years				20 years			
	Non-OA		OA		Non-OA		OA		Non-OA		OA		Non-OA		OA		Non-OA		OA	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Musculoskeletal																				
Ankylosing Spondylitis	132	0.06	215	0.09	674	0.30	1092	0.49	1358	0.61	2163	0.98	1889	0.87	2891	1.33	2158	1.03	3258	1.55
Back pain	4452	2.02	7632	3.44	21452	9.73	30343	13.72	40443	18.44	55334	25.20	54194	25.11	73749	34.11	61835	29.61	84092	40.12
Crystal arthropathy	493	0.22	749	0.34	2162	0.98	3625	1.64	3564	1.62	5995	2.73	4382	2.03	7325	3.38	4829	2.31	8013	3.82
Osteoporosis	632	0.28	1166	0.52	2622	1.19	3680	1.66	4028	1.84	5267	2.39	4664	2.61	5961	2.75	4896	2.34	6260	2.98
Polymyalgia	170	0.08	323	0.14	659	0.29	1285	0.58	980	0.45	1885	0.86	1161	0.54	2129	0.98	1243	0.59	2226	1.06
Rheumatoid Arthritis	91	0.04	367	0.16	387	0.17	1015	0.46	710	0.32	1586	0.72	865	0.40	1904	0.88	972	0.46	1956	0.93
Sjogren's syndrome	18	0.01	48	0.02	71	0.03	171	0.08	137	0.06	265	0.12	173	0.08	314	0.14	202	0.09	340	0.16
Systemic lupus erythematous	5	0.00	11	0.00	26	0.01	46	0.02	49	0.02	80	0.03	73	0.03	109	0.05	81	0.04	122	0.05
Fibromyalgia	115	0.05	404	0.18	490	0.22	1192	0.54	827	0.37	1829	0.83	1027	0.47	2129	0.98	1073	0.51	2162	1.03
Fatigue	218	0.09	360	0.16	915	0.41	1363	0.62	1445	0.66	2099	0.95	1693	0.78	2386	1.10	1739	0.83	2453	1.17
Respiratory																				
Asthma	691	0.31	1081	0.48	3636	1.64	5338	2.41	7458	3.40	10628	4.84	10482	4.86	14770	6.83	12320	5.90	17029	8.12
COPD	602	0.27	927	0.42	2984	1.35	4209	1.90	5886	2.68	8126	3.70	8043	3.72	11088	5.12	9296	4.45	12642	6.05
Genito-Urinary																				
Chronic kidney disease	1566	0.71	2002	0.90	5549	2.51	6789	3.07	7369	3.36	8768	3.99	7513	3.48	8960	4.14	7527	3.60	8965	4.27
Prostate^	639	0.29	989	0.45	2899	1.31	4038	1.83	4884	2.22	6543	2.98	5972	2.76	7901	3.65	6365	3.05	8436	4.02
Renal stone	114	0.05	158	0.07	542	0.24	698	0.31	989	0.45	1261	0.57	1325	0.61	1635	0.76	1567	0.75	1923	0.91
Neuro/Psychiatric																				
Stroke	1354	0.61	1773	0.80	6025	2.73	7141	3.22	10314	4.70	11902	5.42	13042	6.04	14826	6.85	14200	6.80	16158	7.70
Dementia	235	0.11	355	0.16	741	0.33	908	0.41	929	0.42	1036	0.47	978	0.45	1061	0.49	990	0.47	1068	0.51
Epilepsy	72	0.03	144	0.06	414	0.19	511	0.23	702	0.30	897	0.41	966	0.45	1196	0.55	1125	0.54	1376	0.65
Multiple sclerosis	22	0.01	35	0.01	127	0.57	124	0.56	236	0.11	230	0.10	329	0.15	293	0.13	433	0.20	348	0.17
Parkinson's Disease	80	0.03	161	0.07	318	0.14	481	0.22	450	0.20	629	0.29	489	0.23	680	0.31	502	0.24	696	0.33
Migraine	500	0.23	745	0.33	2487	1.12	3561	1.61	5093	2.32	7065	3.21	7203	3.33	9802	4.53	8489	4.06	11359	5.41
Depression	1799	0.82	2978	1.34	9051	4.10	13588	6.14	17610	8.03	25398	11.57	24076	11.15	34044	15.74	27362	13.10	38417	18.32
Psychosis	18	0.001	41	0.01	154	0.07	161	0.07	291	0.13	268	0.12	377	0.17	354	0.16	419	0.20	398	0.19
Schizophrenia	51	0.02	84	0.04	325	0.14	389	0.17	657	0.29	692	0.31	888	0.41	928	0.43	1034	0.49	1073	0.51
Cancer	874	0.39	902	0.41	3697	1.67	4287	1.94	5951	2.71	6795	3.09	7294	3.38	8248	3.81	7984	3.80	8972	4.28
Circulatory																				
Coronary heart disease	967	0.44	1257	0.56	5059	2.29	6390	2.89	9472	4.32	12171	5.54	12496	5.79	16162	7.47	14262	6.83	18302	8.73
Arterial/Venous	116	0.05	176	0.08	513	0.23	731	0.33	825	0.37	1123	0.51	989	0.46	1337	0.62	1062	0.51	1429	0.68
Heart failure	289	0.13	444	0.20	1045	0.47	1795	0.81	1568	0.71	2658	1.21	1777	0.82	3011	1.39	1847	0.88	3113	1.48
Hypertension	3906	1.77	4805	2.16	18204	8.25	20969	9.48	32449	14.80	37418	17.04	41348	19.16	48042	22.22	46012	22.03	53659	25.60
Peripheral vascular disease	413	0.18	753	0.34	1767	0.80	2734	1.23	2939	1.34	4411	2.00	3614	1.67	5191	2.40	3906	1.87	5539	2.64
Metabolic																				
High Cholesterol	2239	1.02	3053	1.37	9875	4.47	12467	5.63	16604	7.57	20458	9.32	20291	9.40	24714	11.43	21865	10.47	26558	12.67
Diabetes Mellitus	1397	0.63	1948	0.88	6188	2.80	7954	3.59	9945	4.53	12677	5.77	11856	5.49	15129	6.99	12656	6.06	16147	7.70
Hyperthyroid	137	0.06	142	0.06	665	0.30	712	0.32	1205	0.55	1294	0.59	1579	0.73	1732	0.80	1843	0.88	2047	0.97
Hypothyroidism	895	0.40	1203	0.54	4075	1.84	5067	2.29	7050	3.21	8732	3.97	8922	4.13	11096	5.13	9793	4.69	12276	5.85
Digestive																				
Gastritis	610	0.28	997	0.45	2771	1.25	4069	1.84	4915	2.24	7070	3.22	6542	3.03	9198	4.25	7551	3.61	10527	5.02
Gastrointestinal bleed	155	0.07	270	0.12	672	0.30	1032	0.47	1133	0.52	1675	0.76	1431	0.66	2072	0.96	1570	0.75	2253	1.07
Gall bladder stone	533	0.24	660	0.30	2490	1.13	3438	1.55	4296	1.95	6077	2.76	5602	2.59	7971	3.68	6461	3.09	9189	4.38
Inflammatory Bowel Disease	578	0.26	805	0.36	2548	1.15	3695	1.67	4514	2.05	6379	2.90	5824	2.69	7983	3.69	6409	3.06	8704	4.15
Liver Disease	73	0.03	135	0.06	329	0.15	508	0.23	506	0.23	796	0.36	637	0.29	950	0.44	689	0.32	1029	0.49
Irritable bowel syndrome	986	0.44	1421	0.63	3134	1.42	4787	2.16	6266	2.85	9261	4.21	8653	4.01	12589	5.82	10015	4.79	14335	6.83
Others																				
Hearing	1666	0.75	2357	1.06	7193	3.26	9172	4.14	11807	5.38	14748	6.71	14329	6.64	17855	8.26	15587	7.46	19315	9.21
Vision problem	130	0.06	136	0.06	510	0.23	625	0.28	860	0.39	1015	0.46	1059	0.49	1209	0.56	1136	0.54	1313	0.62
Psoriasis	277	0.12	439	0.19	1286	0.58	1751	0.79	2379	1.08	3127	1.42	3174	1.47	4086	1.89	3655	1.75	4602	2.19
Scleroderma	2		12		17	0.01	29	0.01	37	0.02	41	0.02	46	0.02	51	0.02	54	0.02	55	0.02
Sleep Disorder	481	0.22	724	0.32	2061	0.93	2877	1.30	3169	1.44	4340	1.97	3677	1.70	4978	2.30	3820	1.82	5148	2.45
Tuberculosis	16	0.01	32	0.01	112	0.05	139	0.06	215	0.09	269	0.12	310	0.14	360	0.17	342	0.16	417	0.19
Anaemia	588	0.26	920	0.41	2389	1.08	3385	1.53	4010	1.83	5268	2.40	4927	2.28	6269	2.90	5406	2.59	6732	3.21
Comorbidities (count)																				
No comorbidity	195859	88.10	184311	82.91	131897	59.33	109920	49.44	95710	43.05	73856	33.22	81303	36.57	61335	27.59	77845	35.01	59752	26.88
Single comorbidity	22891	10.29	31971	14.38	57054	25.66	64354	28.95	60358	27.15	59574	26.80	56157	25.26	51005	22.94	51546	23.18	44541	20.03
Any two comorbidities	3058	1.37	5042	2.26	22617	10.17	30264	13.61	35762	16.08	42647	19.18	39244	17.65	43511	19.57	38897	17.49	41327	18.59
Any three comorbidities	415	0.19	787	0.35	7618	3.42	11901	5.35	17810	8.01	24647	11.09	23302	10.48	30336	13.64	25282	11.37	31429	14.14
Four or more	67	0.03	179	0.08	3104	1.39	5851	2.63	12650	5.69	21568	9.70	22284	10.02	36103	16.24	28720	12.92	45241	20.35

COPD- Chronic Obstructive Pulmonary Disease; SLE: Systemic lupus erythematous; ^only for men

Appendix Table 6 Association between any OA and comorbidities in the past years prior to the index date (Expanded version)

	20 years		15 years		10 years		5 years		1 year	
	Unadjusted OR	Adjusted OR[#]	Unadjusted OR	Adjusted OR[#]	Unadjusted OR	Adjusted OR[#]	Unadjusted OR	Adjusted OR[#]	Unadjusted OR	Adjusted OR[#]
>=2 comorbidities	1.86 (1.83-1.88)*	1.71(1.69-1.74)*	1.80(1.77-1.82)*	1.66(1.63-1.68)*	1.71(1.68-1.73)*	1.58(1.56-1.60)*	1.63(1.60-1.65)*	1.53(1.49-1.55)*	1.64 (1.56-1.71)*	1.52 (1.45-1.59)*
Musculoskeletal										
Ankylosing Spondylitis	1.53 (1.45-1.62)*	1.53 (1.44-1.62)*	1.55 (1.46-1.64)	1.56 (1.46-1.65)*	1.61 (1.50-1.72)*	1.63 (1.52-1.75)*	1.61 (1.46-1.77)*	1.63 (1.47-1.79)*	1.55 (1.24-1.92)*	1.49 (1.19-1.86)*
Back pain	1.70 (1.67-1.72)*	1.67 (1.64-1.69)*	1.61 (1.59-1.64)*	1.59 (1.56-1.61)*	1.52 (1.50-1.55)*	1.51 (1.48-1.53)*	1.47 (1.44-1.50)*	1.45 (1.43-1.48)*	1.61 (1.55-1.68)*	1.60 (1.54-1.69)*
Crystal arthropathy	1.69 (1.64-1.76)*	1.52 (1.46-1.57)*	1.70 (1.63-1.77)*	1.52 (1.45-1.58)*	1.69 (1.63-1.77)*	1.52 (1.45-1.59)*	1.67 (1.58-1.76)*	1.49 (1.41-1.58)*	1.39 (1.24-1.56)*	1.26 (1.11-1.42)*
Osteoporosis	1.27 (1.22-1.32)*	1.41 (1.35-1.47)*	1.27 (1.22-1.32)*	1.41 (1.35-1.46)*	1.29 (1.23-1.34)*	1.42 (1.36-1.49)*	1.36 (1.29-1.43)*	1.49 (1.42-1.58)*	1.59 (1.44-1.76)*	1.74 (1.57-1.93)*
Polymyalgia	1.80 (1.68-1.93)*	1.74 (1.62-1.87)*	1.84 (1.71-1.98)*	1.78 (1.65-1.92)*	1.92 (1.78-2.08)*	1.86 (1.72-2.01)*	1.93 (1.76-2.12)*	1.86 (1.69-2.05)*	1.80 (1.49-2.17)*	1.71 (1.41-2.08)*
Rheumatoid Arthritis	1.97 (1.83-2.13)*	1.95 (1.80-2.11)*	2.16 (1.99-2.34)*	2.14 (1.97-2.32)*	2.18 (1.99-2.39)*	2.17 (1.98-2.38)*	2.51 (2.23-2.83)*	2.50 (2.21-2.82)*	3.60 (2.85-4.54)*	3.69 (2.90-4.68)*
Sjogren's syndrome	1.64 (1.38-1.96)*	1.67 (1.39-2.00)*	1.77 (1.47-2.13)*	1.82 (1.50-2.20)*	1.86 (1.52-2.29)*	1.94 (1.56-2.40)*	2.31 (1.75-3.05)*	2.47 (1.85-3.30)*	2.27 (1.30-3.96)	2.60 (1.44-4.69)
Systemic lupus erythematous	1.49 (1.12-1.98)	1.54 (1.15-2.07)	1.48 (1.09-1.99)	1.54 (1.10-2.05)	1.63 (1.14-2.33)	1.59 (1.09-2.29)	1.77 (1.09-2.86)	1.72 (1.05-2.82)	2.19 (0.76-6.33)	2.31 (0.76-7.05)
Fibromyalgia	1.95 (1.81-2.10)*	1.89 (1.75-2.04)*	2.01 (1.86-2.16)*	1.95 (1.80-2.11)*	2.12 (1.95-2.31)*	2.07 (1.89-2.25)*	2.27 (2.03-2.53)*	2.19 (1.96-2.45)*	2.71 (2.18-3.36)*	2.77 (2.21-3.46)*
Fatigue	1.42 (1.33-1.51)*	1.42 (1.32-1.51)*	1.42 (1.33-1.51)*	1.42 (1.33-1.52)*	1.46 (1.36-1.56)*	1.46 (1.36-1.57)*	1.49 (1.36-1.62)*	1.48 (1.36-1.62)*	1.59 (1.34-1.88)*	1.56 (1.30-1.86)*
Respiratory										
Asthma	1.41 (1.38-1.45)*	1.33 (1.30-1.37)*	1.44 (1.40-1.48)*	1.35 (1.31-1.39)*	1.44 (1.40-1.49)*	1.35 (1.31-1.39)*	1.46 (1.40-1.53)*	1.37 (1.31-1.43)*	1.47 (1.34-1.63)*	1.36 (1.23-1.51)*
COPD	1.40 (1.37-1.45)*	1.35 (1.31-1.39)*	1.42 (1.38-1.46)*	1.36 (1.32-1.41)*	1.40 (1.36-1.46)*	1.36 (1.31-1.41)*	1.40 (1.34-1.47)*	1.37 (1.30-1.43)*	1.46 (1.31-1.62)*	1.42 (1.28-1.58)*
Genito-Urinary										
Chronic kidney disease	1.25 (1.20-1.29)*	1.12 (1.08-1.16)*	1.24 (1.20-1.29)*	1.12 (1.08-1.16)*	1.24 (1.20-1.29)*	1.12(1.08-1.16)*	1.27 (1.22-1.32)*	1.15 (1.10-1.19)*	1.27 (1.18-1.36)*	1.16 (1.08-1.24)*
Prostate	1.38 (1.32-1.43)*	1.38 (1.33-1.43)*	1.37 (1.32-1.42)*	1.37 (1.32-1.42)*	1.37 (1.31-1.42)*	1.37 (1.32-1.43)*	1.39 (1.32-1.46)*	1.37 (1.32-1.46)*	1.39 (1.25-1.53)*	1.37 (1.24-1.53)*
Renal stone	1.22 (1.14-1.31)	1.16 (1.09-1.25)*	1.23 (1.15-1.33)*	1.16 (1.08-1.26)*	1.28 (1.17-1.39)*	1.21 (1.11-1.32)*	1.28 (1.14-1.43)*	1.21 (1.08-1.36)*	1.34 (1.05-1.71)	1.31 (1.02-1.68)
Neuro/Psychiatric										
Stroke	1.17 (1.14-1.20)*	1.15 (1.11-1.19)*	1.17 (1.14-1.20)*	1.15 (1.11-1.19)*	1.18 (1.15-1.21)*	1.15 (1.12-1.19)*	1.19 (1.15-1.24)*	1.17 (1.13-1.22)*	1.26 (1.18-1.37)*	1.24 (1.15-1.34)*
Dementia	1.07 (0.97-1.17)	1.09 (0.99-1.19)	1.07 (0.98-1.17)	1.09 (1.00-1.20)	1.10 (1.01-1.21)	1.13 (1.03-1.24)	1.21 (1.09-1.33)*	1.23 (1.11-1.36)*	1.45 (1.23-1.72)*	1.44 (1.21-1.71)*
Epilepsy	1.20 (1.11-1.30)*	1.18 (1.08-1.29)*	1.22 (1.12-1.33)*	1.20 (1.10-1.31)*	1.26 (1.14-1.39)*	1.24 (1.11-1.37)*	1.20 (1.05-1.37)	1.17 (1.03-1.35)	1.85 (1.39-2.47)*	1.89 (1.40-2.54)*
Multiple sclerosis	0.79 (0.68-0.91)*	0.80 (0.69-0.93)*	0.88 (0.75-1.03)	0.89 (0.75-1.04)	0.97 (0.81-1.16)	0.95 (0.78-1.14)	0.98 (0.76-1.25)	0.95 (0.72-1.20)	1.59 (0.93-2.71)	1.55 (0.89-2.67)
Parkinson's Disease	1.36 (1.21-1.53)*	1.39 (1.23-1.57)*	1.36 (1.21-1.53)*	1.39 (1.24-1.57)*	1.36 (1.21-1.54)*	1.39 (1.22-1.57)*	1.46 (1.27-1.68)*	1.47 (1.27-1.70)*	1.79 (1.36-2.35)*	1.75 (1.33-2.31)*
Migraine	1.36 (1.32-1.39)*	1.37 (1.33-1.41)*	1.38 (1.33-1.42)*	1.39 (1.34-1.43)*	1.40 (1.35-1.45)*	1.42 (1.36-1.47)*	1.42 (1.34-1.49)*	1.44 (1.37-1.53)*	1.36 (1.21-1.53)*	1.40 (1.25-1.59)*
Depression	1.53 (1.50-1.56)*	1.49 (1.46-1.52)*	1.52 (1.49-1.55)*	1.49 (1.46-1.52)*	1.52 (1.49-1.55)*	1.49 (1.46-1.52)*	1.52 (1.48-1.56)*	1.49 (1.45-1.54)*	1.54 (1.45-1.64)*	1.51 (1.42-1.61)*
Psychosis	0.94 (0.82-1.08)	0.86 (0.75-1.00)	0.93 (0.81-1.08)	0.86 (0.74-0.99)	0.90 (0.76-1.07)	0.83 (0.69-0.98)	1.02 (0.82-1.27)	0.95 (0.75-1.19)	2.05 (1.17-3.61)*	1.89 (1.06-3.39)*
Schizophrenia	1.03 (0.95-1.12)	0.95 (0.87-1.04)	1.03 (0.95-1.14)	0.95 (0.86-1.05)	1.05 (0.94-1.17)	0.97 (0.87-1.08)	1.17 (1.01-1.36)	1.08 (0.92-1.26)	1.59 (1.12-2.27)	1.36 (0.95-1.96)
Cancer	1.13 (1.09-1.17)*	1.12 (1.09-1.16)*	1.14 (1.09-1.17)*	1.12 (1.09-1.16)*	1.15 (1.10-1.18)*	1.12 (1.08-1.17)*	1.15 (1.10-1.20)*	1.12 (1.08-1.18)*	0.98 (0.89-1.08)	0.96 (0.87-1.05)
Circulatory										
Coronary heart disease	1.33 (1.30-1.36)*	1.24 (1.21-1.27)*	1.34 (1.31-1.37)*	1.24 (1.20-1.27)*	1.31 (1.28-1.35)*	1.22 (1.18-1.25)*	1.26 (1.21-1.32)*	1.17 (1.12-1.21)*	1.22 (1.12-1.33)*	1.12 (1.03-1.23)
Arterial/Venous	1.34 (1.23-1.45)*	1.29 (1.19-1.41)*	1.34 (1.24-1.46)*	1.30 (1.19-1.42)*	1.34 (1.23-1.47)*	1.30 (1.19-1.43)*	1.39 (1.24-1.57)*	1.35 (1.20-1.52)*	1.39 (1.09-1.77)*	1.41 (1.10-1.81)
Heart failure	1.72 (1.62-1.82)*	1.52 (1.43-1.62)*	1.73 (1.62-1.83)*	1.52 (1.44-1.62)*	1.72 (1.61-1.84)*	1.52 (1.43-1.63)*	1.72 (1.59-1.86)*	1.53 (1.41-1.65)*	1.48 (1.27-1.72)*	1.30 (1.11-1.52)*
Hypertension	1.24 (1.22-1.26)*	1.08 (1.06-1.10)*	1.22 (1.20-1.24)*	1.08 (1.05-1.09)*	1.18 (1.16-1.20)*	1.06 (1.04-1.07)*	1.15 (1.12-1.17)*	1.04 (1.02-1.06)*	1.13 (1.08-1.18)*	1.03 (0.98-1.08)
Peripheral vascular disease	1.41 (1.35-1.47)*	1.45 (1.39-1.51)*	1.42 (1.36-1.49)*	1.45 (1.39-1.53)*	1.48 (1.41-1.55)*	1.51 (1.44-1.59)*	1.49 (1.41-1.59)*	1.54 (1.45-1.64)*	1.58 (1.40-1.79)*	1.62 (1.43-1.84)*
Metabolic/Endocrine										
High Cholesterol	1.27 (1.24-1.29)*	1.18 (1.16-1.20)*	1.26 (1.24-1.29)*	1.18 (1.15-1.20)*	1.27 (1.24-1.29)*	1.18 (1.15-1.21)*	1.27 (1.23-1.31)*	1.20 (1.16-1.23)*	1.27 (1.20-1.35)*	1.20 (1.13-1.28)*
Diabetes Mellitus	1.31 (1.27-1.34)*	1.06 (1.03-1.09)*	1.30 (1.27-1.34)*	1.06 (1.03-1.08)*	1.29 (1.26-1.33)*	1.06 (1.02-1.10)*	1.29 (1.24-1.33)*	1.06 (1.02-1.09)*	1.35 (1.26-1.45)*	1.12 (1.04-1.20)*
Hyperthyroid	1.10 (1.03-1.17)*	1.09 (1.02-1.16)*	1.09 (1.01-1.17)	1.08 (1.00-1.15)	1.06 (0.98-1.15)	1.05 (0.97-1.14)	1.05 (0.94- 1.16)	1.04 (0.93-1.15)	0.94 (0.74-1.20)	0.92 (0.71-1.17)
Hypothyroidism	1.27 (1.23-1.30)*	1.18 (1.15-1.22)*	1.25 (1.22-1.29)*	1.18 (1.14-1.21)*	1.24 (1.20-1.28)*	1.17 (1.12-1.20)*	1.22 (1.17-1.28)*	1.16 (1.11-1.21)*	1.25 (1.14-1.36)*	1.19 (1.08-1.30)*
Digestive										
Gastritis	1.42 (1.37-1.46)*	1.42 (1.36-1.45)*	1.43 (1.37-1.47)*	1.42 (1.37-1.47)*	1.45 (1.39-1.50)*	1.45 (1.39-1.50)*	1.46 (1.39-1.53)*	1.45 (1.38-1.52)*	1.55 (1.39-1.71)*	1.55 (1.39-1.72)*
Gastro-intestinal bleed	1.43 (1.34-1.53)*	1.42 (1.33-1.52)*	1.45 (1.36-1.55)*	1.43 (1.34-1.54)*	1.47 (1.36-1.59)*	1.44 (1.33-1.56)*	1.52 (1.38-1.68)*	1.49 (1.34-1.64)*	1.69 (1.39-2.07)*	1.66 (1.36-2.03)*
Gall bladder stone	1.44 (1.39-1.49)*	1.27 (1.22-1.31)*	1.43 (1.36-1.49)*	1.27 (1.22-1.31)*	1.42 (1.37-1.48)*	1.26 (1.21-1.31)*	1.37 (1.30-1.44)*	1.23 (1.17-1.30)*	1.18 (1.06-1.33)	1.05 (0.93-1.18)
Inflammatory Bowel Disease	1.38 (1.33-1.43)*	1.36 (1.32-1.41)*	1.39 (1.34-1.44)*	1.38 (1.33-1.43)*	1.43 (1.37-1.48)*	1.42 (1.36-1.47)*	1.45 (1.38-1.53)*	1.44 (1.36-1.52)*	1.34 (1.20-1.49)*	1.33 (1.19-1.48)*
Liver Disease	1.47 (1.33-1.62)*	1.42 (1.29-1.57)*	1.46 (1.32-1.62)*	1.42 (1.27-1.56)*	1.55 (1.38-1.73)*	1.48 (1.32-1.67)*	1.49 (1.30-1.72)*	1.45 (1.26-1.68)*	1.64 (1.23-2.19)*	1.47 (1.09-1.99)
Irritable Bowel Syndrome	1.47 (1.43-1.51)*	1.52(1.47-1.56)*	1.50(1.45-1.54)*	1.54(1.49-1.58)*	1.50(1.46-1.55)*	1.55(1.49-1.60)*	1.54(1.47-1.61)*	1.58(1.51-1.66)*	1.62(1.20-2.04)*	1.59(1.23-1.95)*
Others										
HIV infection/AIDS	1.99 (0.75-5.32)	2.08 (0.76-5.75)	1.66 (0.65-4.22)	1.64 (0.62-4.33)	1.38 (0.52-3.62)	1.49 (0.54-4.16)	2.99 (0.81-11.08)	3.17 (0.84-12.03)	-	-
Hearing	1.26 (1.24-1.29)*	1.26 (1.23-1.29)*	1.27 (1.24-1.30)*	1.26 (1.23-1.29)*	1.26 (1.23-1.29)*	1.26 (1.22-1.29)*	1.27 (1.23-1.31)*	1.26 (1.22-1.30)*	1.32 (1.24-1.41)*	1.30 (1.22-1.39)*
Psoriasis	1.24 (1.19-1.30)*	1.20 (1.14-1.25)*	1.27 (1.22-1.33)*	1.22 (1.16-1.28)*	1.30 (1.23-1.37)*	1.24 (1.17-1.31)*	1.32 (1.23-1.42)*	1.26 (1.17-1.36)*	1.39 (1.19-1.64)*	1.32 (1.12-1.55)*
Scleroderma	0.98 (0.67-1.43)	0.97 (0.65-1.44)	1.06 (0.71-1.59)	1.05 (0.69-1.59)	1.05 (0.67-1.65)	1.02 (0.64-1.64)	1.59 (0.86-2.91)	1.76 (0.94-3.30)	4.99 (1.09-22.82)	5.75 (1.22-22.09)
Sleep Disorder	1.43 (1.36-1.49)*	1.35 (1.28-1.41)*	1.44 (1.37-1.51)*	1.35 (1.29-1.42)*	1.45 (1.38-1.53)*	1.37 (1.30-1.44)*	1.45 (1.37-1.55)*	1.37 (1.28-1.46)*	1.46 (1.29-1.64)*	1.41 (1.24-1.59)*
Tuberculosis	1.21 (1.04-1.39)	1.25 (1.08-1.45)	1.16 (0.99-1.35)	1.19 (1.02-1.39)	1.25 (1.04-1.50)	1.24 (1.04-1.50)	1.21 (0.94-1.55)	1.23 (0.95-1.59)	1.87 (1.02-3.44)	1.71 (0.91-3.18)
Anaemia	1.25 (1.20-1.29)*	1.25 (1.21-1.30)*	1.28(1.23-1.33)*	1.28(1.23-1.33)*	1.32(1.26-1.37)*	1.31(1.26-1.37)*	1.41(1.33-1.49)*	1.40(1.32-1.48)*	1.47(1.32-1.63)*	1.42(1.28-1.59)*
Vision problem	1.15 (1.07-1.25)	1.11 (1.02-1.21)	1.14 (1.05-1.24)	1.11 (1.00-1.19)	1.17 (1.07-1.29)	1.13 (1.03-1.24)	1.21 (1.07-1.36)	1.17 (1.03-1.32)	0.99 (0.78-1.27)	0.96 (0.74-1.23)

*P value <0.01 adjusted for multiple testing using 'False discovery rate'; [#]Adjusted for age, gender, BMI, Smoking, Alcohol, multimorbidity and index year ^Only for men; COPD- Chronic Obstructive Pulmonary Disease

Appendix Table 7 Association between joint specific OA and comorbidities prior to the index date (Expanded version)

	Hip			Knee			Wrist/Hand			Ankle/Foot		
	1 year	10 years	20 years	1 year	10 years	20 years	1 year	10 years	20 years	1 year	10 years	20 years
Musculoskeletal												
Ankylosing Spondylitis	1.37(0.76-2.49)	1.82(1.49-2.22)*	1.62(1.39-1.90)*	0.91(0.57-1.43)	1.55(1.34-1.79)*	1.55(1.37-1.73)*	2.01(0.90-4.46)	1.57(1.20-2.06)	1.57(1.24-1.96)*	3.58(0.57-22.34)	1.41(0.91-2.18)	1.40(0.96-2.00)
Back pain	2.22(1.99-2.47)*	1.53(1.47-1.61)*	1.66(1.59-1.73)*	1.21(1.11-1.31)*	1.36(1.33-1.41)*	1.51(1.47-1.56)*	1.28(1.06-1.54)	1.40(1.31-1.50)*	1.58(1.49-1.69)*	1.25(0.98-1.60)	1.31(1.19-1.44)*	1.59(1.45-1.73)*
Crystal arthropathy	0.74(0.50-1.09)	1.14(1.01-1.29)	1.21(1.09-1.35)*	1.03(0.81-1.31)	1.47(1.35-1.60)*	1.49(1.39-1.61)*	0.86(0.45-1.59)	1.78(1.42-2.23)*	1.70(1.39-2.08)*	3.07(1.61-5.84)	2.66(2.08-3.40)*	2.56(2.01-3.14)*
Osteoporosis	1.62(1.23-2.14)	1.28(1.13-1.46)*	1.30(1.16-1.46)*	1.25(1.00-1.56)	1.25(1.16-1.40)*	1.25(1.13-1.34)*	0.88(0.56-1.38)	1.25(1.01-1.53)	1.26(1.05-1.53)	1.69(0.83-3.42)	1.33(0.98-1.82)	1.34(1.04-1.85)
Polymyalgia	2.16(1.24-3.74)	1.53(1.22-1.90)*	1.39(1.14-1.69)*	1.20(0.79-1.83)	1.58(1.35-1.86)*	1.56(1.32-1.77)*	1.65(0.53-5.13)	1.78(1.15-2.76)	1.58(1.07-2.35)	0.38(0.07-1.91)	1.37(0.76-2.48)	1.38(0.81-2.37)
Rheumatoid Arthritis	2.95(1.29-6.72)	1.25(0.93-1.67)	1.25(0.99-1.63)	1.21(0.75-1.93)	1.38(1.13-1.69)	1.43(1.21-1.70)*	4.76(1.54-14.79)	1.55(1.07-2.26)	1.57(0.99-1.99)	1.98(0.47-8.30)	1.35(0.77-2.36)	1.30(0.62-1.72)
Sjogren's syndrome	0.80(0.05-13.88)	1.86(0.94-3.64)	1.93(1.08-3.47)	1.59(0.49-5.09)	1.81(1.20-2.71)	1.47(1.04-2.09)	-	0.88(0.36-2.15)	1.32(0.63-2.74)	-	1.24(0.26-5.93)	1.30(0.32-5.22)
SLE	-		-	1.92(0.10-36.03)	1.21(0.54-2.69)	1.19(0.62-2.29)	-	0.28(0.03-2.62)	0.38(0.09-1.38)	-	-	-
Fibromyalgia	1.12(0.52-2.45)	1.49(1.13-1.96)	1.51(1.17-1.92)	1.87(1.21-2.87)	1.84(1.54-2.21)*	1.75(1.49-2.05)*	0.97(0.35-2.74)	1.57(1.13-2.19)	1.53(1.14-2.07)	1.46(0.49-4.31)	1.47(0.90-2.41)	1.29(0.81-2.03)
Fatigue	1.54(0.93-2.55)	1.31(1.06-1.62)	1.32(1.09-1.60)	1.21(0.85-1.71)	1.36(1.17-1.58)*	1.38(1.21-1.59)*	2.32(1.05-5.12)	1.43(1.07-1.89)	1.42(1.09-1.84)	1.41(0.58-3.40)	1.11(0.72-1.70)	1.10(0.66-1.53)
Respiratory												
Asthma	1.16(0.87-1.55)	1.14(1.04-1.25)	1.19(1.11-1.28)*	1.45(1.17-1.79)	1.45(1.36-1.55)*	1.38(1.31-1.46)*	1.46(0.90-2.37)	1.34(1.16-1.54)*	1.31(1.18-1.47)*	1.01(0.54-1.88)	1.47(1.22-1.79)*	1.38(1.18-1.62)*
COPD	1.42(1.06-1.90)	1.21(1.10-1.35)*	1.20(1.11-1.31)*	1.56(1.25-1.94)*	1.36(1.26-1.46)*	1.33(1.25-1.41)*	0.73(0.41-1.28)	1.24(1.05-1.47)	1.23(1.07-1.41)	0.71(0.32-1.57)	1.18(0.93-1.50)	1.25(1.02-1.52)
Genito-Urinary												
CKD	0.97(0.80-1.18)	1.09(0.99-1.20)	1.10(0.99-1.21)	1.02(0.88-1.18)	1.04(0.97-1.12)	1.04(0.97-1.13)	0.80(0.54-1.19)	0.89(0.74-1.07)	0.92(0.77-1.10)	1.08(0.68-1.71)	1.05(0.82-1.34)	1.05(0.82-1.34)
Prostate^	1.67(1.26-2.21)*	1.42(1.27-1.58)*	1.40(1.27-1.55)*	1.09(0.90-1.33)	1.31(1.21-1.41)*	1.32(1.25-1.43)*	1.80(1.06-3.07)	1.56(1.21-1.87)*	1.56(1.29-1.89)*	1.54(0.88-2.71)	1.44(1.14-1.82)	1.40(1.10-1.72)*
Renal stone	0.99(0.47-2.12)	1.13(0.89-1.44)	1.05(0.87-1.28)	2.30(1.30-4.08)	1.41(1.19-1.68)*	1.41(1.15-1.51)*	1.47(0.32-6.78)	0.90(0.58-1.38)	1.22(0.86-1.73)	0.49(0.12-2.04)	1.22(0.76-1.97)	1.60(1.07-2.39)
Neuro/Psychiatric												
Stroke	1.09(0.89-1.34)	1.12(1.04-1.22)	1.09(1.00-1.16)	1.20(1.03-1.39)	1.20(1.13-1.27)*	1.20(1.10-1.22)*	1.40(0.98-1.99)	1.19(1.04-1.37)	1.24(1.09-1.40)*	1.37(0.83-2.24)	1.18(0.97-1.43)	1.17(0.90-1.28)
Dementia	1.44(0.89-2.34)	1.16(0.89-1.50)	1.11(0.86-1.44)	1.28(0.89-1.82)	0.96(0.79-1.16)	0.93(0.77-1.12)	0.63(0.23-1.70)	0.72(0.44-1.18)	0.72(0.44-1.17)	1.89(0.47-7.53)	0.97(0.46-2.06)	0.96(0.45-2.01)
Epilepsy	1.95(0.76-4.97)	1.21(0.90-1.63)	1.27(0.99-1.61)	1.21(0.67-2.18)	1.33(1.08-1.64)	1.29(1.09-1.51)	1.89(0.41-8.73)	1.11(0.66-1.87)	1.12(0.76-1.66)	0.96(0.06-15.51)	1.11(0.50-2.43)	0.81(0.45-1.44)
Multiple sclerosis	1.09(0.29-4.21)	0.95(0.56-1.63)	0.73(0.48-1.10)	0.72(0.19-2.79)	0.97(0.65-1.46)	0.96(0.70-1.29)	1.71(0.10-29.60)	0.90(0.39-2.10)	0.67(0.34-1.32)	2.09(0.33-13.28)	0.95(0.25-3.70)	0.63(0.25-1.61)
Parkinson's Disease	1.34(0.66-2.73)	0.91(0.64-1.29)	0.87(0.62-1.23)	1.43(0.84-2.45)	1.21(0.95-1.55)	1.20(0.96-1.52)	0.15(0.02-1.17)	1.04(0.50-2.17)	1.11(0.55-2.24)	-	1.30(0.55-3.12)	1.49(0.66-3.35)
Migraine	1.05(0.73-1.53)	1.24(1.10-1.40)*	1.16(1.06-1.28)*	1.29(0.99-1.68)	1.38(1.28-1.49)*	1.35(1.27-1.44)*	1.35(0.83-2.18)	1.40(1.20-1.63)*	1.47(1.30-1.67)*	1.23(0.56-2.69)	1.42(1.13-1.79)	1.38(1.14-1.67)*
Depression	1.43(1.18-1.73)*	1.37(1.28-1.46)*	1.32(1.25-1.39)*	1.52(1.34-1.73)*	1.47(1.41-1.54)*	1.46(1.43-1.49)*	1.26(0.97-1.64)	1.48(1.35-1.63)*	1.48(1.34-1.57)*	1.50(1.05-2.15)	1.34(1.18-1.53)*	1.40(1.27-1.60)*
Psychosis	-	1.13(0.67-1.89)	1.09(0.70-1.71)	2.59(0.94-7.16)	1.04(0.75-1.47)	1.05(0.82-1.44)	-	0.54(0.23-1.27)	0.58(0.30-1.12)			
Schizophrenia	0.90(0.31-2.58)	0.98(0.70-1.36)	0.99(0.77-1.29)	1.51(0.69-3.27)	1.09(0.87-1.37)	1.15(0.96-1.38)	1.98(0.18-21.97)	0.83(0.49-1.42)	0.83(0.55-1.25)	-	0.54(0.25-1.18)	0.56(0.38-1.13)
Cancer	1.13(0.86-1.47)	1.27(1.15-1.41)*	1.24(1.13-1.35)*	0.84(0.68-1.01)	1.11(1.03-1.19)	1.11(1.03-1.17)	0.92(0.56-1.51)	0.92(0.78-1.09)	0.92(0.79-1.07)	1.25(0.68-2.31)	1.30(1.01-1.67)	1.16(0.93-1.45)
Circulatory												
CHD	1.06(0.83-1.36)	1.17(1.08-1.27)*	1.18(1.10-1.26)*	1.09(0.92-1.29)	1.15(1.08-1.21)*	1.15(1.09-1.21)*	0.94(0.58-1.53)	1.05(0.90-1.23)	1.02(0.90-1.16)	1.07(0.59-1.93)	1.36(1.11-1.66)	1.38(1.17-1.63)*
Arterial/Venous	1.00(0.55-1.85)	1.35(1.07-1.71)	1.34(1.09-1.65)	1.87(1.13-3.10)	1.17(0.97-1.40)	1.21(1.03-1.42)	1.88(0.51-6.93)	0.96(0.60-1.54)	0.96(0.63-1.52)	0.56(0.09-3.19)	0.68(0.35-1.33)	0.84(0.47-1.51)
Heart failure	1.72(1.11-2.67)	1.34(1.12-1.61)	1.38(1.16-1.63)*	1.27(0.93-1.74)	1.33(1.17-1.52)*	1.34(1.20-1.53)*	1.44(0.44-4.69)	1.17(0.79-1.72)	1.17(0.89-1.83)	1.13(0.33-3.90)	1.82(1.14-2.89)	1.57(1.03-2.38)
Hypertension	1.10(0.97-1.24)	1.10(1.05-1.15)*	1.12(1.07-1.17)*	1.06(0.97-1.16)	1.06(1.03-1.10)*	1.10(1.07-1.15)*	0.85(0.67-1.07)	1.00(0.92-1.09)	1.02(0.95-1.10)	0.93(0.69-1.24)	1.05(0.93-1.17)	1.08(0.97-1.20)
PVD	1.86(1.34-2.59)*	1.41(1.23-1.62)*	1.37(1.21-1.55)*	1.58(1.22-2.04)*	1.29(1.17-1.43)*	1.29(1.16-1.38)*	1.89(1.04-3.43)	1.50(1.18-1.90)	1.50(1.22-1.86)*	1.60(0.82-3.11)	1.56(1.15-2.12)	1.44(1.09-1.89)
Metabolic/Endocrine												
High Cholesterol	1.18(0.99-1.39)	1.13(1.06-1.21)*	1.15(1.09-1.22)*	1.13(1.01-1.26)	1.14(1.09-1.19)*	1.14(1.09-1.19)*	1.29(0.99-1.67)	1.23(1.11-1.37)*	1.22(1.11-1.35)*	1.16(0.81-1.66)	1.10(0.95-1.27)	1.11(1.01-1.33)
Diabetes Mellitus	1.06(0.86-1.31)	1.06(0.98-1.16)	1.06(0.98-1.13)	1.04(0.90-1.19)	1.01(0.96-1.07)	1.02(0.98-1.08)	1.11(0.78-1.57)	1.01(0.87-1.16)	0.97(0.85-1.10)	1.23(0.78-1.96)	1.01(0.83-1.22)	0.95(0.79-1.13)
Hyperthyroid	0.93(0.45-1.92)	1.09(0.85-1.39)	1.13(0.94-1.38)	0.96(0.57-1.64)	1.08(0.92-1.28)	1.13(0.99-1.30)	0.38(0.10-1.43)	0.90(0.64-1.28)	1.05(0.79-1.39)	0.56(0.14-2.25)	1.22(0.71-2.12)	1.15(0.72-1.85)
Hypothyroidism	1.34(1.02-1.74)	1.16(1.05-1.28)	1.23(1.13-1.34)*	1.04(0.94-1.39)	1.10(1.02-1.17)	1.17(1.10-1.24)*	1.13(0.75-1.75)	1.23(1.06-1.43)	1.21(1.07-1.38)*	1.21(0.65-2.24)	0.99(0.78-1.25)	1.11(0.91-1.37)
Digestive												
Gastritis	1.21(0.90-1.61)	1.23(1.10-1.37)*	1.22(1.11-1.34)*	1.53(1.23-1.89)*	1.42(1.32-1.54)*	1.39(1.30-1.47)*	1.30(0.79-2.14)	1.32(1.10-1.59)	1.26(1.09-1.45)*	0.76(0.33-1.74)	1.68(1.31-2.16)*	1.45(1.18-1.78)*
GI bleed	2.62(1.44-4.79)	1.43(1.14-1.78)	1.49(1.23-1.80)*	1.86(1.27-2.72)	1.43(1.23-1.66)*	1.37(1.21-1.56)*	1.06(0.42-2.68)	1.47(1.04-2.09)	1.35(0.99-1.83)	0.92(0.19-4.31)	1.69(1.01-2.85)	1.48(0.96-2.31)
Gall bladder stone	0.95(0.67-1.34)	1.19(1.05-1.34)	1.22(1.11-1.35)*	1.11(0.87-1.42)	1.30(1.20-1.41)*	1.33(1.25-1.43)*	1.08(0.65-1.77)	1.25(1.05-1.49)	1.31(1.13-1.52)*	0.31(0.11-0.85)	1.43(1.09-1.88)	1.45(1.14-1.83)*
IBD	1.14(0.84-1.54)	1.28(1.13-1.44)*	1.23(1.11-1.37)*	1.25(0.99-1.56)	1.38(1.27-1.50)*	1.35(1.26-1.44)*	1.32(0.82-2.12)	1.34(1.12-1.59)	1.22(1.04-1.40)	0.92(0.42-2.05)	1.45(1.12-1.88)	1.63(1.29-2.06)*
Liver Disease	0.91(0.37-2.25)	1.24(0.88-1.76)	1.14(0.85-1.55)	1.33(0.75-2.33)	1.33(1.14-1.83)	1.32(1.08-1.62)	-	0.95(0.56-1.76)	1.18(0.71-1.96)	-	1.58(0.65-3.84)	1.51(0.74-3.06)
Irritable Bowel syndrome	1.17(0.89-1.57)	1.26(1.13-1.39)*	1.33(1.22-1.46)*	1.28(0.99-1.56)	1.35(1.23-1.47)*	1.38(1.29-1.48)*	1.30(0.80-2.10)	1.26(1.10-1.65)	1.25(1.08-1.42)	0.98(0.47-2.02)	1.51(1.23-1.78)*	1.61(1.24-2.02)*
Others												
HIV infection/AIDS	-	2.25(0.33-15.46)	0.86(0.13-5.86)	-	1.94(0.26-14.61)	2.05(0.20-20.69)	-	-	-	-	-	-
Hearing	1.19(0.98-1.45)	1.13(1.05-1.21)	1.14(1.06-1.21)*	1.32(1.16-1.50)*	1.24(1.17-1.30)*	1.24(1.18-1.29)*	1.26(0.94-1.69)	1.32(1.17-1.49)*	1.31(1.18-1.46)*	1.07(0.72-1.60)	1.40(1.18-1.66)*	1.41(1.23-1.67)*
Psoriasis	1.04(0.66-1.63)	1.10(0.93-1.30)	1.07(0.93-1.22)	1.10(.80-1.53)	1.13(1.01-1.26)	1.13(1.04-1.25)	2.31(1.02-5.25)	1.07(0.83-1.39)	1.07(0.86-1.29)	1.92(0.61-6.04)	1.11(0.79-1.56)	1.15(1.01-1.81)
Scleroderma	-	1.88(0.43-8.25)	1.29(0.33-5.06)	1.15(0.08-16.35)	0.78(0.29-2.14)	072(0.26-1.52)	-	-	-			
Sleep Disorder	1.15(0.81-1.62)	1.28(1.09-1.48)	1.25(1.-8-1.43)	1.29(1.02-1.65)	1.44(1.30-1.60)*	1.44(1.26-1.53)*	1.74(1.05-2.86)	1.45(1.15-1.85)	1.44(1.15-1.78)*	1.87(0.85-4.10)	1.47(1.06-2.06)	1.48(1.09-2.02)
Tuberculosis	0.92(0.7-12.02)	0.72(0.43-1.21)	0.86(0.56-1.32)	0.90(0.16-5.04)	1.45(0.98-2.15)	1.35(1.00-1.84)	-	2.32(0.53-10.13)	3.44(1.23-9.58)	-	1.99(0.69-5.71)	2.56(0.93-7.07)
Anaemia	1.35(0.96-1.90)	1.27(1.11-1.46)*	1.21(1.08-1.36)*	1.29(1.04-1.59)	1.26(1.16-1.38)*	1.26(1.16-1.35)*	1.19(0.76-1.89)	1.31(1.09-1.59)	1.31(1.10-1.53)	0.83(0.39-1.79)	1.24(0.92-1.68)	1.14(0.87-1.49)
Vision problem	0.88(0.43-1.80)	1.13(0.87-1.48)	1.05(0.83-1.33)	0.87(0.51-1.48)	1.11(0.93-1.34)	1.11(0.95-1.31)	1.21(0.32-4.58)	2.6(0.79-2.01)	1.27(0.85-1.90)	-	0.59(0.26-1.32)	0.77(0.39-1.51)

*P-value <0.05 adjusted for multiple testing using 'False discovery rate'; Adjusted for age, gender, body mass index (BMI), Smoking, Alcohol use and index year; ^for men only

SLE – Systemic Lupus Erythematous; COPD- Chronic Obstructive Pulmonary Disease; CHD- Coronary Heart Disease; PVD- Peripheral vascular disease; GI – Gastrointestinal; IBD- Inflammatory Bowel Disease

Appendix Table 8 Cumulative probabilities (%) of incident comorbidities after index date

	Osteoarthritis cases					Non-Osteoarthritis controls				
	1 year	5 years	10 years	15 years	20 years	1 year	5 years	10 years	15 years	20 years
Musculoskeletal										
Ankylosing Spondylitis	0.01	0.05	0.14	0.26	0.26	0.01	0.02	0.03	0.10	0.21
Back pain	0.87	2.57	6.30	13.13	18.60	0.36	1.43	4.13	9.20	12.32
Crystal arthropathy	0.03	0.15	0.53	1.70	2.47	0.01	0.06	0.25	0.81	1.41
Osteoporosis	0.05	0.14	0.48	1.58	2.31	0.02	0.05	0.25	1.15	2.36
Polymyalgia	0.01	0.04	0.09	0.18	0.51	0.00	0.01	0.07	0.15	0.29
Rheumatoid Arthritis	0.04	0.06	0.23	0.53	1.00	0.01	0.02	0.07	0.13	0.16
Sjogren's Disease	0.00	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.04	0.07
Systemic lupus erythematous	0.00	0.00	0.00	0.03	0.03	0.00	0.00	0.00	0.02	0.02
Fibromyalgia	0.02	0.05	0.13	0.31	0.45	0.00	0.01	0.01	0.05	0.08
Fatigue	0.01	0.04	0.21	0.51	0.62	0.00	0.02	0.12	0.22	0.49
Respiratory										
Asthma	0.09	0.32	0.80	1.73	2.42	0.09	0.28	0.60	1.24	1.50
COPD	0.04	0.13	0.46	1.29	2.17	0.04	0.11	0.31	0.96	1.48
Genito-Urinary										
Chronic kidney disease	0.01	0.02	1.20	3.09	4.48	0.00	0.00	0.97	2.85	4.51
Prostate^	0.06	0.21	0.53	1.79	2.61	0.03	0.09	0.40	0.92	1.38
Renal stone	0.00	0.02	0.07	0.36	0.63	0.01	0.04	0.11	0.32	0.47
Neuro/Psychiatric										
Stroke	0.24	0.77	2.53	6.17	9.74	0.22	0.69	2.00	4.91	8.61
Dementia	0.01	0.06	0.27	0.95	1.78	0.00	0.01	0.10	0.52	1.11
Epilepsy	0.02	0.05	0.07	0.13	0.13	0.00	0.01	0.06	0.11	0.14
Multiple sclerosis	0.01	0.01	0.03	0.05	0.05	0.00	0.01	0.03	0.05	0.05
Parkinson's Disease	0.01	0.02	0.04	0.15	0.26	0.00	0.01	0.04	0.10	0.35
Migraine	0.08	0.17	0.39	0.87	1.17	0.03	0.10	0.31	0.71	0.92
Depression	0.38	1.15	2.48	4.78	7.69	0.17	0.63	1.43	2.80	4.41
Psychosis	0.01	0.01	0.01	0.06	0.06	0.00	0.00	0.01	0.03	0.14
Schizophrenia	0.01	0.01	0.03	0.15	0.22	0.01	0.01	0.05	0.14	0.17
Cancer	0.07	0.23	0.87	2.57	4.74	0.02	0.13	0.51	1.71	3.31
Circulatory										
Coronary heart disease	0.11	0.35	0.89	1.79	2.61	0.08	0.29	0.57	1.08	1.51
Arterial/Venous	0.00	0.02	0.06	0.32	0.40	0.00	0.01	0.05	0.16	0.29
Heart failure	0.01	0.06	0.16	0.54	0.70	0.01	0.04	0.11	0.28	0.56
Hypertension	0.46	1.55	4.46	7.79	10.58	0.30	0.93	3.22	7.00	9.89
Peripheral vascular disease	0.02	0.10	0.32	0.91	1.17	0.03	0.09	0.18	0.52	0.65
Metabolic/Endocrine										
High Cholesterol	0.11	0.42	1.57	4.11	5.86	0.06	0.30	1.29	3.31	5.10
Diabetes Mellitus	0.09	0.36	1.20	3.67	6.18	0.03	0.12	0.60	2.31	4.01
Hyperthyroid	0.02	0.06	0.13	0.13	0.23	0.00	0.03	0.12	0.20	0.25
Hypothyroidism	0.10	0.20	0.64	1.45	1.81	0.02	0.12	0.48	1.26	1.67
Digestive										
Gastritis	0.09	0.19	0.49	1.32	1.92	0.01	0.06	0.23	0.82	1.94
Gastro-intestinal bleed	0.01	0.03	0.11	0.25	0.60	0.01	0.01	0.06	0.21	0.35
Gall bladder stone	0.04	0.17	0.42	1.22	2.06	0.04	0.09	0.24	0.68	1.09
IBD	0.04	0.15	0.43	1.14	1.45	0.03	0.13	0.35	0.79	1.29
Liver Disease	0.01	0.04	0.09	0.29	0.44	0.00	0.00	0.04	0.12	0.41
Irritable bowel syndrome	0.65	2.00	3.48	4.63	5.66	0.32	1.33	2.31	3.13	3.85
Others										
HIV infection/AIDS										
Hearing	0.14	0.42	1.30	3.80	6.77	0.06	0.23	0.94	2.76	4.73
Psoriasis	0.02	0.07	0.22	0.72	1.00	0.02	0.07	0.17	0.46	0.66
Scleroderma	0.00	0.01	0.02	0.04	0.11	0.00	0.00	0.00	0.00	0.00
Sleep Disorder	0.03	0.10	0.34	0.84	1.56	0.02	0.03	0.20	0.45	1.06
Tuberculosis	0.00	0.00	0.00	0.03	0.03	0.00	0.01	0.06	0.06	0.06
Anaemia	0.06	0.18	0.46	1.41	2.27	0.00	0.08	0.32	0.67	1.26
Vision problem	0.01	0.04	0.11	0.12	0.25	0.01	0.03	0.05	0.11	0.18

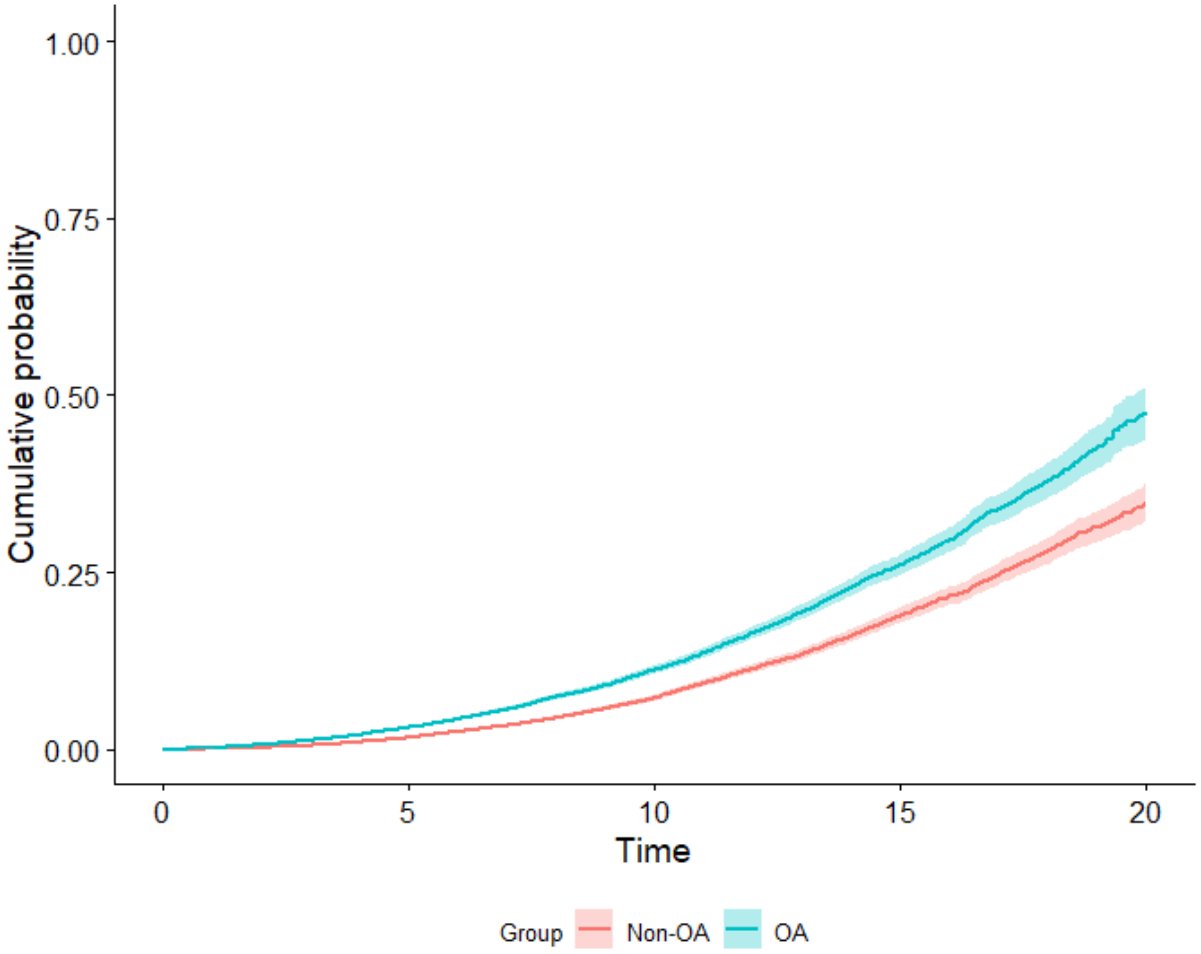
COPD- Chronic Obstructive Pulmonary Disease; ^only men

Appendix Table 9 Hazard ratio and 95% confidence interval for each comorbidity comparing incident OA cases and controls without any comorbidities at the index date

	Unadjusted HR (95% CI)	Adjusted HR (95% CI)	p value
Two or more comorbidities	1.38(1.31-1.45)	1.34(1.28-1.41)	0.001*
Musculoskeletal			
Ankylosing Spondylitis	1.86(1.13-3.05)	1.85(1.13-3.10)	0.028*
Back pain	1.46(1.37-1.55)	1.45(1.36-1.54)	0.001*
Gout	1.57(1.30-1.91)	1.40(1.15-1.70)	0.002*
Osteoporosis	1.38(1.13-1.69)	1.61(1.32-1.98)	0.001*
Polymyalgia	1.48(0.92-2.38)	1.60(0.99-2.59)	0.088
Rheumatoid Arthritis	4.25(2.65-6.82)	4.31(2.68-6.92)	0.001*
Sjogren's syndrome	2.12(0.62-7.26)	2.22(0.64-7.70)	0.279
Systemic lupus erythematosus	2.24(0.41-12.30)	2.45(0.44-13.62)	0.369
Fibromyalgia	5.28(2.66-10.48)	5.29(2.65-10.50)	0.001*
Fatigue	1.25(0.89-1.76)	1.25(0.89-1.77)	0.265
Respiratory			
Asthma	1.15(0.97-1.36)	1.09(0.92-1.29)	0.368
COPD	1.22(0.99-1.49)	1.19(0.98-1.46)	0.088
Genito-Urinary			
Chronic Kidney Disease	1.17(1.02-1.35)	1.14(0.99-1.32)	0.098
Benign prostatic hypertrophy^	1.55(1.27-1.88)	1.56(1.28-1.90)	0.001*
Renal stone	0.91(0.61-1.37)	0.81(0.54-1.22)	0.369
Neuro/Psychiatric			
Stroke	1.15(1.05-1.22)	1.14(1.06-1.24)	0.001*
Dementia	1.43(1.07-1.90)	1.77(1.32-2.38)	0.001*
Epilepsy	0.87(0.49-1.54)	0.88(0.49-1.56)	0.698
Multiple sclerosis	0.85(0.32-2.28)	0.75(0.28-2.03)	0.608
Parkinson's disease	1.21(0.68-2.12)	1.32(0.74-2.34)	0.398
Migraine	1.28(1.02-1.59)	1.27(1.02-1.59)	0.064
Depression	1.58(1.43-1.74)	1.55(1.40-1.71)	0.001*
Psychosis	1.44(0.63-3.35)	1.38(0.59-3.24)	0.488
Schizophrenia	1.30(0.73-2.31)	1.25(0.70-2.23)	0.488
Cancer	1.46(1.27-1.69)	1.43(1.24-1.65)	0.001*
Circulatory			
Coronary Heart Disease	1.27(1.07-1.51)	1.19(1.01-.142)	0.075
Arterial/Venous	1.18(0.72-1.96)	1.27(0.76-2.11)	0.398
Heart failure	1.61(1.09-2.35)	1.62(1.10-2.39)	0.022*
Hypertension	1.15(1.06-1.24)	1.06(0.98-1.14)	0.225
Peripheral Vascular Disease	1.56(1.18-2.05)	1.57(1.19-2.09)	0.002*
Metabolic/Endocrine			
High Cholesterol	1.19(1.08-1.33)	1.15(1.04-1.29)	0.014*
Diabetes Mellitus	1.43(1.26-1.62)	1.26(1.11-1.43)	0.001*
Hyperthyroid	1.06(0.68-1.66)	1.05(0.67-1.66)	0.832
Hypothyroidism	1.21(1.02-1.45)	1.15(0.96-1.37)	0.204
Digestive			
Gastritis	1.46(1.18-1.79)	1.41(1.15-1.74)	0.002*
Gastrointestinal bleed	1.95(1.16-3.28)	1.93(1.14-3.27)	0.027*
Gall bladder stone	1.48(1.18-1.85)	1.31(1.05-1.64)	0.034*
Inflammatory bowel disease	1.33(1.08-1.65)	1.31(1.06-1.62)	0.026*
Liver Disease	3.55(2.01-6.26)	3.36(1.89-5.97)	0.001*
Irritable bowel syndrome	1.43(1.26-1.62)	1.43(1.27-1.63)	0.001*
Others			
HIV infection/AIDS	1.23(0.08-19.65)	0.85(0.50-14.21)	0.907
Hearing	1.30(1.15-1.46)	1.31(1.16-1.48)	0.001*
Psoriasis	1.37(1.05-1.79)	1.31(1.00-1.72)	0.082
Scleroderma	3.57(0.37-34.52)	3.79(0.39-37.19)	0.324
Sleep Disorder	1.95(1.50-2.53)	1.95(1.50-2.55)	0.001*
Tuberculosis	0.47(0.16-1.33)	0.48(0.17-1.38)	0.251
Anaemia	1.58(1.28-1.93)	1.57(1.27-1.95)	0.001*
Vision problem	1.35(0.64-2.85)	1.56(0.73-3.34)	0.324
Cataract	1.07(0.99-1.5)	1.12(1.04-1.21)	0.005*

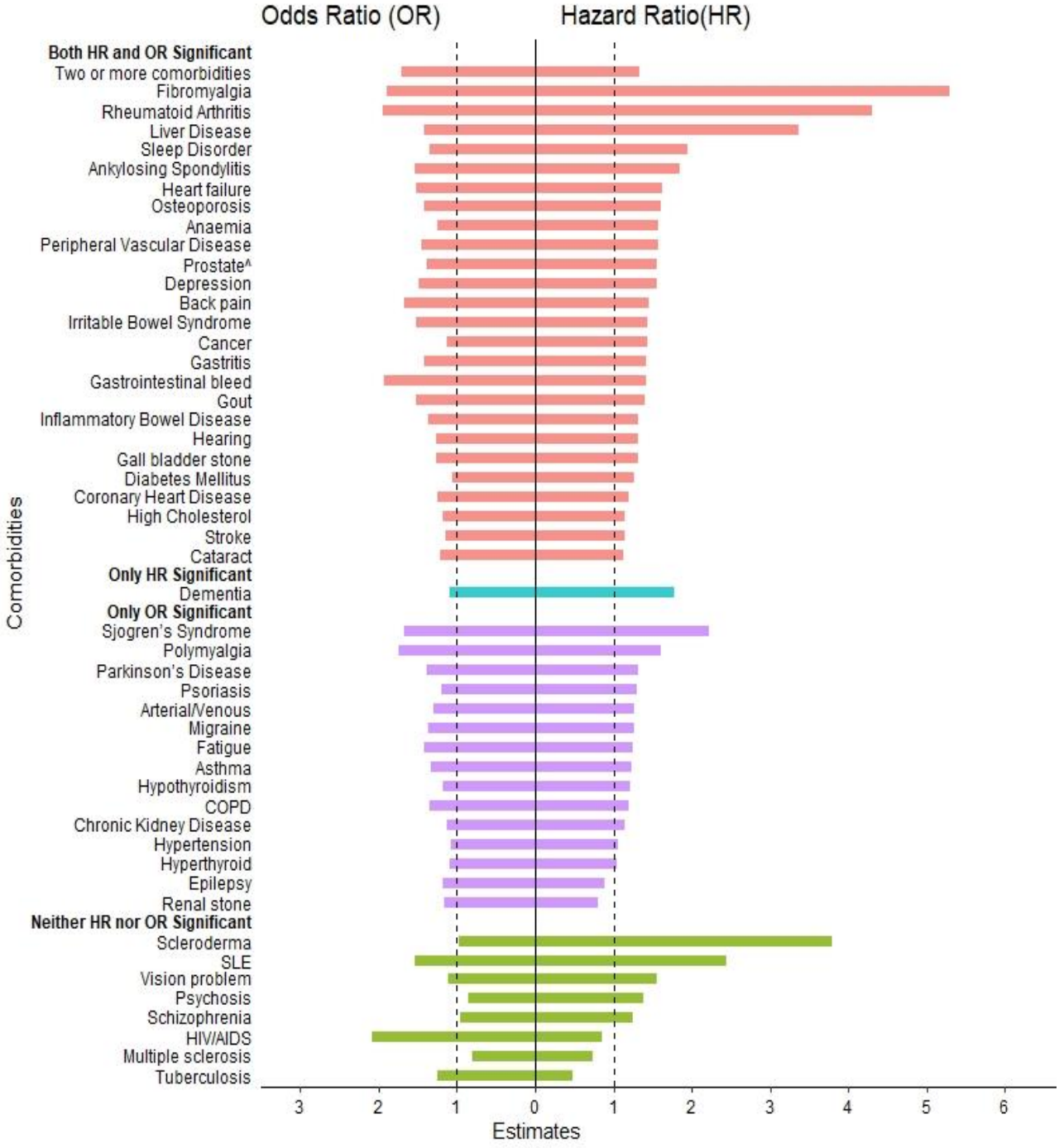
Adjusted for age, sex, BMI, alcohol use, smoking and index date; *p <0.05 'False discovery rate' (FDR) adjusted; p-y person years; COPD- Chronic Obstructive Pulmonary Disease

Appendix Figure 8 Cumulative probabilities of developing multimorbidity in cases with OA and matched non-OA controls without any comorbidities at the index date



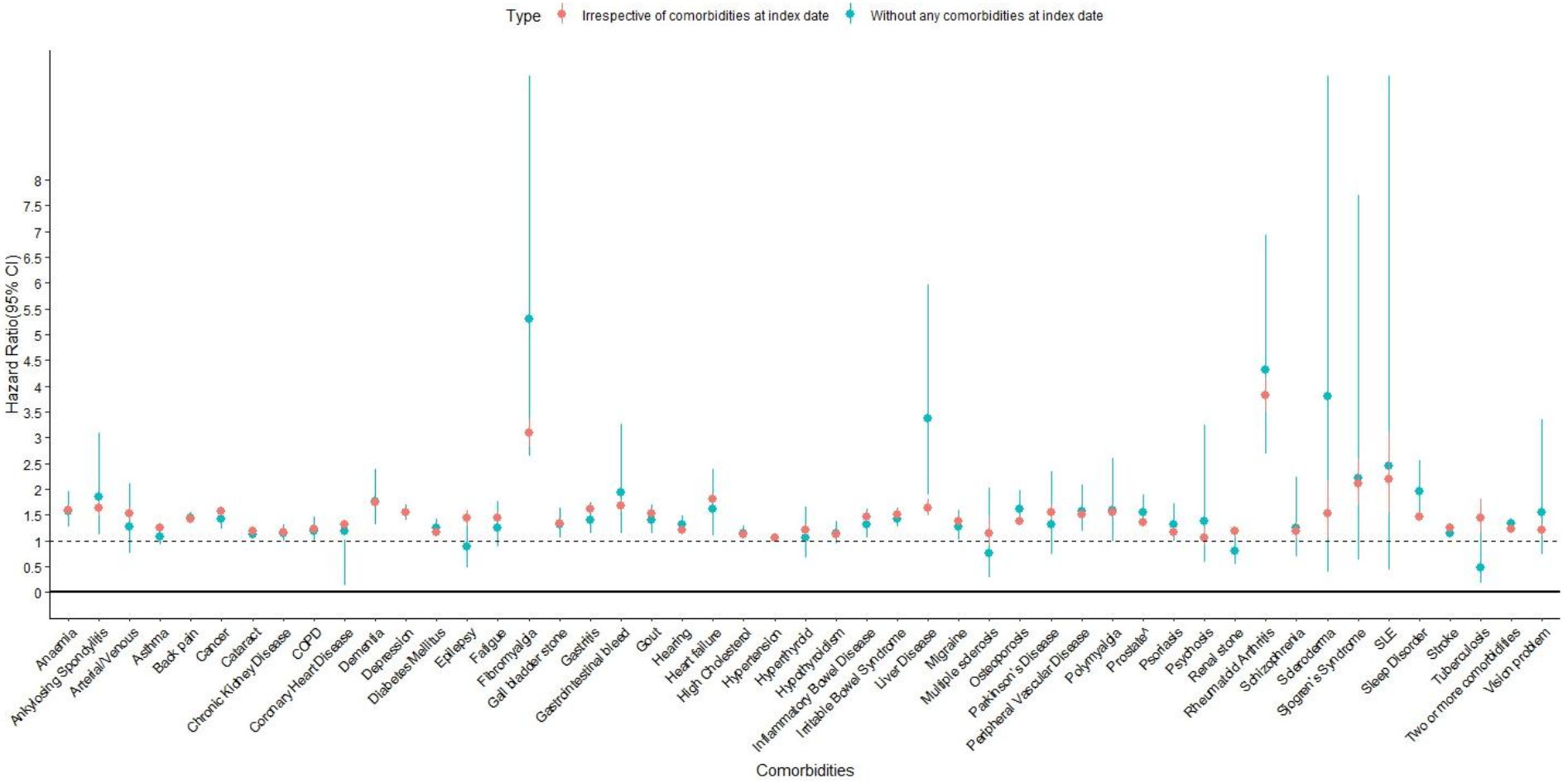
OA: Osteoarthritis (cases); Non-OA: Non-Osteoarthritis (controls)

Appendix Figure 9 Comparison of adjusted Odds Ratio and Hazard Ratio for comorbidities in OA for 20 years observation period among OA and matched controls without any comorbidities at the index date



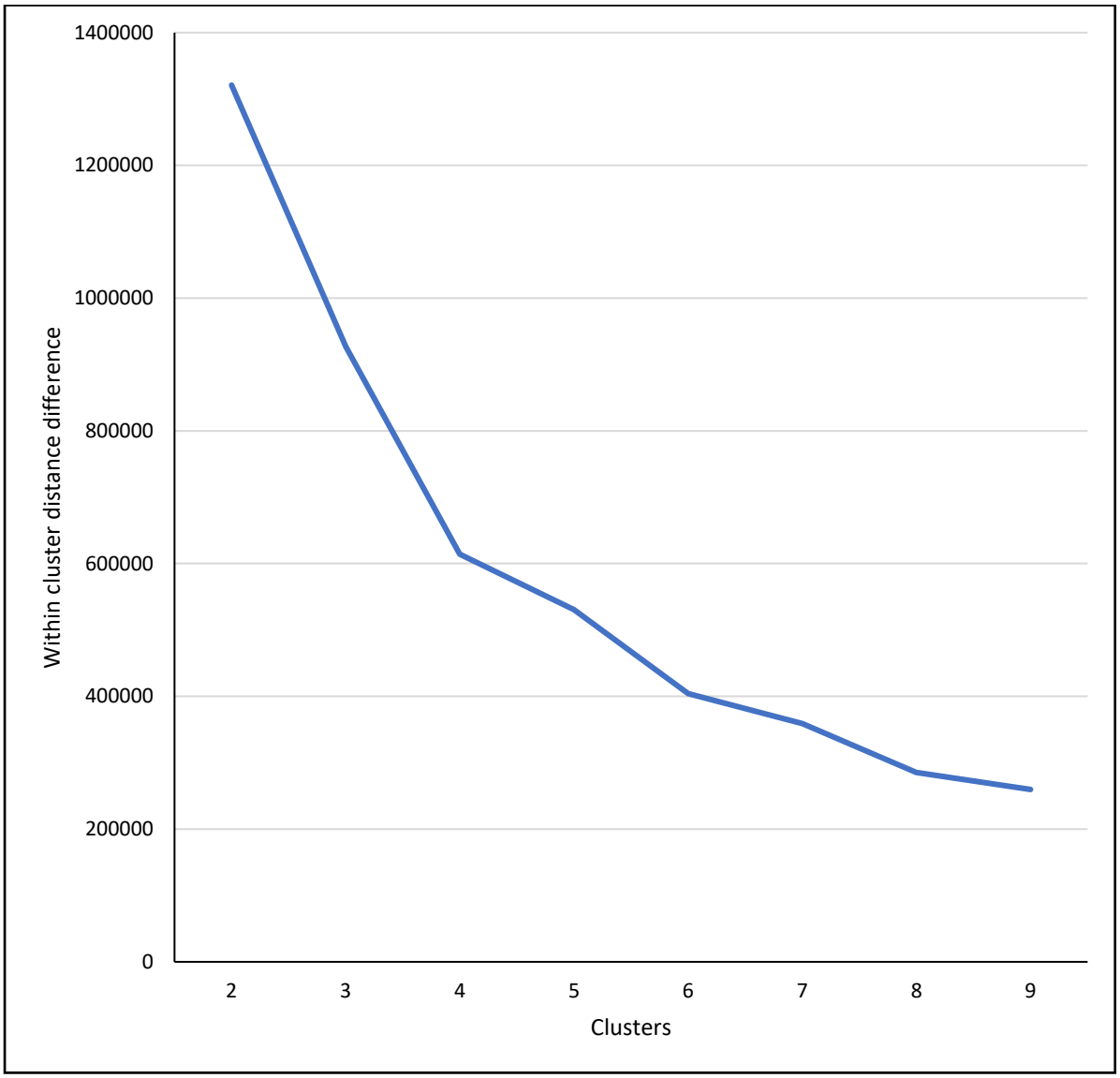
COPD- Chronic obstructive pulmonary diseases; SLE- Systemic lupus erythematous; *p <0.05; ^Benign prostate hypertrophy -Only men
Red: Both HR and OR significant; Blue: Only HR significant; Purple: Only OR significant; Green: Neither HR nor OR significant

Appendix Figure 10 Comparison of the adjusted hazard ratios comparing the analyses for “OA without any comorbidity” at index date and “OA without the specific comorbidity” at the index date with respective matched controls

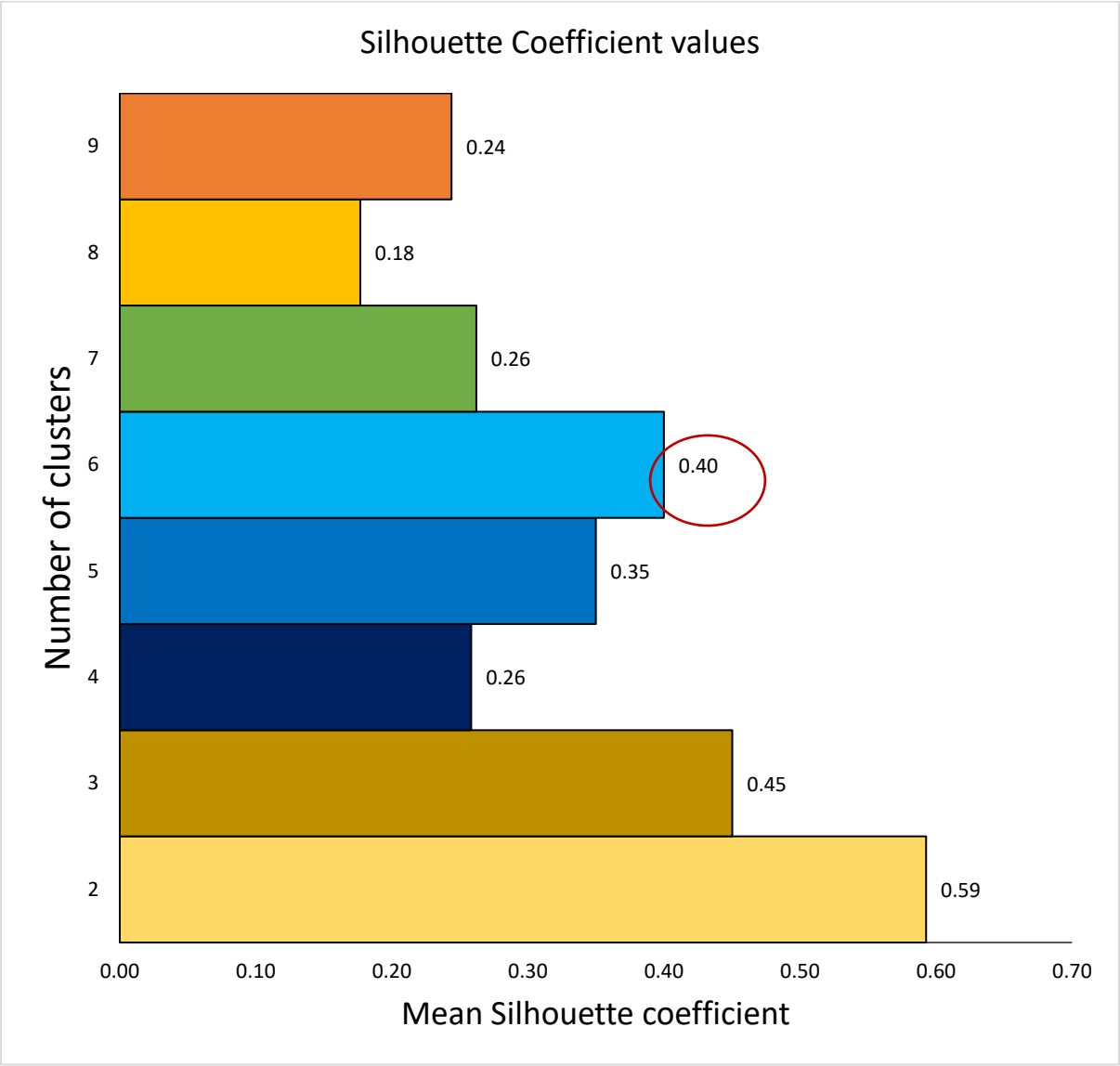


COPD- Chronic Obstructive Pulmonary Disease; SLE- Systemic Lupus Erythematous ; ^Benign prostate hypertrophy -Only men

Appendix Figure 11 Within cluster distance difference across the K-mode classes



Appendix Figure 12 Silhouette coefficient index in K-mode



1.

Appendix Table 10 K-mode (Machine learning) approach for clustering categorical data

	Class 1 (MSK)	Class2 (IBS)	Class3 (Healthy)	Class 4 (MSK-MH)	Class 5 CV-MH	Class 6 CV-MSK-Metabolic
	22.22	2.90	57.29	12.27	1.40	3.92
Anaemia	4.95	2.95	1.95	8.39	7.71	11.92
Ankylosing Spondylitis	1.84	0.03	0.03	2.71	0.05	2.8
Arterial/Venous	0.65	0.17	0.22	0.84	1.3	2.54
Asthma	9.5	5.73	3.98	14.97	12.65	14.27
Back pain	**100**	0	0	**100**	0	**94.85**
Benign prostatic hypertrophy^	4.4	1.43	1.29	4	4.75	10.92
Cancer	6.28	2.6	2.54	7.45	9.93	15.08
Cataract	2.09	3.57	3.34	4.88	10.99	50.96
Chronic Heart Disease	5.84	1.61	2.11	7.96	11.97	20.34
Chronic Kidney Disease	6.18	2.32	2.66	9.18	18.68	31.77
COPD	4.67	1.66	1.46	8.43	7.81	11.57
Gout	4.12	1.16	1.65	4.16	7.03	12.81
Dementia	0.97	0.4	0.55	2.05	4.56	5.77
Depression	0	16.72	8.75	**100**	**100**	8.11
Diabetes	10.21	3.92	4.48	14.59	25.47	29.71
Epilepsy	0.89	0.43	0.5	1.59	1.79	1.39
Fatigue	1.79	1.72	0.62	4.47	2.73	2.88
Fibromyalgia	0.84	1.16	0.24	3.76	1.57	1.19
Gall stones	5.37	3.51	1.87	9.7	8.7	12.92
Gastritis	6.61	3.67	1.99	11.38	8.79	14.07
Gastrointestinal bleed	1.23	0.7	0.44	2.43	1.95	2.77
Hearing problem	12.47	5.24	4.5	16.2	16.77	31.3
Heart failure	0.8	0.21	0.35	1.15	2.11	4.33
High Cholesterol	14.93	5.45	5.41	19.55	28.21	36.94
HIV/AIDS	0.01	0.01	0.01	0.02	0.03	0.01
Hypertension	19.61	6.33	9.03	27.9	**100**	**92.46**
Hyperthyroid	1.34	0.83	0.56	2.16	2.19	2.69
Hypothyroidism	6.09	3.76	2.6	10.31	11.76	13.9
Inflammatory bowel disease	5.91	4.31	1.91	9.98	5.87	8.72
Irritable Bowel Syndrome	7.81	**100**	0	16.2	8.03	9.52
Liver Diseases	0.61	0.28	0.34	1.17	1.38	0.89
Migraine	8.08	6.33	2.85	15.4	8.81	8.17
Multiple sclerosis	0.36	0.28	0.21	0.69	0.53	0.25
Osteoarthritis	9.53	12.52	7.73	21.4	21.27	**84.45**
Osteoporosis	3.72	1.5	1.09	5.85	5.92	13.84
Parkinson's Disease	0.38	0.09	0.13	0.54	0.86	1.02
Peripheral Vascular Diseases	2.47	1.19	0.85	3.92	3.94	7.12
Polymyalgia	0.77	0.25	0.23	1.06	1.48	3.92
Psoriasis	3.43	1.79	1.54	4.37	4.34	4.74
Psychosis	0.15	0.21	0.21	0.8	1.11	0.2
Renal stone	1.69	0.7	0.57	2.01	1.78	2.72
Rheumatoid Arthritis	0.85	0.44	0.33	1.47	1.36	2.19
Schizophrenia	0.29	0.57	0.51	2.32	3.35	0.44
Scleroderma	0.04	0.03	0.02	0.07	0.09	0.08
Sjogren's Syndrome	0.15	0.15	0.04	0.28	0.29	0.51
Sleep problem	2.36	1.49	0.86	6.11	4.67	5.08
Stroke	7.85	7.79	6.62	9.12	13.84	17.41
Systemic Lupus Erythematosus	0.07	0.05	0.02	0.13	0.09	0.11
Tuberculosis	0.39	0.14	0.16	0.43	0.37	0.8
Vision problem	0.36	0.12	0.18	0.66	0.95	2.09

COPD- Chronic obstructive pulmonary disease; CVD- Cardiovascular; IBS- Irritable bowel syndrome; DEP- Depression; MSK- Musculoskeletal; ^only men

Appendix Table 11 Janssen Shannon Index for similarity

		Training data					
		Healthy	MSK	CV	CV-MSK	MSK-MH	THY
Testing data	Healthy	0.026	2.33	6.85	15.68	13.78	19.89
	MSK	2.35	0.029	0.89	5.59	4.4	8.45
	CVD	5.69	1.12	0.003	2.28	1.52	4.36
	MSK, CVD	14.32	6.24	2.03	0.001	0.065	0.42
	MSK, DEP	13.41	5.61	1.65	0.029	0.013	0.63
	THY	18.78	9.58	4.24	0.42	0.865	0.001

Janssen Shannon Index for similarity

Jensen-Shannon distance (JSD, range = [0-1], high values indicate a higher degree of divergence) for the similarities between corresponding clusters profiles (for 50 chronic conditions) in the training (rows) and test sets (columns). Each cluster in the test set is matched with a cluster in the training set with the smallest JSD (with zero indicating perfect similarity). Matched clusters are shaded in dark grey (secondary choices ae shaded in light grey).

Appendix Table 12 Class distribution in Training data

	Healthy (41.90%)	MSK (26.68%)	CV (15.97%)	CV-MSK (5.55%)	MSK-MH (7.20%)	Metabolic (2.67%)
Anaemia	0.29	4.38	3.31	18.35	10.93	10.23
Ankylosing Spondylitis	0.00	0.81	0.52	2.96	5.45	0.51
Arterial/Venous	0.01	0.06	1.10	4.80	0.43	0.19
Asthma	0.99	9.78	9.28	15.40	19.88	11.65
Back pain	6.16	**55.65**	52.85	**80.77**	**90.10**	50.16
Benign Prostatic Hypertrophy^	0.12	1.15	8.27	13.61	4.06	0.81
Cancer	0.38	2.66	11.36	19.20	8.81	7.34
Cataract	2.63	0.08	10.12	36.26	4.96	4.09
Chronic Heart Disease	0.12	0.73	11.70	34.28	5.61	2.80
Chronic Kidney Disease	0.03	0.33	14.46	46.42	5.50	7.64
COPD	0.10	2.05	6.76	17.53	10.39	4.96
Dementia	0.03	0.07	2.18	10.27	0.89	2.57
Depression	1.77	31.78	15.34	37.45	**62.52**	45.07
Diabetes	0.51	3.31	22.00	38.22	13.39	20.66
Epilepsy	0.11	0.87	1.13	2.24	1.78	2.13
Fatigue	0.02	1.57	0.54	4.05	9.25	3.66
Fibromyalgia	0.00	0.50	0.00	1.31	8.82	1.57
Gall stones	0.20	3.42	5.54	17.34	16.08	6.82
Gastritis	0.15	4.14	5.54	21.79	19.40	4.70
Gastrointestinal Bleed	0.01	0.96	0.98	4.95	3.55	1.00
Gout	0.27	1.22	9.31	15.31	2.59	0.82
Hearing Problem	0.74	7.03	18.06	39.09	19.90	10.53
Heart Failure	0.00	0.01	1.40	8.68	0.11	0.40
High Cholesterol	0.61	4.86	29.72	43.17	23.69	15.56
HIV/AIDS	0.00	0.03	0.02	0.01	0.01	0.03
Hypertension	1.21	5.88	**57.15**	**75.66**	29.30	23.60
Hyperthyroidism	0.04	0.00	0.00	4.23	1.09	26.35
Hypothyroidism	0.28	2.35	5.39	19.72	11.87	**51.60**
Inflammatory Bowel Disease	0.29	5.74	3.40	11.85	16.92	4.22
Irritable Bowel Syndrome	3.90	7.84	2.51	10.72	30.17	8.44
Liver Disease	0.06	0.58	0.76	1.65	1.55	1.22
Migraine	0.52	9.68	3.60	8.78	25.27	9.57
Multiple Sclerosis	0.04	0.55	0.28	0.28	0.93	0.83
Osteoarthritis	6.81	3.22	22.66	54.73	33.76	12.39
Osteoporosis	0.08	0.96	4.79	18.54	8.61	5.51
Parkinson’s Disease	0.01	0.05	0.75	1.79	0.36	0.73
Peripheral Vascular Disease	0.10	1.29	2.94	11.85	4.44	2.14
Polymyalgia	0.01	0.05	1.22	5.33	1.24	0.68
Psoriasis	0.47	3.36	3.96	5.62	4.94	3.40
Psychosis	0.00	0.00	0.00	0.37	0.03	9.54
Renal Stone	0.10	1.05	2.44	3.49	2.36	0.78
Rheumatoid Arthritis	0.05	0.44	0.92	2.51	2.85	1.09
Schizophrenia	0.01	0.49	0.02	1.27	0.99	16.14
Scleroderma	0.00	0.02	0.04	0.13	0.13	0.13
Sjogren’s Syndrome	0.00	0.05	0.07	0.59	0.66	0.45
Sleep Problem	0.00	2.25	2.23	7.26	8.31	4.07
Stroke	6.03	4.33	11.90	24.66	8.24	7.70
Systemic Lupus Erythematosus	0.00	0.04	0.03	0.14	0.29	0.20
Tuberculosis	0.05	0.30	0.40	1.04	0.49	0.39
Vision Problem	0.03	0.09	0.55	3.05	0.49	1.17

COPD- Chronic obstructive pulmonary disease; ^only men; CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

Appendix Table 13 Class distribution in Testing data

	Healthy (43.78%)	MSK (24.92%)	CV (16.47%)	CV-MSK (5.73%)	MSK-MH (6.53%)	Metabolic (2.57%)
Anaemia	0.39	4.89	3.27	18.20	11.39	9.05
Ankylosing Spondylitis	0.00	0.98	0.57	2.57	5.69	0.24
Arterial/Venous	0.00	0.07	1.02	4.71	0.48	0.15
Asthma	1.26	10.49	9.43	14.59	20.78	11.45
Back pain	7.15	**58.97**	53.26	**78.64**	**90.62**	47.01
Benign Prostatic Hypertrophy^	0.12	1.25	7.89	13.95	3.95	0.97
Cancer	0.38	2.86	10.84	19.58	8.81	7.23
Cataract	2.48	0.03	9.56	36.60	5.96	4.29
Chronic Heart Disease	0.12	0.79	10.99	33.95	6.04	3.40
Chronic Kidney Disease	0.04	0.35	13.22	47.20	6.46	8.02
COPD	0.12	2.24	6.24	17.78	11.09	4.86
Dementia	0.02	0.07	1.97	10.68	0.91	2.68
Depression	2.42	34.00	15.47	35.97	**64.74**	45.84
Diabetes	0.52	3.31	21.49	38.23	14.37	21.93
Epilepsy	0.12	0.92	1.01	2.66	1.67	1.96
Fatigue	0.02	1.98	0.51	3.32	10.02	2.82
Fibromyalgia	0.00	0.72	0.00	0.93	9.71	1.20
Gall stones	0.27	3.66	5.48	17.19	16.74	6.41
Gastritis	0.16	4.88	5.38	21.15	19.75	4.15
Gastrointestinal Bleed	0.02	1.09	0.89	4.95	3.65	1.30
Gout	0.33	1.29	8.84	15.85	2.97	1.29
Hearing Problem	0.83	7.57	17.23	38.95	20.57	10.00
Heart Failure	0.00	0.00	1.07	9.05	0.17	0.56
High Cholesterol	0.57	5.20	28.95	42.75	25.46	14.93
HIV/AIDS	0.00	0.02	0.01	0.00	0.04	0.01
Hypertension	1.22	5.51	**55.95**	**75.90**	32.13	24.08
Hyperthyroidism	0.04	0.00	0.00	4.16	1.82	27.27
Hypothyroidism	0.30	2.70	4.98	19.13	13.53	**50.05**
Inflammatory Bowel Disease	0.42	6.29	3.39	10.84	16.97	4.36
Irritable Bowel Syndrome	3.96	8.71	2.73	9.69	30.22	7.08
Liver Disease	0.07	0.62	0.79	1.70	1.35	1.41
Migraine	0.64	10.93	3.74	7.73	25.40	9.03
Multiple Sclerosis	0.05	0.60	0.29	0.14	1.00	0.66
Osteoarthritis	6.56	3.56	22.04	53.94	36.82	11.41
Osteoporosis	0.08	1.04	4.54	17.99	9.19	4.81
Parkinson Disease	0.01	0.07	0.64	1.71	0.49	0.67
Peripheral Vascular Disease	0.11	1.35	2.83	11.75	4.74	1.73
Polymyalgia	0.01	0.07	1.13	5.30	1.47	0.42
Psoriasis	0.53	3.31	4.01	5.45	5.20	3.42
Psychosis	0.00	0.00	0.00	0.37	0.09	10.54
Renal Stone	0.10	1.13	2.38	3.78	2.40	1.11
Rheumatoid Arthritis	0.05	0.57	0.89	2.49	3.58	0.84
Schizophrenia	0.02	0.45	0.00	1.40	1.24	18.42
Scleroderma	0.00	0.02	0.04	0.16	0.19	0.12
Sjogren's Disease	0.01	0.04	0.09	0.61	0.80	0.25
Sleep Problem	0.04	2.50	2.31	6.89	8.80	4.03
Stroke	5.87	4.36	11.50	24.65	9.24	7.96
Systemic Lupus Erythematous	0.00	0.05	0.02	0.16	0.32	0.22
Tuberculosis	0.07	0.35	0.40	1.01	0.51	0.30
Vision Problem	0.01	0.12	0.50	3.23	0.58	1.20

COPD- Chronic obstructive pulmonary disease; ^only men; CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

Appendix Table 14 Descriptive statistics Training dataset (N= 11,40,658)

Variables	Healthy(n=465234)	MSK (n=299323)	CV (n=168796)	CV-MSK (n=53201)	MSK-MH (n=65265)	Metabolic (n=18629)
Gender						
Men	249,736 (53.68)	135,165 (45.16)	97,024 (57.48)	22,446 (42.19)	16,510 (25.30)	4,665 (25.04)
Women	215,498 (46.32)	164,158 (54.84)	71,772 (42.52)	30,755 (57.81)	48,755 (74.70)	13,964 (74.96)
Age						
20-39	156,189 (33.57)	70,028 (23.40)	1,800 (1.07)	16 (0.03)	1,613 (2.47)	1,538 (8.26)
40-59	188,756 (40.57)	169,747 (56.71)	38,715 (22.94)	1,845 (3.47)	26,134 (40.04)	8,319 (44.66)
60-79	97,769 (21.02)	56,378 (18.84)	97,738 (57.90)	24,198 (45.48)	33,014 (50.58)	7,434 (39.91)
>80	22,520 (4.84)	3,170 (1.06)	30,543 (18.09)	27,142 (51.02)	4,504 (6.90)	1,338 (7.18)
Smoking						
Never smoked	251,165 (53.99)	165,247 (55.21)	92,545 (54.83)	26,359 (49.55)	33,229 (50.91)	9,735 (52.26)
Current smoker	91,914 (19.76)	77,355 (25.84)	27,572 (16.33)	7,665 (14.41)	15,837 (24.27)	4,893 (26.27)
Ex-smoker	70,359 (15.12)	53,567 (17.90)	48,033 (28.46)	19,043 (35.79)	16,083 (24.64)	3,909 (20.98)
Missing	51,796 (11.13)	3,154 (1.05)	646 (0.38)	134 (0.25)	116 (0.18)	92 (0.49)
Alcohol use						
Never	65,565 (14.09)	42,709 (14.27)	27,429 (16.25)	14,089 (26.48)	13,098 (20.07)	4,539 (24.37)
Ex-drinker	5,293 (1.14)	5,076 (1.70)	4,012 (2.38)	2,320 (4.36)	2,063 (3.16)	742 (3.98)
Current (1-9)	124,391 (26.74)	95,186 (31.80)	52,051 (30.84)	15,314 (28.79)	21,777 (33.37)	5,722 (30.72)
Current (>=10)	59,386 (12.76)	51,731 (17.28)	35,787 (21.20)	6,868 (12.91)	9,405 (14.41)	2,004 (10.76)
Current (Unknown)	91,023 (19.56)	67,720 (22.62)	42,431 (25.14)	13,416 (25.22)	16,658 (25.52)	4,673 (25.08)
Missing	119,576 (25.70)	36,901 (12.33)	7,086 (4.20)	1,194 (2.24)	2,264 (3.47)	949 (5.09)
BMI						
Underweight	12,964 (2.79)	7,181 (2.40)	1,661 (0.98)	891 (1.67)	1,194 (1.83)	384 (2.06)
Normal	146,825 (31.56)	115,370 (38.54)	42,375 (25.10)	13,365 (25.12)	20,694 (31.71)	6,413 (34.42)
Overweight	116,024 (24.94)	87,705 (29.30)	64,455 (38.19)	19,676 (36.98)	21,294 (32.63)	5,757 (30.90)
Obese	73,667 (15.83)	52,655 (17.59)	52,529 (31.12)	17,766 (33.39)	19,780 (30.31)	5,061 (27.17)
Missing	115,754 (24.88)	36,412 (12.16)	7,776 (4.61)	1,503 (2.83)	2,303 (3.53)	1,014 (5.44)
Mean age (SD)	48.94 (16.62)	49.20 (12.43)	68.20 (12.07)	78.78 (9.59)	62.01 (11.68)	58.94 (13.77)
Mean BMI (SD)	26.30 (5.43)	26.22 (5.22)	28.37 (5.31)	28.47 (5.54)	27.95 (5.91)	27.51 (5.89)
Mean Multimorbidity (SD)	0.29 (0.56)	2.01 (0.98)	3.93 (1.37)	8.72 (2.00)	6.04 (1.61)	4.63 (1.72)

BMI- Body mass index; SD- Standard deviation; CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

Appendix Table 15 Descriptive statistics Testing dataset (n=285167)

	Healthy(n=159020)	MSK (n=97411)	CV (n=56923)	CV-MSK (n=18744)	MSK-MH (n=19920)	Metabolic (n=5981)
Gender						
Men	84,504 (53.14)	43,417 (44.57)	32,371 (56.87)	8,090 (43.16)	4,818 (24.19)	1,589 (26.57)
Women	74,516 (46.86)	53,994 (55.43)	24,552 (43.13)	10,654 (56.84)	15,102 (75.81)	4,392 (73.43)
Age						
20-39	53,242 (33.48)	22,379 (22.97)	676 (1.19)	5 (0.03)	390 (1.96)	493 (8.24)
40-59	64,660 (40.66)	55,434 (56.91)	13,543 (23.79)	579 (3.09)	7,423 (37.26)	2,671 (44.66)
60-79	33,607 (21.13)	18,611 (19.11)	32,916 (57.83)	8,410 (44.87)	10,604 (53.23)	2,402 (40.16)
>80	7,511 (4.72)	987 (1.01)	9,788 (17.20)	9,750 (52.02)	1,503 (7.55)	415 (6.94)
Smoking						
Never smoked	86,292 (54.26)	53,276 (54.69)	31,500 (55.34)	9,363 (49.95)	10,124 (50.82)	3,136 (52.43)
Current smoker	31,136 (19.58)	25,738 (26.42)	9,151 (16.08)	2,558 (13.65)	4,858 (24.39)	1,539 (25.73)
Ex-smoker	24,171 (15.20)	17,382 (17.84)	16,050 (28.20)	6,767 (36.10)	4,898 (24.59)	1,283 (21.45)
Missing	17,421 (10.96)	1,015 (1.04)	222 (0.39)	56 (0.30)	40 (0.20)	23 (0.38)
Alcohol use						
Never	22,497 (14.15)	14,036 (14.41)	9,217 (16.19)	4,826 (25.75)	4,119 (20.68)	1,462 (24.44)
Ex-drinker	1,839 (1.16)	1,767 (1.81)	1,345 (2.36)	777 (4.15)	691 (3.47)	267 (4.46)
Current (1-9)	42,405 (26.67)	30,859 (31.68)	17,507 (30.76)	5,524 (29.47)	6,553 (32.90)	1,768 (29.56)
Current (>=10)	20,527 (12.91)	16,841 (17.29)	12,163 (21.37)	2,450 (13.07)	2,909 (14.60)	633 (10.58)
Current (Unknown)	31,249 (19.65)	22,267 (22.86)	14,342 (25.20)	4,747 (25.33)	5,042 (25.31)	1,540 (25.75)
Missing	40,503 (25.47)	11,641 (11.95)	2,349 (4.13)	420 (2.24)	606 (3.04)	311 (5.20)
BMI						
Underweight	4,378 (2.75)	2,234 (2.29)	565 (0.99)	300 (1.60)	362 (1.82)	125 (2.09)
Normal	50,338 (31.66)	37,824 (38.83)	14,453 (25.39)	4,625 (24.67)	6,142 (30.83)	2,054 (34.34)
Overweight	39,589 (24.90)	28,470 (29.23)	21,447 (37.68)	6,984 (37.26)	6,493 (32.60)	1,803 (30.15)
Obese	25,169 (15.83)	17,355 (17.82)	17,922 (31.48)	6,317 (33.70)	6,255 (31.40)	1,707 (28.54)
Missing	39,546 (24.87)	11,528 (11.83)	2,536 (4.46)	518 (2.76)	668 (3.35)	292 (4.88)
Mean age (SD)	49.95 (16.54)	49.30 (12.39)	67.83 (12.05)	79.08 (9.39)	62.94 (11.51)	58.87 (13.72)
Mean BMI (SD)	26.29 (5.42)	26.24 (5.24)	28.37 (5.32)	28.46 (5.47)	28.08 (5.94)	27.63 (6.02)
Mean Multimorbidity (SD)	0.32 (0.57)	2.10 (1.04)	3.88 (1.34)	8.61 (2.03)	6.31 (1.66)	4.58 (1.73)

BMI- Body mass index; SD- Standard deviation; CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

Appendix Table 16 Summary statistics of different models across the age group in the total population (n=1.4 million)

Age group	Class	log-likelihood	BIC	SABIC	likelihood-ratio	AIC	Entropy	class 1	class 2	class 3	class 4	class 5	class 6	class 7	class 8	class 9
20-39	1	-1718985	-	3438911	213464.53	3438161		100.0								
	2	-1659095	3320688	3320081	93684.34	3318573	0.49	32.0	68.0							
	3	-1652903	3309559	3308647	81300.32	3306381	0.45	5.2	61.3	33.6						
	4	**-1649374**	**3303754**	**3302537**	**74240.8**	**3299513**	**0.46**	**34.1**	**0.6**	**4.3**	**61.0**					
	5	-1646748	3299759	3298237	68990.27	3294454	0.44	0.6	3.6	39.3	2.0	54.4				
	6	-1645008	3297534	3295707	65510.56	3291167	0.43	44.9	0.5	0.6	46.7	2.0	5.3			
	7	-1643412	3295597	3293465	62318.47	3288167	0.42	0.9	10.6	1.0	40.4	2.0	0.6	44.5		
	8	-1641847	3293721	3291283	59186.8	3285227	0.47	4.4	0.6	0.5	0.9	1.7	1.0	48.3	42.7	
	9	-1640820	3292923	3290180	57133.51	3283366	0.4	13.7	0.8	53.1	1.3	9.7	0.6	3.1	17.3	0.4
40-59	1	-3548501	-	7097968	682613.61	7097191		100.0								
	2	-3391442	6785436	6784829	368497.4	6783267	0.6	41.6	58.4							
	3	-3376753	6757339	6756427	339118.34	6754080	0.53	14.6	56.9	28.5						
	4	-3363678	6732472	6731255	312968.92	6728122	0.5	12.6	34.4	44.7	8.3					
	5	**-3354553**	**6715503**	**6713981**	**294718.15**	**6710064**	**0.52**	**3.3**	**10.6**	**35.8**	**6.0**	**44.4**				
	6	-3346478	6700636	6698809	278568.72	6694106	0.54	0.9	10.3	2.3	36.6	44.0	5.9			
	7	-3343006	6694974	6692842	271624.35	6687354	0.54	4.5	2.3	1.6	35.6	44.8	0.9	10.4		
	8	-3341037	6692318	6689880	267685.71	6683607	0.5	2.1	35.1	9.5	3.3	0.9	1.7	9.6	38.0	
	9	-3339240	6690007	6687264	264092.5	6680206	0.48	0.8	2.1	6.8	8.7	1.3	33.5	8.8	35.9	2.0
60-79	1	-4006131	-	8012733	1209659	8012357		100.0								
	2	-3810862	7622987	7622679	819121.97	7621918	0.73	58.0	42.0							
	3	-3785035	7571970	7571506	767467.46	7570362	0.63	28.7	18.5	52.8						
	4	-3766934	7536406	7535787	731265.79	7534258	0.59	12.8	32.1	23.9	31.3					
	5	**-3759021**	**7521219**	**7520444**	**715440.86**	**7518531**	**0.59**	**22.3**	**5.2**	**31.3**	**31.7**	**11.2**				
	6	-3753196	7510206	7509275	703789.75	7506978	0.61	30.6	22.5	0.9	11.2	3.6	31.3			
	7	-3747887	7500226	7499139	693171.51	7496458	0.59	30.4	9.7	6.9	28.3	0.8	20.9	3.1		
	8	-3744862	7494815	7493572	687122.47	7490507	0.58	8.3	3.1	8.7	29.9	4.6	24.2	20.4	0.8	
	9	-3742212	7490152	7488753	681821.4	7485303	0.58	17.1	24.3	4.6	0.8	4.4	9.3	28.8	2.8	7.9
>= 80	1	-1553617	-	3107648	769866.79	3107330		100.0								
	2	-1454818	2910782	2910474	572269.48	2909831	0.9	27.8	72.2							
	3	-1442426	2886575	2886111	547483.76	2885143	0.7	48.8	24.4	26.8						
	4	-1437531	2877365	2876746	537694.85	2875452	0.64	26.1	26.7	22.3	24.9					
	5	**-1434154**	**2871190**	**2870414**	**530940.31**	**2868796**	**0.62**	**13.5**	**24.1**	**24.4**	**15.6**	**22.4**				
	6	-1432432	2868325	2867394	527497.23	2865451	0.61	13.6	23.2	12.5	24.5	21.1	5.1			
	7	-1430977	2865993	2864906	524586	2862637	0.62	13.2	20.4	12.7	3.3	23.2	2.8	24.4		
	8	-1429598	2863814	2862571	521828.28	2859978	0.61	0.7	2.7	21.6	17.7	23.2	8.6	15.3	10.2	
	9	-1428522	2862242	2860843	519677.56	2857925	0.59	8.6	8.0	13.9	22.5	16.7	17.1	3.1	2.8	7.3

BIC- Bayesian information criteria; AIC- Akaike information criteria; SABIC- Sample size adjusted BIC

Appendix Table 17 Three-class cluster model in the age group 20-39 years

	Relative Healthy (78.43%)	Musculoskeletal (10.42%)	Mental Health (2.13%)
Anaemia	0.48	5.51	7.22
Ankylosing Spondylitis	0	0.85	0.21
Arterial/Venous	0	0.05	0.17
Asthma	1.13	8.67	7.68
Back pain	9.37	**58.40**	35.14
Benign Prostatic Hypertrophy^	0.04	0.34	0.06
Cataract	0.12	0.18	0.72
Chronic Heart Disease	0.02	0.13	0.77
Chronic Kidney Disease	0.06	0.75	3.08
COPD	0.04	0.51	0.46
Dementia	0.01	0.06	0.36
Depression	4.51	40.54	**49.17**
Diabetes	0.43	3.40	20.14
Epilepsy	0.15	0.86	1.59
Fatigue	0.10	2.77	2.86
Fibromyalgia	0	1.32	1.31
Gall stones	0.20	4.03	2.93
Gastritis	0.28	4.62	2.68
Gastrointestinal Bleed	0.08	1.58	0.82
Gout	0.17	0.83	1.59
Hearing Problem	0.79	5.33	4.40
Heart Failure	0	0.02	0.21
High Cholesterol	0.11	1.15	6.19
HIV/AIDS	0.01	0.02	0.04
Hypertension	0.25	2.47	8.17
Hyperthyroidism	0	0	13.28
Hypothyroidism	0.21	1.82	27.95
Inflammatory Bowel Disease	0.72	8.25	4.23
Irritable Bowel Syndrome	3.22	11.91	6.25
Liver Disease	0.1	0.53	0.91
Migraine	1.23	14.62	8.31
Multiple Sclerosis	0.05	0.42	0.54
Osteoarthritis	0.24	0.91	0.89
Osteoporosis	0.01	0.19	0.22
Parkinson Disease	0	0	0.02
Peripheral Vascular Disease	0.15	1.37	0.87
Polymyalgia	0	0.01	0.02
Psoriasis	0.61	3.15	2.17
Psychosis	0	0	9.10
Renal Stone	0.11	0.97	0.43
Rheumatoid Arthritis	0.03	0.47	0.61
Schizophrenia	0.02	0.30	15.21
Scleroderma	0	0.02	0.06
Sjogren's syndrome	0	0.05	0.22
Sleep Problem	0.10	2.85	2.95
Stroke	0.42	3.39	5.19
Systemic Lupus Erythematous	0	0.07	0.23
Tuberculosis	0.07	0.36	0.54
Vision Problem	0.02	0.13	0.89

^For men only

Appendix Table 18 Five-class cluster model in the age group 40-59 years

	Healthy (47.48%)	MSK (31.86%)	CV-MSK (11.56%)	MSK-MH (5.94%)	Metabolic (3.16%)
Anaemia	0.62	5.48	3.69	14.74	10.57
Ankylosing Spondylitis	0	1.38	0.52	4.99	0.32
Arterial/Venous	0.01	0.08	0.44	0.53	0.08
Asthma	2.07	12.17	10.99	23.12	12.04
Back pain	11.09	**65.4**	44.68	**88.64**	49.04
Benign Prostatic Hypertrophy^	0.18	1.19	2.33	2.01	0.35
Cancer	0.55	3.05	3.75	4.84	4.89
Cataract	0.52	0.41	2.03	1.95	1.13
Chronic Heart Disease	0.13	0.6	6.94	5.34	1.52
Chronic Kidney Disease	0.07	0.6	7.35	5.88	4.26
COPD	0.18	2.54	3.01	8.56	3.09
Dementia	0	0.06	0.08	0.34	1.12
Depression	3.31	37.91	23.11	**73.82**	49.08
Diabetes	0.8	2.68	24.11	18.42	20.68
Epilepsy	0.18	1.02	1.12	2.25	2.22
Fatigue	0.05	2.09	0.63	12.59	3.4
Fibromyalgia	0	0.86	0.01	13.64	1.34
Gall stones	0.4	4.25	4.09	16.31	5.35
Gastritis	0.3	5.33	4.95	19.74	3.66
Gastrointestinal Bleed	0.03	1.14	1.14	4.38	1.03
Gout	0.52	1.27	9.08	2.88	0.7
Hearing Problem	1.14	8.49	8.33	15.56	7.51
Heart Failure	0	0	0.89	0.44	0.3
High Cholesterol	1.13	5.33	25.74	20.2	11.58
HIV/AIDS	0	0.03	0.06	0.02	0.05
Hypertension	1.66	5.15	**57.61**	27.31	13.85
Hyperthyroidism	0.05	0	0	2.02	23.78
Hypothyroidism	0.47	2.76	3.36	13.42	**56.7**
Inflammatory Bowel Disease	0.52	7.13	4.62	19.68	3.9
Irritable Bowel Syndrome	3.89	10.17	2.92	33.84	7.75
Liver Disease	0.1	0.65	1.21	2.01	1.32
Migraine	0.91	11.98	5.27	30.98	9.77
Multiple Sclerosis	0.09	0.67	0.31	1.08	0.86
Osteoarthritis	3.34	4.82	8.86	**21.46**	5.34
Osteoporosis	0.05	0.8	0.69	3.27	1.23
Parkinson Disease	0.01	0.05	0.07	0.12	0.14
Peripheral Vascular Disease	0.17	1.61	1.66	4.93	1.7
Polymyalgia	0	0.07	0.1	0.52	0.07
Psoriasis	0.81	3.82	4.35	4.88	3.24
Psychosis	0	0	0	0.3	10.09
Renal Stone	0.19	1.27	2.58	2.37	0.94
Rheumatoid Arthritis	0.1	0.7	0.7	3.11	0.93
Schizophrenia	0.03	0.31	0.12	2.11	18.44
Scleroderma	0	0.02	0.04	0.1	0.07
Sjogren's syndrome	0.01	0.07	0.03	0.59	0.27
Sleep Problem	0.09	2.75	3	10.15	3.89
Stroke	0.64	5.36	8.46	8.44	5.85
Systemic Lupus Erythematous	0.01	0.05	0.04	0.45	0.12
Tuberculosis	0.08	0.32	0.52	0.56	0.38
Vision Problem	0.02	0.09	0.46	0.71	1.25

CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal; ^only for men

Appendix Table 19 Five-class cluster model in the age group 60-79 years

	Healthy (28.89%)	MSK (29.70%)	CV-MSK (9.56%)	CV (2.09%)	MSK-MH (10.96%)
Anaemia	0.19	3.66	13.94	3.32	12.22
Ankylosing Spondylitis	0.00	1.61	3.36	0.00	4.70
Arterial/Venous	0.02	0.45	4.92	1.15	0.42
Asthma	0.91	10.09	14.80	9.25	20.57
Back pain	4.13	**65.02**	**83.77**	45.55	**84.74**
Benign Prostatic Hypertrophy^	0.25	7.17	16.23	6.49	2.16
Cancer	0.87	10.26	16.71	10.91	12.29
Cataract	5.81	4.00	20.52	8.29	9.75
Chronic Heart Disease	0.38	4.95	35.54	12.60	7.19
Chronic Kidney Disease	0.00	1.35	32.51	17.86	13.98
COPD	0.28	7.71	19.90	5.47	11.47
Dementia	0.04	0.88	4.01	0.88	1.85
Depression	1.28	25.80	42.00	13.53	**61.51**
Diabetes	0.56	5.02	43.49	28.64	18.82
Epilepsy	0.09	1.15	2.65	0.92	1.91
Fatigue	0.01	1.33	3.31	0.53	9.10
Fibromyalgia	0.00	0.31	1.25	0.00	8.52
Gall stones	0.33	5.70	14.61	5.72	18.99
Gastritis	0.13	7.72	24.32	3.58	17.95
Gastrointestinal Bleed	0.00	1.20	5.46	0.54	2.63
Gout	0.30	2.75	17.85	10.99	2.54
Hearing Problem	1.01	17.77	31.39	13.71	21.54
Heart Failure	0.01	0.13	6.90	1.58	0.34
High Cholesterol	1.16	17.22	46.82	33.54	32.04
HIV/AIDS	0.00	0.01	0.01	0.01	0.02
Hypertension	2.38	20.83	**72.27**	**68.89**	42.73
Hyperthyroidism	0.04	0.14	1.17	2.12	8.39
Hypothyroidism	0.46	4.53	9.70	9.96	28.01
Inflammatory Bowel Disease	0.15	5.64	12.82	2.48	13.75
Irritable Bowel Syndrome	5.11	7.51	11.17	1.68	27.38
Liver Disease	0.03	0.66	2.47	0.81	1.54
Migraine	0.27	7.93	9.47	3.05	22.23
Multiple Sclerosis	0.05	0.70	0.35	0.21	0.83
Osteoarthritis	18.39	30.25	**57.54**	19.66	**55.93**
Osteoporosis	0.31	6.29	8.45	2.61	14.79
Parkinson Disease	0.02	0.67	1.40	0.48	0.85
Peripheral Vascular Disease	0.06	2.47	12.83	2.94	4.59
Polymyalgia	0.05	0.74	2.40	1.07	2.61
Psoriasis	0.37	3.90	6.53	3.91	4.91
Psychosis	0.00	0.15	0.21	0.00	2.06
Renal Stone	0.09	1.76	5.13	2.23	1.81
Rheumatoid Arthritis	0.05	0.96	2.48	0.79	3.74
Schizophrenia	0.03	0.66	0.89	0.25	4.34
Scleroderma	0.00	0.05	0.14	0.03	0.24
Sjogren's syndrome	0.01	0.15	0.34	0.07	1.12
Sleep Problem	0.02	2.73	6.97	1.70	7.96
Stroke	1.36	7.02	21.15	11.17	10.51
Systemic Lupus Erythematous	0.00	0.04	0.13	0.03	0.33
Tuberculosis	0.04	0.34	0.69	0.28	0.46
Vision Problem	0.03	0.32	1.89	0.36	0.85

CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal; ^only for men

Appendix Table 20 Five-class cluster model in the age group >=80 years

	Healthy (24.83%)	MSK (21.12%)	CV (24.64%)	MSK-CV-Renal (13.71%)	MSK- CV-MH (15.69%)
Anaemia	0.10	5.48	7.29	19.37	21.02
Ankylosing Spondylitis	0.00	1.58	0.00	2.31	4.69
Arterial/Venous	0.02	1.38	1.46	8.91	2.03
Asthma	0.34	10.09	8.52	11.91	17.59
Back pain	1.04	**66.14**	41.86	**78.03**	**91.21**
Benign Prostatic Hypertrophy^	0.19	16.28	5.43	32.06	2.21
Cancer	0.43	18.75	14.92	23.45	17.01
Cataract	24.12	26.25	28.31	43.71	**55.95**
Chronic Heart Disease	0.21	11.43	16.45	49.96	24.12
Chronic Kidney Disease	0.00	7.31	44.48	**67.88**	47.49
COPD	0.26	11.37	6.74	19.24	15.67
Dementia	0.52	10.79	11.03	9.41	15.74
Depression	0.27	18.00	12.12	21.35	**57.76**
Diabetes	0.25	7.88	26.99	**52.40**	26.06
Epilepsy	0.05	1.64	1.18	1.93	1.65
Fatigue	0.00	1.84	0.63	3.28	6.00
Fibromyalgia	0.00	0.22	0.00	0.21	1.85
Gall stones	0.20	8.73	7.52	16.32	21.59
Gastritis	0.11	11.13	4.85	23.84	21.73
Gastrointestinal Bleed	0.01	2.28	0.97	6.35	4.56
Gout	0.05	3.33	8.78	25.45	7.92
Hearing Problem	0.68	38.26	25.83	47.38	43.17
Heart Failure	0.03	1.29	3.74	15.55	5.86
High Cholesterol	0.16	16.20	29.45	45.07	40.70
HIV/AIDS	0.00	0.00	0.00	0.00	0.01
Hypertension	1.86	38.27	**80.48**	**77.93**	**74.64**
Hyperthyroidism	0.00	0.56	2.88	0.76	7.10
Hypothyroidism	0.19	6.95	15.46	9.96	28.75
Inflammatory Bowel Disease	0.06	5.80	2.19	10.00	15.45
Irritable Bowel Syndrome	3.70	6.43	1.22	6.03	18.00
Liver Disease	0.01	0.41	0.25	1.06	1.11
Migraine	0.04	5.07	2.12	4.77	12.34
Multiple Sclerosis	0.01	0.27	0.16	0.12	0.24
Osteoarthritis	30.15	39.39	29.51	**62.15**	**68.45**
Osteoporosis	0.29	17.61	11.53	7.56	37.88
Parkinson Disease	0.05	2.33	1.09	1.78	1.67
Peripheral Vascular Disease	0.05	3.33	3.64	15.40	8.79
Polymyalgia	0.06	4.09	3.18	4.14	8.88
Psoriasis	0.09	3.43	3.11	5.39	5.71
Psychosis	0.00	0.22	0.31	0.01	0.77
Renal Stone	0.05	2.01	1.36	5.37	1.73
Rheumatoid Arthritis	0.02	1.19	0.70	1.28	2.90
Schizophrenia	0.02	0.84	0.81	0.09	1.96
Scleroderma	0.00	0.09	0.05	0.03	0.15
Sjogren's syndrome	0.00	0.22	0.10	0.13	1.20
Sleep Problem	0.05	4.55	2.17	6.10	9.36
Stroke	5.99	13.70	19.18	29.21	23.51
Systemic Lupus Erythematous	0.00	0.08	0.04	0.07	0.09
Tuberculosis	0.02	1.01	0.44	0.99	1.42
Vision Problem	0.04	1.68	1.47	3.07	3.85

CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal; ^only for men

Appendix Table 21 Summary statistics of different models across gender in the OA population (n=221k)

	Class	log-likelihood	BIC	SABIC	likelihood-ratio	AIC	Entropy	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
Men	1	-707711	-	1415835	213962.6	1415521		100									
	2	-683958	1369073	1368752	166457.6	1368118	0.59	35.42	64.58								
	3	-680328	1362397	1361914	159198.1	1360961	0.54	25.62	58.62	15.76							
	4	-678182	1358688	1358043	154904.8	1356770	0.5	24.77	8.48	19.74	47.01						
	5	**-677217**	**1357342**	**1356535**	**152975.1**	**1354942**	**0.5**	**2.09**	**23.85**	**18.54**	**47.39**	**8.13**					
	6	-676495	1356482	1355512	151530.5	1353599	0.52	8.22	0.72	23.91	0.64	19.35	47.18				
	7	-675869	1355813	1354682	150278.3	1352449	0.51	20.81	0.72	48.07	6.93	6.47	0.65	16.34			
	8	-675337	1355334	1354041	149215.1	1351488	0.53	16.33	49.09	6.5	0.63	5.14	4.22	0.74	17.36		
	9	-674825	1354894	1353438	148190.8	1350566	0.52	5.46	8.45	0.64	4.09	16.61	0.62	10.38	4.93	48.82	
	10	-674425	1354677	1353060	147390.3	1349867	0.49	5.44	4.12	4.21	15.31	17.24	0.61	8.33	40	4.09	0.66
Women	1	-1046505	-	2093439	313363.8	2093110		100									
	2	-1016959	2035106	2034785	254272.5	2034120	0.56	61.22	38.78								
	3	-1007064	2015915	2015432	234481.4	2014431	0.57	21.96	21.62	56.42							
	4	-1004124	2010636	2009991	228602.7	2008655	0.53	22.4	7.55	25.05	45						
	5	-1002057	2007100	2006293	224467.4	2004621	0.55	24.01	20.63	2.68	45.96	6.74					
	6	-1000467	2004520	2003550	221287.1	2001543	0.56	45.84	20.65	0.78	6.62	23.51	2.59				
	7	-999195	2002577	2001446	218744.8	1999103	0.55	9.87	20.15	45.84	5.93	15.14	0.76	2.3			
	8	-998348	2001482	2000189	217050.2	1997510	0.53	15.1	5.37	43.62	10.01	4.11	2.32	0.76	18.71		
	9	-997683	2000752	1999296	215720.2	1996282	0.5	0.74	5.42	30.34	2.18	8.57	23.09	15.8	9.19	4.66	
	10	-996996	1999978	1998360	214346.1	1995010	0.54	4.66	2.16	5.12	31.18	3.41	5.41	15.58	22.76	8.98	0.74

BIC- Bayesian information criteria; AIC- Akaike information criteria; SABIC- Sample size adjusted BIC

Appendix Table 22 Summary statistics of different models across the age group in the OA population (n=221K)

	Model	log-likelihood	BIC	SABIC	likelihood-ratio	AIC	Entropy	C1	C2	C3	C4	C5	C6	C7	C8	C9
20-39	1	-84301.8	-	168896.7	30831.18	168697.6		100.00								
	2	-81300.6	163495.5	163193.6	24828.76	162791.1	0.57	43.04	56.96							
	3	**-80795.7**	**162937.6**	**162483.1**	**23818.94**	**161877.3**	**0.56**	**52.81**	**38.80**	**8.39**						
	4	-80474.7	162747.5	162140.5	23176.95	161331.3	0.53	49.67	32.85	10.94	6.52					
	5	-80296.8	162843.5	162084	22821.14	161071.5	0.54	53.18	23.92	15.02	5.56	2.30				
	6	-80122.3	162946.4	162034.4	22472.15	160818.5	0.55	27.61	0.97	1.97	5.80	53.23	10.41			
	7	-80002.6	163158.9	162094.3	22232.75	160675.1	0.53	30.05	5.74	1.91	9.87	8.03	0.96	43.44		
	8	-79888	163381.6	162164.5	22003.59	160542	0.53	46.98	19.73	1.07	3.46	8.85	12.79	1.99	5.13	
	9	-79777.6	163612.8	162243.2	21782.9	160417.3	0.52	8.46	0.39	5.93	1.12	23.75	2.02	4.98	41.31	12.04
40-59	1	-794856	-	1590100	260018.3	1589806		100.00								
	2	-768549	1538182	1537880	207403.3	1537287	0.61	56.83	43.17							
	3	-762821	1527274	1526820	195947.4	1525927	0.56	38.97	37.19	23.84						
	4	-759436	1521053	1520446	189177.4	1519253	0.55	33.81	28.66	26.32	11.20					
	5	**-757514**	**1517756**	**1516997**	**185333.2**	**1515505**	**0.56**	**31.17**	**27.60**	**26.29**	**9.79**	**5.14**				
	6	-755975	1515227	1514314	182255.1	1512523	0.58	34.19	22.58	20.26	10.77	7.40	4.76			
	7	-754631	1513088	1512023	179568.5	1509932	0.56	9.90	35.29	3.79	1.01	22.08	20.60	7.32		
	8	-753657	1511688	1510471	177620.7	1508081	0.55	3.57	5.22	1.00	29.10	10.34	8.94	19.57	22.26	
	9	-753089	1511100	1509731	176484.5	1507040	0.55	4.30	21.66	2.37	20.03	3.70	28.79	1.01	7.39	10.75
60-79																
	1	-1112361	-	2225115	466617.9	2224817		100.00								
	2	-1075437	2151969	2151667	392769.3	2151064	0.7	69.49	30.51							
	3	-1066509	2134666	2134211	374912.8	2133304	0.63	26.38	15.78	57.85						
	4	-1062240	2126681	2126074	366374.6	2124862	0.57	33.89	29.35	19.66	17.09					
	5	**-1059898**	**2122551**	**2121791**	**361691**	**2120274**	**0.56**	**30.16**	**27.09**	**15.72**	**14.35**	**12.66**				
	6	-1058402	2120113	2119201	358699.7	2117379	0.57	29.70	26.05	14.92	13.77	11.61	3.92			
	7	-1057169	2118199	2117135	356233.2	2115008	0.58	29.41	19.45	13.84	13.28	12.43	6.26	5.30		
	8	-1056061	2116537	2115320	354017.8	2112889	0.57	14.12	3.76	9.32	11.17	8.26	1.06	26.05	26.27	
	9	-1055061	2115091	2113721	352017.9	2110985	0.59	23.02	7.47	13.77	2.38	13.99	5.05	7.28	1.04	26.00
>=80																
	1	-206667	-	413636.6	128573.5	413426.4		100.00								
	2	-199829	400565.5	400269.9	114898	399844.9	0.85	18.33	81.67							
	3	-198582	398529.5	398084.6	112403.9	397444.8	0.61	13.97	55.95	30.08						
	4	-197873	397569.3	396975	110985.5	396120.4	0.57	23.82	14.55	29.82	31.81					
	5	**-197329**	**396938.5**	**396194.9**	**109896.5**	**395125.4**	**0.56**	**28.44**	**25.12**	**17.45**	**14.96**	**14.00**				
	6	-197082	396903.9	396010.9	109403.8	394726.7	0.57	22.90	20.17	16.07	14.62	13.38	12.84			
	7	-196842	396881.3	395838.9	108923	394339.9	0.57	22.71	20.03	16.15	13.83	12.96	8.77	5.51		
	8	-196632	396918.9	395727.2	108502.5	394013.4	0.56	22.57	13.04	3.36	7.38	15.27	13.71	3.17	21.51	
	9	-196449	397012	395670.9	108137.4	393742.3	0.56	21.01	3.20	14.48	21.81	13.05	3.58	8.37	6.77	7.73

BIC- Bayesian information criteria; AIC- Akaike information criteria; SABIC- Sample size adjusted BIC

Appendix Table 23 Summary statistics of different models across the gender in the non-OA population (n=221K)

	Model	log-likelihood	BIC	SABIC	likelihood-ratio	AIC	Entropy	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
Men	1	-586257	-	1172927	160617.3	1172613		100									
	2	-564671	1130498	1130177	117445.9	1129544	0.59	31.95	68.05								
	3	-561526	1124792	1124309	111155.7	1123356	0.54	22.77	63.18	14.04							
	4	-559814	1121953	1121308	107732.4	1120034	0.49	53.04	22.01	6.95	17.99						
	5	-558735	1120378	1119570	105573.3	1117977	0.5	1.77	21.33	17.15	53.14	6.62					
	6	-558035	1119562	1118593	104173.7	1116680	0.51	0.61	17.71	21.35	0.76	6.54	53.03				
	7	-557550	1119176	1118044	103203.4	1115812	0.49	19.64	0.59	53.22	7.85	5.99	0.7	12.02			
	8	-557029	1118718	1117424	102161.4	1114872	0.51	12.72	54.79	7.41	0.71	4.95	4.24	0.58	14.6		
	9	-556645	1118533	1117078	101393	1114205	0.49	4.62	7.57	0.59	3.52	14.83	0.72	14.88	3.76	49.51	
	10	-556330	1118488	1116870	100763.9	1113678	0.49	0.72	49.94	3.41	3.27	2.16	0.59	4.52	14.55	6.92	13.94
Women	1	-850709	-	1701847	224654.9	1701518		100									
	2	-824269	1649726	1649405	171775.6	1648740	0.55	67	33								
	3	-816692	1635171	1634688	156620.9	1633688	0.56	62.8	17.79	19.42							
	4	-814589	1631566	1630921	152416.2	1629585	0.51	5.72	20.64	20.23	53.41						
	5	-812760	1628506	1627699	148756.4	1626027	0.52	4.1	18	18.72	54.09	5.09					
	6	-811217	1626021	1625052	145671.5	1623044	0.54	19.02	0.69	18.24	4.98	54.45	2.62				
	7	-810405	1624997	1623865	144047.7	1621522	0.54	5.09	17.74	2.77	0.72	19.63	1.28	52.77			
	8	-809761	1624307	1623014	142758.7	1620335	0.52	0.71	4.16	0.86	3.55	21.75	48.03	18	2.94		
	9	-809125	1623637	1622181	141488.2	1619167	0.5	2.32	47.41	3.77	18.83	4.62	6.25	0.71	0.9	15.2	
	10	-808547	1623078	1621461	140330.3	1618111	0.49	2.98	5.51	45.68	18.85	15.54	2.6	0.48	3.36	0.7	4.3

BIC- Bayesian information criteria; AIC- Akaike information criteria; SABIC- Sample size adjusted BIC

Appendix Table 24 Summary statistics of different models across the age groups in the non-OA population (n=221K)

	Model	log-likelihood	BIC	SABIC	likelihood-ratio	AIC	Entropy	C1	C2	C3	C4	C5	C6	C7	C8	C9
20-39																
	1	-71006.1	-	142308.2	19871.23	142106.3										
	2	-68852.6	138605.2	138303.3	15564.17	137895.2	0.51	60.9	39.1							
	3	**-68462.9**	**138280.6**	**137826.1**	**14784.77**	**137211.8**	**0.49**	**59.6**	**29.7**	**10.6**						
	4	-68213.2	138236	137629	14285.45	136808.5	0.51	59.4	28.3	11.3	0.97					
	5	-67960.8	138185.9	137426.4	13780.6	136399.6	0.55	58.4	30.0	7.30	3.24	0.95				
	6	-67798.6	138316.2	137404.2	13456.15	136171.2	0.5	1.0	40.6	6.3	2.5	7.1	42.5			
	7	-67702.2	138578.2	137513.6	13263.33	136074.4	0.52	1.0	6.0	2.2	1.0	48.1	34.8	6.9		
	8	-67617.7	138864.1	137646.9	13094.47	136001.5	0.51	7.2	39.3	2.1	0.8	7.4	37.2	5.0	1.0	
	9	-67554.6	139192.6	137822.9	12968.25	135971.3	0.53	18.7	0.8	0.5	42.6	9.9	2.0	4.4	14.5	6.6
40-59																
	1	-679652	-	1359692	184619	1359398										
	2	-659274	1319633	1319331	143862.3	1318738	0.56	49.5	50.5							
	3	-654441	1310515	1310061	134196.4	1309168	0.54	21.9	47.9	30.2						
	4	-652322	1306826	1306219	129959.6	1305027	0.51	34.5	34.2	22.9	8.40					
	5	**-650611**	**1303951**	**1303191**	**126535.9**	**1301699**	**0.54**	**36.0**	**31.6**	**22.2**	**6.98**	**3.18**				
	6	-649160	1301598	1300686	123634.4	1298894	0.55	36.2	32.9	31.0	22.8	5.79	0.93			
	7	-648293	1300412	1299348	121900.7	1297256	0.55	33.4	0.9	2.2	3.4	6.0	34.1	19.9		
	8	-647768	1299911	1298694	120851.2	1296302	0.52	8.6	28.2	1.9	18.5	33.4	3.2	0.9	5.2	
	9	-647269	1299461	1298092	119853.4	1295401	0.52	3.9	1.3	27.1	0.9	3.5	3.1	9.8	20.0	30.4
60-79																
	1	-994852	-	1990095	368056.2	1989797										
	2	-962423	1925940	1925638	303199.1	1925036	0.66	63.4	36.6							
	3	-956037	1913721	1913266	290427	1912360	0.60	56.5	19.3	24.2						
	4	-951818	1905835	1905228	281988.2	1904017	0.56	30.9	30.7	21.6	16.8					
	5	**-949594**	**1901940**	**1901180**	**277540.1**	**1899665**	**0.55**	**29.8**	**26.5**	**19.7**	**14.1**	**9.93**				
	6	-948146	1899596	1898684	274644.1	1896865	0.58	30.2	24.9	18.1	12.9	8.5	5.3			
	7	-946885	1897628	1896563	272122.5	1894440	0.57	29.4	19.4	13.8	13.3	12.4	6.3	5.3		
	8	-946059	1896529	1895311	270470.9	1892884	0.54	9.4	13.9	23.7	8.5	1.0	19.1	3.0	21.3	
	9	-945487	1895938	1894568	269327.1	1891836	0.54	3.5	8.1	2.9	21.2	22.6	0.9	13.6	19.0	8.3
>=80																
	1	-194509	-	389326.3	111852.9	389111.3										
	2	-188296	377518.8	377216.9	99428.05	376782.5	0.78	22.8	77.2							
	3	-187228	375851.3	375396.8	97292.49	374742.9	0.6	54.7	15.5	29.8						
	4	-186429	374721.1	374114.1	95694.28	373240.7	0.57	32.1	28.1	23.3	16.5					
	5	-186044	374419.1	373659.5	94924.19	372566.6	0.54	28.0	22.0	19.2	15.6	15.1				
	6	**-185749**	**374296.8**	**373384.7**	**94333.87**	**372072.3**	**0.56**	**27.3**	**19.9**	**17.8**	**15.1**	**14.3**	**5.5**			
	7	-185477	374221.3	373156.7	93790.41	371624.8	0.57	27.8	20.8	16.9	15.0	13.4	5.1	0.96		
	8	-185320	374374.1	373157	93475.17	371405.6	0.56	4.1	13.1	26.6	12.6	0.8	15.9	12.0	14.9	
	9	-185173	374548.9	373179.2	93181.91	371208.3	0.55	9.0	12.3	21.3	0.8	16.7	3.6	11.2	15.0	10.0

BIC- Bayesian information criteria; AIC- Akaike information criteria; SABIC- Sample size adjusted BIC

Appendix Table 25 Statistical parameters for each class using LTA

Class	AIC	BIC	Log-likelihood	Number of parameters	Entropy
Osteoarthritis					
Class 1	9087641	8956381	-2613941	36	100
Class 2	8585302	8403212	-2803813	102	0.66
Class 3	8083211	7983918	-3092181	167	0.65
Class 4	7478432	7410985	-3287312	233	0.64
Class 5	**6877108**	**6880180**	**-3438256**	**298**	**0.69**
Class 6	6889342	6890183	-3421843	365	0.64
Class 7	7456932	7450932	-3302814	431	0.64
Class 8	7985400	7984513	-3039471	497	0.61
Class 9	8487212	8510393	-2735715	562	0.59
Class 10	8954901	8972381	-2345750	628	0.59
Non-Osteoarthritis					
Class 1	7011091	7012094	-3421392	36	100
Class 2	6796783	6795883	-3382123	102	0.61
Class 3	6703452	6707631	-3348726	167	0.63
Class 4	6632157	6634559	-3315846	233	0.63
Class 5	**6561408**	**6564480**	**-3280406**	**298**	**0.67**
Class 6	6505843	6509606	-3252557	365	0.65
Class 7	6561601	6564502	-3281201	431	0.61
Class 8	6622315	6645408	-3334653	497	0.57
Class 9	6762432	6695720	-3340032	562	0.55
Class 10	6973465	7063701	-3350016	628	0.55

BIC- Bayesian information criteria; AIC- Akaike information criteria

Appendix Table 26 Conditional probabilities in OA using LTA

	Healthy (34.85%)	CV-MSK (9.78%)	CV (26.36%)	MSK (15.38%)	MH (13.18%)
Anaemia	1.07	13.23	3.33	8.61	3.56
Ankylosing Spondylitis	0.11	3.38	0.13	8.66	0.07
Asthma	5.50	15.80	10.63	20.84	14.24
Back pain	27.16	**64.42**	25.59	**77.73**	33.05
Benign Prostatic Hypertrophy^	1.01	12.66	7.79	3.46	2.51
Cataract	3.91	25.83	8.87	2.85	3.13
Chronic Heart Disease	0.01	41.26	15.59	2.70	5.01
Hypercholesterolemia	0.99	42.66	23.50	13.76	11.56
Arterial and venous	0.00	4.36	0.92	0.21	0.18
Chronic Kidney Disease	0.01	28.96	9.63	2.31	3.17
COPD	1.56	19.35	6.90	11.98	5.84
Gout	0.89	13.78	8.51	1.64	3.02
Dementia	0.13	4.04	1.09	0.29	0.95
Depression	0.01	46.76	0.01	**43.86**	**95.70**
Diabetes	0.27	32.48	15.97	4.33	7.53
Epilepsy	0.78	2.40	1.14	2.41	2.06
Fatigue	0.17	3.01	0.46	4.93	1.11
Fibromyalgia	0.06	1.70	0.07	6.27	1.18
Gastritis	1.27	21.00	5.17	17.46	4.02
Gall stones	1.16	16.17	5.64	13.50	4.30
Gastrointestinal bleed	0.11	5.06	0.93	3.20	0.79
Hearing Problem	3.66	28.67	14.35	14.73	9.81
Heart Failure	0.00	9.68	1.89	0.03	0.28
HIV/AIDS	0.01	0.02	0.01	0.00	0.02
Hypertension	0.04	**72.75**	**60.76**	18.02	25.95
Hyperthyroid	0.00	4.25	1.54	2.96	1.33
Hypothyroid	1.20	17.70	7.40	11.50	7.27
Inflammatory Bowel Disease	1.36	12.00	2.88	15.65	3.29
Irritable Bowel Syndrome	4.91	11.80	2.10	28.39	6.15
Liver Disease	0.12	1.37	0.56	0.96	0.86
Migraine	2.24	9.09	3.10	26.87	6.98
Multiple sclerosis	0.12	0.29	0.18	0.54	0.44
Osteoporosis	0.74	12.18	3.36	7.02	2.30
Parkinson Disease	0.10	1.30	0.55	0.23	0.58
Polymyalgia	0.03	4.38	1.85	1.19	0.59
Psoriasis	1.67	5.14	3.69	4.61	3.84
Psychosis	0.00	0.52	0.00	0.00	1.64
Peripheral Vascular Disease	0.43	12.31	3.13	3.43	1.75
Rheumatoid Arthritis	0.58	1.98	1.13	3.22	1.37
Renal Stone	0.38	3.26	1.70	1.64	0.81
Schizophrenia	0.00	1.65	0.00	0.25	4.74
Scleroderma	0.00	0.13	0.05	0.08	0.02
Sjogren's syndrome	0.04	0.62	0.09	0.55	0.08
Systemic Lupus Erythematous	0.02	0.21	0.03	0.22	0.04
Sleep Problem	0.36	8.15	2.20	5.01	3.73
Stroke	7.61	22.48	9.54	4.03	6.20
Tuberculosis	0.21	1.59	0.73	1.16	0.46
Vision problem	0.01	3.55	0.90	0.48	0.41

CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal; ^only for men

Appendix Table 27 Conditional probabilities in non-OA using LTA

	Healthy (39.45%)	CV-MSK (9.52%)	Metabolic (7.54%)	CV (19.41%)	MH (24.07%)
Anaemia	0.87	10.12	11.1	0.68	4.71
Ankylosing spondylitis	0	3.62	0.31	0.01	3.37
Arterial/Venous	0	3.88	0.37	0.78	0.07
Asthma	4.65	11.65	12.3	7.77	13.66
Back pain	16.03	**52.7**	21.97	16.2	40.91
Benign Prostatic Hypertrophy	0.91	13.62	0.79	6.72	2.35
Cancer (any)	1.83	12.46	10.54	6.42	4.23
Cataract	3.9	24.06	15.02	7.27	1.41
Chronic Kidney Disease	0.01	26.84	11.91	10.37	1.08
COPD	1.55	17.86	4.42	5.02	7.27
Coronary Heart Disease	0.01	36.99	6.41	15.47	1.46
Dementia	0.13	3.68	2.41	0.66	0.41
Depression	0.07	34.96	15.16	3.24	**54.24**
Diabetes	0.13	26.45	10.64	16.41	2.6
Epilepsy	0.53	2.22	1.71	0.75	1.81
Fatigue	0.04	2.47	1.49	0.15	2.71
Fibromyalgia	0	1	0.24	0	1.97
Gall stones	0.96	13.15	6.46	3.51	6.39
Gastritis	0.24	19.53	2.42	2.73	10.38
Gastrointestinal bleed	0.01	4.87	0.44	0.32	1.78
Gout	0.17	9.9	1.55	7.18	0.9
Hearing impairment	3.75	28.89	15.77	11.28	9.88
Heart Failure	0	6.81	0.95	1.23	0.01
HIV/AIDS	0	0	0	0.01	0.01
Hypocholesteraemia	0.51	39.41	15.81	25.53	8.04
Hypertension	0	**67.72**	41.79	**66.22**	12.57
Hyperthyroidism	0	1.47	13.78	0	0.17
Hypothyroidism	0.01	9.76	**49.55**	0	3.27
Inflammatory Bowel Disease	0.66	10.4	3.13	1.88	9.25
Irritable Bowel Syndrome	2.38	11.09	3.73	0.44	15.85
Liver Disease	0.1	0.97	0.53	0.3	0.78
Migraine	0.25	8.65	5.35	1.69	17.23
Multiple Sclerosis	0.12	0.3	0.86	0	0.7
Osteoporosis	0.53	9.79	15.66	0.26	3.07
Parkinson's Disease	0.08	1.09	0.67	0.26	0.3
Peripheral Vascular Disease	0.29	10.55	2.12	2.4	1.77
Polymyalgia	0.07	2.93	4.66	0.19	0.24
Psoriasis	1.31	4.46	2.9	3.08	3.74
Psychosis	0	0.58	0.01	0	1.01
Renal Stone	0.28	3.16	0.33	1.82	1.13
Rheumatoid Arthritis	0.14	1.02	2.55	0.09	0.67
Schizophrenia	0	1.75	0.25	0	2.92
Scleroderma	0.01	0.12	0.24	0	0.02
Sjogren’s disease	0	0.36	0.54	0	0.15
Sleep problem	0.32	7.69	3.74	1.22	3.28
Stroke	6.79	20.97	8.45	8.83	4.5
Systemic Lupus Erythematous	0.01	0.14	0.18	0	0.09
Tuberculosis	0.25	1.55	2.08	0.13	0.81
Vision problem	0.02	3.33	3.58	0	0.11

CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal; ^only for men

Appendix Table 28 Transition probabilities across clusters in OA

	CV-MSK	CV	MSK	Relative Healthy	MH
Index date	Transition at year 5				
CV-MSK	100.0	0.0	0.0	0.0	0.0
CV	11.6	88.4	0.0	0.0	0.0
MSK	5.2	0.0	94.8	0.0	0.0
Relative Healthy	0.0	12.6	1.4	81.7	4.3
MH	3.3	0.0	0.0	0.0	96.7
Year 5	Transition at year 10				
CV-MSK	100.0	0.0	0.0	0.0	0.0
CV	15.5	84.1	0.0	0.0	0.4
MSK	11.7	0.0	88.2	0.0	0.0
Relative Healthy	0.0	14.1	2.5	80.7	2.7
MH	7.9	0.1	0.4	0.0	91.7
Year 10	Transition at year 15				
CV-MSK	82.0	9.3	2.4	3.2	3.1
CV	22.5	70.5	2.8	2.4	1.8
MSK	6.0	2.0	85.1	2.6	4.4
Relative Healthy	0.5	14.5	7.0	72.2	5.9
MH	22.5	3.4	2.7	1.9	69.5

CV-Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

Shaded cell is the cases remaining in the same LTA class on successive years.

The transition is to be read from row to column. For example, at year 5; 100% cases moved to cluster 1 and 11.6% cases from cluster 2 (at index date) moved to cluster 1 at year 5.

Appendix Table 29 Transition probabilities across clusters in non-OA

	Metabolic	Cardiovascular	Relative healthy	CV-MSK	Mental Health
Index date	Transition at year 5				
Metabolic	99.8	0	0	0.2	0
Cardiovascular	0.0	91.3	0	8.7	0
Relative Healthy	3.1	10.4	85.8	0	0.8
CV-MSK	0	0	0	100	0
Mental Health	0	0	0	6.0	93.9
Year 5	Transition at year 10				
Metabolic	98.5	0.0	0	1.5	0
Cardiovascular	1.4	83.9	0	14.7	0
Relative Healthy	3.8	11.8	81.1	0.0	3.3
CV-MSK	0.0	0	0	100.0	0
Mental Health	0.0	0	0	12.3	87.7
Year 10	Transition at year 15				
Metabolic	87.6	0.2	1.7	8.8	1.6
Cardiovascular	7.3	72.2	0.4	19.1	1.0
Relative Healthy	3.9	9.5	80.6	0.7	5.4
CV-MSK	0.3	2.8	2.1	93.3	1.5
Mental Health	2.5	1.2	1.0	11.7	83.6

CV-Cardiovascular; MSK- Musculoskeletal

Appendix Table 30 Transition pattern with more than 1% of total in OA

	At index	Year 0	Year 5	Year 10	Year 15	% of total transition
Path 1	Relative Healthy	Relative Healthy	Relative Healthy	Relative Healthy	Relative Healthy	30.5
Path 2	MH	MH	MH	MH	MH	20.1
Path 3	CV	CV	CV	CV	CV	12.2
Path 4	CV-MSK	CV-MSK	CV-MSK	CV-MSK	CV-MSK	8.1
Path 5	MSK	MSK	MSK	MSK	CV-MSK	6.3
Path 6	MSK	MSK	MSK	MSK	MSK	4.9
Path 7	MSK	MSK	MSK	CV-MSK	CV-MSK	2.1
Path 8	Relative Healthy	CV	CV	CV	CV	1.6
Path 9	MSK	CV-MSK	CV-MSK	CV-MSK	CV-MSK	1.6
Path 10	CV	CV	CV	CV	CV-MSK	1.3
Path 11	MSK	MSK	CV-MSK	CV-MSK	CV-MSK	1.1
Path 12	CV	CV-MSK	CV-MSK	CV-MSK	CV-MSK	1.0

Names in each cell at each time point represent the leading conditions in the cluster.

CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal

Appendix Table 31 Statistical parameters for each class IN RMLCA in people with OA

	Model	log-likelihood	BIC	SABIC	likelihood-ratio	AIC	Entropy	npar	C1	C2	C3	C4	C5	C6	C7	C8	C9
At year 5	1	-36607	-	73526.43	8174.411	73308.04	47	NA	100								
	2	-34666.4	70266.08	69964.18	4293.128	69522.76	0.546	95	85.97	14.03							
	3	-34504.9	70414.71	69960.26	3970.182	69295.81	0.498	143	6.66	83.59	9.75						
	4	-34438	70752.42	70145.43	3836.317	69257.95	0.501	191	7.66	8.21	0.56	83.57					
	5	-34388.7	71125.49	70365.96	3737.82	69255.45	0.504	239	0.87	83.45	6.3	8.68	0.71				
	6	-34338	71495.62	70583.55	3636.379	69250.01	0.498	287	82.33	1.03	5.53	10	0.45	0.65			
	7	-34307.7	71906.56	70841.95	3575.745	69285.37	0.513	335	0.13	1.26	82.51	0.29	0.79	5.09	9.95		
	8	-34276.9	72316.6	71099.44	3514.205	69319.83	0.461	383	5.48	0.26	1.15	12	2.24	1.82	0.52	76.52	
	9	-34242.6	72719.57	71349.87	3445.602	69347.23	0.459	431	0.26	0.28	1.09	2.59	11.91	76.11	1.73	1.49	4.54
At year 10	1	-34330.8	-	68956.59	13877.73	68759.59	49	NA	100								
	2	-31449.5	63809.67	63495.06	8115.187	63097.05	0.716	99	76.86	23.14							
	3	-31209.5	63789.54	63316.04	7635.152	62717.02	0.651	149	13.77	74.19	12.04						
	4	-31099.9	64030.17	63397.78	7415.868	62597.73	0.621	199	4.47	12.31	10.29	72.93					
	5	-31034.3	64358.86	63567.57	7284.651	62566.52	0.62	249	10.46	72.93	1.08	12.15	3.39				
	6	-30981.8	64713.77	63763.59	7179.654	62561.52	0.63	299	71.53	1.89	12.8	2	1.7	10.08			
	7	-30941.3	65092.69	63983.62	7098.665	62580.53	0.545	349	1.21	66.28	4.69	1.85	10.19	13.53	2.25		
	8	-30901.2	65472.41	64204.45	7018.477	62600.34	0.637	399	0.9	1.18	12.19	10.79	1.12	1.27	71.96	0.58	
	9	-30857.8	65845.65	64418.79	6931.806	62613.67	0.583	449	4.05	8.33	14.05	0.69	2.25	67.3	1.36	1.48	0.49
At year 15	1	-16252.6	-	32732.54	9454.857	32597.19	46	NA	100								
	2	-14583.1	29921.4	29625.89	6115.909	29352.24	0.812	93	31.42	68.58							
	3	-14457	30050.84	29606	5863.712	29194.04	0.715	140	67.5	16.82	15.68						
	4	-14389.2	30296.93	29702.75	5728.161	29152.49	0.711	187	18.43	12.44	67.04	2.09					
	5	-14337.4	30574.87	29831.35	5624.464	29142.79	0.698	234	6.81	14.36	66.99	2.92	8.92				
	6	-14298.1	30877.86	29985	5545.816	29158.15	0.684	281	9.95	6.13	1.26	65.83	7.06	9.77			
	7	-14268.9	31201.08	30158.87	5487.387	29193.72	0.353	328	65.29	0.22	2.17	13.68	6.11	7.14	5.39		
	8	-14232.8	31510.68	30319.13	5415.348	29215.68	0.706	375	8.97	1.67	0.67	66.55	1.66	12.53	2.64	5.32	
	9	-14213.2	31853.05	30512.16	5376.079	29270.41	0.744	422	1.05	11.98	1.96	0.68	4.14	65.76	1.05	10.72	2.66

BIC- Bayesian information criteria; AIC- Akaike information criteria; SABIC- Sample size adjusted BIC

Appendix Table 32 Statistical parameters for each class IN RMLCA in people with non-OA

	Model	log-likelihood	BIC	SABIC	likelihood-ratio	AIC	Entropy	npar	C1	C2	C3	C4	C5	C6	C7	C8	C9
At year 5	1	-32360.3	-	65038.7	5982.273	64814.69	47	NA	100								
	2	-30977.5	62899.74	62597.83	3216.642	62145.06	0.493	95	10.23	89.77							
	3	-30821	63063.89	62609.44	2903.483	61927.9	0.413	143	4.16	83.23	12.61						
	4	-30750.3	63399.8	62792.81	2762.085	61882.51	0.43	191	83.01	3.87	1.24	11.88					
	5	-30714.2	63804.94	63045.41	2689.915	61906.34	0.394	239	3.08	81.28	1.01	0.83	13.8				
	6	-30673.2	64200.35	63288.27	2608.012	61920.43	0.435	287	0.45	81.56	0.33	1.82	4.89	10.95			
	7	-30645.1	64621.33	63556.72	2551.686	61960.11	0.407	335	4.13	0.79	0.26	10.93	0.44	80.31	3.14		
	8	-30609.8	65028.13	63810.98	2481.178	61985.6	0.234	383	3.87	3.17	0.44	43.35	0.28	4.47	0.19	44.23	
	9	-30570.7	65427.26	64057.56	2402.998	62003.42	0.425	431	0.45	0.64	1.18	2.84	2.66	78.55	13.39	0.02	0.27
At year 10	1	-35804.0	-	71915.58	11510.43	71706.01	49	NA	100								
	2	-33329.3	67594.59	67279.98	6560.98	66856.57	0.642	99	19.25	80.75							
	3	-33100.6	67610.05	67136.54	6103.701	66499.29	0.57	149	76.76	15.58	7.66						
	4	-33016.8	67915.17	67282.77	5936.082	66431.67	0.592	199	1.16	13.9	77.37	7.57					
	5	-32942.2	68238.6	67447.31	5786.78	66382.37	0.57	249	2.43	76.13	13.27	6.67	1.51				
	6	-32888.1	68603.09	67652.9	5678.529	66374.11	0.563	299	0.55	14.46	6.65	1.34	1.94	75.06			
	7	-32863.1	69025.98	67916.9	5628.682	66424.27	0.343	349	1.38	30.8	0.56	6.92	1.44	48.82	10.08		
	8	-32798.4	69369.16	68101.18	5499.118	66394.7	0.574	399	5.02	1.53	1.44	14.47	0.44	1.09	3.56	72.45	
	9	-32773.3	69791.71	68364.83	5448.925	66444.51	0.552	449	8.42	5.07	6.87	3.25	1.08	72.42	0.76	0.86	1.28
At year 15	1	-19936.8	-	40127.72	9838.909	39969.64	48	NA	100								
	2	-18193.4	37208.52	36900.29	6352.092	36580.82	0.748	97	72.29	27.71							
	3	-18013.8	37264.47	36800.54	5992.959	36319.69	0.673	146	12.7	69.21	18.09						
	4	-17935.2	37522.2	36902.56	5835.599	36260.32	0.662	195	17.52	68.15	4.64	9.69					
	5	-17872.5	37812.04	37036.69	5710.35	36233.08	0.67	244	9.32	1.32	68.22	4.12	17.01				
	6	-17828.3	38138.59	37207.54	5621.816	36242.54	0.671	293	2.43	3.39	8.09	16.86	67.75	1.48			
	7	-17785.3	38467.72	37380.97	5535.858	36254.58	0.701	342	1.94	2.5	2.4	0.92	67.43	17.11	7.72		
	8	-17750.0	38812.21	37569.76	5465.266	36281.99	0.644	391	2.61	0.52	9.61	1.51	7.69	10.51	1.3	66.25	
	9	-17725.6	39178.6	37780.44	5416.566	36331.29	0.673	440	2.89	1.72	1.04	1.07	66.57	12.16	7	5.46	2.08

BIC- Bayesian information criteria; AIC- Akaike information criteria; SABIC- Sample size adjusted BIC

Appendix Table 33 Posterior probabilities distribution of conditions in OA after 5 years (sensitivity analysis)

Conditions	Relative Healthy (83.58%)	Cardiovascular (6.66%)	Musculoskeletal (9.75%)
Anaemia	0	2	3
Ankylosing Spondylitis	0	0	1
Arterial and venous diseases	0	0	0
Asthma	0	4	4
Backpain	1	27	**43**
Benign Prostatic Hypertrophy^	0	3	3
Cancer	0	5	1
Cataract	6	6	1
Cerebral stroke	3	4	2
Chronic Kidney Disease	0	7	0
COPD	0	2	2
Coronary Heart Disease	0	9	1
Dementia	0	1	0
Depression	0	10	18
Diabetes	0	13	0
Epilepsy	0	1	1
Fatigue	0	0	1
Fibromyalgia	0	0	1
Gall stones	0	2	2
Gastritis	0	3	2
Gastrointestinal bleed	0	1	0
Gout	0	3	1
Hearing problem	0	7	4
Heart failure	0	1	0
Hypercholesteremia	0	18	3
Hypertension	1	**41**	5
Hyperthyroidism	0	1	0
Hypothyroidism	0	6	1
Inflammatory Bowel Disease	0	2	2
Irritable Bowel Syndrome	7	3	8
Liver disease	0	0	0
Migraine	0	0	4
Multiple Sclerosis	0	0	0
Osteoporosis	0	3	1
Parkinson Disease	0	0	0
Peripheral Vascular Disease	0	2	1
Polymyalgia	0	1	0
Psoriasis	0	1	1
Psychosis	0	0	0
Renal stone	0	0	0
Rheumatoid arthritis	0	1	1
Schizophrenia	0	1	0
Sjogren's syndrome	0	0	0
Sleep problem	0	2	1
Systemic Lupus Erythematous	0	0	0
Tuberculosis	0	0	0
Vision problem	0	1	0

^only for men

Appendix Table 34 Posterior probabilities after 10 years in OA

Conditions	Relative Healthy (74.19%)	Musculoskeletal (13.76%)	Cardiovascular (12.04%)
Anaemia	0	3	4
Ankylosing Spondylitis	0	2	0
Arterial and venous diseases	0	0	1
Asthma	0	5	7
Backpain	1	**58**	42
Benign Prostatic Hypertrophy^	0	2	7
Cancer	0	3	8
Cataract	9	0	13
Cerebral stroke	7	2	9
Chronic Kidney Disease	0	2	15
COPD	0	2	4
Coronary Heart Disease	0	1	12
Dementia	0	0	3
Depression	0	26	13
Diabetes	0	2	20
Epilepsy	0	0	1
Fatigue	0	2	1
Fibromyalgia	0	2	0
Gall stones	0	3	3
Gastritis	0	4	5
Gout	0	2	7
Gastrointestinal bleed	0	1	1
Hearing problem	1	8	12
Heart failure	0	0	2
HIV/AIDS	0	0	0
Hypercholesteremia	0	7	22
Hypertension	1	8	**54**
Hyperthyroidism	0	0	2
Hypothyroidism	0	3	7
Inflammatory Bowel Disease	0	5	3
Irritable Bowel Syndrome	8	9	4
Liver disease	0	1	1
Migraine	0	6	2
Multiple Sclerosis	0	0	0
Osteoporosis	0	3	6
Parkinson Disease	0	0	0
Peripheral Vascular Disease	0	2	4
Polymyalgia	0	0	1
Psoriasis	0	2	2
Psychosis	0	0	1
Renal stone	0	0	1
Rheumatoid arthritis	0	1	2
Schizophrenia	0	0	1
Scleroderma	0	0	0
Sjogren's syndrome	0	0	0
Sleep problem	0	3	3
Systemic Lupus Erythematous	0	0	0
Tuberculosis	0	0	0
Vision problem	0	0	1

^only for men

Appendix Table 35 Posterior probabilities after 15 years in OA

Conditions	Healthy (66.99%)	CV-MSK (2.91%)	MSK (6.81%)	CV (14.36%)	MSK-MH (8.92%)
Anaemia	0	8	4	4	7
Ankylosing Spondylitis	0	2	5	0	0
Arterial and venous diseases	0	9	1	0	0
Asthma	0	18	9	5	11
Backpain	2	63	70	44	67
Benign Prostatic Hypertrophy^	0	14	14	7	0
Cancer	1	16	9	9	4
Cataract	11	31	6	6	0
Cerebral stroke	11	14	6	8	0
Chronic Kidney Disease	0	42	0	12	5
COPD	0	9	2	4	6
Coronary Heart Disease	0	34	0	10	3
Dementia	0	11	3	0	0
Depression	1	25	16	11	48
Diabetes	0	25	0	26	6
Epilepsy	0	3	1	0	1
Fatigue	0	6	0	0	6
Fibromyalgia	0	3	0	0	5
Gall stones	0	10	3	2	9
Gastritis	0	9	8	4	6
Gout	0	15	10	10	0
Gastrointestinal bleed	0	2	0	1	0
Hearing problem	0	31	20	13	8
Heart failure	0	11	0	0	0
Hypercholesteremia	0	26	13	29	10
Hypertension	1	64	7	59	24
Hyperthyroidism	0	8	0	0	0
Hypothyroidism	0	20	0	6	9
Inflammatory Bowel Disease	0	5	6	2	6
Irritable Bowel Syndrome	9	2	7	3	16
Liver disease	0	1	3	1	0
Migraine	0	0	5	2	11
Multiple Sclerosis	0	0	0	0	1
Osteoporosis	0	14	11	4	3
Parkinson Disease	0	3	2	0	0
Peripheral Vascular Disease	0	12	5	2	3
Polymyalgia	0	5	2	1	0
Psoriasis	0	5	5	1	5
Psychosis	0	0	0	0	0
Renal stone	0	1	3	1	1
Rheumatoid arthritis	0	0	0	3	3
Schizophrenia	0	0	2	0	0
Scleroderma	0	0	1	0	0
Sjogren's syndrome	0	0	0	0	1
Sleep problem	0	7	5	2	3
Vision problem	0	4	0	0	0

CV- Cardiovascular; MH- Mental Health; MSK- Musculoskeletal; ^only for men

Appendix Table 36 Posterior probabilities after year 5 in non-OA

Conditions	Healthy (83.00%)	CV-MSK (1.24%)	CV (3.87%)	MSK (11.88%)
Anaemia	0	8	0	1
Ankylosing Spondylitis	0	0	0	0
Arterial and venous diseases	0	0	1	0
Asthma	0	2	3	4
Backpain	0	**21**	17	**27**
Benign Prostatic Hypertrophy^	0	3	2	1
Cancer	0	7	2	2
Cataract	7	13	4	0
Cerebral stroke	3	12	5	0
Chronic Kidney Disease	0	13	4	0
COPD	0	6	1	1
Coronary Heart Disease	0	16	7	1
Dementia	0	3	0	0
Depression	0	6	8	9
Diabetes	0	7	11	1
Epilepsy	0	1	0	0
Fatigue	0	3	0	1
Fibromyalgia	0	1	0	0
Gall stones	0	8	0	1
Gastritis	0	8	0	1
Gout	0	4	3	1
Gastrointestinal bleed	0	1	0	0
Hearing problem	0	12	3	4
Heart failure	0	5	0	0
Hypercholesteremia	0	8	19	3
Hypertension	1	**31**	**56**	0
Hyperthyroidism	0	4	0	0
Hypothyroidism	0	17	1	2
Inflammatory Bowel Disease	0	6	0	3
Irritable Bowel Syndrome	5	8	1	4
Liver disease	0	0	0	0
Migraine	0	0	1	2
Multiple Sclerosis	0	0	0	0
Osteoporosis	0	1	1	1
Parkinson Disease	0	1	0	0
Peripheral Vascular Disease	0	5	1	0
Polymyalgia	0	1	0	0
Psoriasis	0	1	1	1
Psychosis	0	0	0	0
Renal stone	0	0	1	0
Rheumatoid arthritis	0	1	0	0
Schizophrenia	0	0	0	0
Scleroderma	0	0	0	0
Sjogren’s syndrome	0	0	0	0
Sleep problem	0	0	2	0
Tuberculosis	0	0	1	0
Vision problem	0	2	0	0

CV- Cardiovascular; MSK- Musculoskeletal; ^only for men

Appendix Table 37 Posterior probabilities after year 10 in non-OA

Conditions	Healthy (76.12%)	CV-MSK (2.42%)	MSK (13.26%)	CV (6.67%)	Metabolic (1.51%)
Anaemia	0	5	2	3	2
Ankylosing Spondylitis	0	0	1	0	0
Arterial and venous diseases	0	1	0	1	0
Asthma	0	8	6	3	2
Backpain	1	**48**	**45**	22	17
Benign Prostatic Hypertrophy^	0	7	2	5	4
Cancer	0	9	4	7	5
Cataract	9	21	0	8	0
Cerebral stroke	7	15	1	9	1
Chronic Kidney Disease	0	18	0	18	5
COPD	0	9	2	4	0
Coronary Heart Disease	0	24	2	6	0
Dementia	0	5	0	1	0
Depression	0	15	16	5	22
Diabetes	0	12	3	16	6
Epilepsy	0	1	0	0	1
Fatigue	0	3	1	0	3
Fibromyalgia	0	1	0	0	1
Gall stones	0	11	2	0	3
Gastritis	0	11	2	2	0
Gout	0	4	1	6	0
Gastrointestinal bleed	0	0	0	0	0
Hearing problem	0	27	7	6	0
Heart failure	0	5	0	0	0
HIV/AIDS	0	0	0	0	1
Hypercholesteremia	0	17	6	24	4
Hypertension	1	**42**	10	**58**	11
Hyperthyroidism	0	1	0	0	13
Hypothyroidism	0	11	0	3	**50**
Inflammatory Bowel Disease	0	10	4	0	1
Irritable Bowel Syndrome	6	9	5	1	3
Liver disease	0	0	0	1	0
Migraine	0	0	3	2	4
Multiple Sclerosis	0	0	0	0	2
Osteoporosis	0	6	2	3	1
Parkinson Disease	0	1	0	0	3
Peripheral Vascular Disease	0	7	0	3	0
Polymyalgia	0	1	0	1	0
Psoriasis	0	4	2	0	3
Psychosis	0	0	0	0	2
Renal stone	0	0	1	2	1
Rheumatoid arthritis	0	0	1	0	1
Schizophrenia	0	1	0	0	5
Scleroderma	0	0	0	0	0
Sjogren’s syndrome	0	0	0	0	1
Sleep problem	0	4	2	0	0
Systemic Lupus Erythematous	0	0	0	0	0
Tuberculosis	0	0	0	1	0
Vision problem	0	1	0	0	0

CV- Cardiovascular; MSK- Musculoskeletal; ^only for men

Appendix Table 38 Posterior probabilities 15 years in non-OA

Conditions	Healthy (68.22%)	CV-MSK (4.12%)	MSK (17.01%)	CV (9.32%)	Metabolic (1.32%)
Anaemia	0	12	3	0	10
Ankylosing Spondylitis	0	0	1	0	0
Arterial and venous	0	3	0	1	2
Asthma	0	11	7	5	5
Backpain	1	**68**	**54**	32	43
Benign Prostatic Hypertrophy^	0	10	3	5	0
Cancer	0	10	6	10	1
Cataract	11	31	1	10	0
Cerebral stroke	10	28	1	9	0
Chronic Kidney Disease	0	20	1	26	0
COPD	0	12	5	3	0
Coronary Heart Disease	0	22	4	10	0
Dementia	0	8	1	2	0
Depression	1	22	21	5	24
Diabetes	0	27	0	19	12
Epilepsy	0	2	1	0	4
Fatigue	0	5	1	0	5
Fibromyalgia	0	1	0	0	0
Gall stones	0	4	3	6	6
Gastritis	0	12	3	3	0
Gout	0	0	2	9	0
Gastrointestinal bleed	0	2	1	0	0
Hearing problem	1	30	10	12	0
Heart failure	0	3	0	2	0
HIV/AIDS	0	0	0	0	0
Hypercholesteremia	0	27	9	30	9
Hypertension	1	**50**	16	**68**	14
Hyperthyroidism	0	1	0	0	31
Hypothyroidism	0	10	1	4	**75**
Inflammatory Bowel Disease	0	9	6	0	3
Irritable Bowel Syndrome	5	8	7	2	6
Liver disease	0	0	0	1	0
Migraine	0	2	6	1	5
Multiple Sclerosis	0	0	0	0	2
Osteoporosis	0	5	4	7	2
Parkinson Disease	0	1	0	0	0
Peripheral Vascular Disease	0	11	1	3	0
Polymyalgia	0	0	1	2	0
Psoriasis	0	3	2	3	4
Psychosis	0	1	0	0	0
Renal stone	0	1	2	2	0
Rheumatoid arthritis	0	0	1	0	2
Schizophrenia	0	2	1	0	0
Sjogren's syndrome	0	0	0	0	5
Sleep problem	0	8	3	0	3
Systemic Lupus Erythematous	0	0	0	0	0
Tuberculosis	0	2	0	0	0
Vision problem	0	3	0	0	0

CV- Cardiovascular; MSK- Musculoskeletal; ^only for men

Appendix Table 39 Checklist for Latent class growth analysis

GRoLTS checklist item	Yes/No	Comments
1. Is the metric of time used in the statistical model reported?	Yes	Time metrics is years
2. Is information presented about the mean and variance of time within a wave?	NA	Exact time of the measurement is used in the analysis, so time-structured data is not relevant in this study.
3a. Is the missing data mechanism reported?	Yes	
3b. Is a description provided of what variables are related to attrition/missing data?	Yes	Described in the model specification in the manuscript.
3c. Is a description provided of how missing data in the analyses were dealt with?	Yes	
4. Is information about the distribution of the observed variables included?	Yes	
5. Is the software mentioned?	Yes	R with 'lcmm' package
6a. Are alternative specifications of within-class heterogeneity considered (e.g., LGCA vs. LGMM) and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	Yes	
6b. Are alternative specifications of the between-class differences in variance–covariance matrix structure considered and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	Yes	
7. Are alternative shape/functional forms of the trajectories described?	Yes	Linear, cubic, and quadratic functions were explored
8. If covariates have been used, can analyses still be replicated?	No	
9. Is information reported about the number of random start values and final iterations included?	Yes	
10. Are the model comparison (and selection) tools described from a statistical perspective?	Yes	BIC and AIC were used
11. Are the total number of fitted models reported, including a one-class solution?	Yes	
12. Are the number of cases per class reported for each model (absolute sample size, or proportion)?	Yes	
13. If classification of cases in a trajectory is the goal, is entropy reported?	Yes	
14a. Is a plot included with the estimated mean trajectories of the final solution?	Yes	
14b. Are plots included with the estimated mean trajectories for each model?	Yes	
14c. Is a plot included of the combination of estimated means of the final model and the observed individual trajectories split out for each latent class?	No.	Only observed trajectories are reported.
15. Are characteristics of the final class solution numerically described (i.e., means, SD/SE, n, CI, etc.)?	Yes	Proportion and different SES
16. Are the syntax files available (either in the appendix, supplementary materials, or from the authors)?	Yes	Available on request from the authors.

Appendix Table 40 Disability weights for UK and Europe

Conditions	Disability weight	Severity/Type
Anaemia	0.118	Severe
Ankylosing Spondylitis		
Arterial/Venous	0.647	
Asthma	0.045	Partly controlled
Back pain	0.365	Severe without leg pain
Benign Prostate Hypertrophy	0.067	
Cancer	0.451	Metastatic
Cancer	0.288	Non metastatic
Cataract	0.17	
Chronic Kidney Disease	0.104	Stage 4
COPD	0.225	Moderate
Coronary Heart Disease	0.432	MI
Dementia	0.377	Moderate
Depression	0.396	Moderate
Diabetes Mellitus	0.015	
Epilepsy	0.552	Severe
Fatigue		
Fibromyalgia		
Gall stones	0.448	
Gastritis	0.003	
Gastrointestinal bleed	0.325	
Gout	0.295	Acute
Hearing problem	0.158	Severe
Heart failure	0.179	Severe
High Cholesterol	0.304	
Hypertension	0.502	
Hyperthyroid	0.145	
Hypothyroidism	0.019	
Inflammatory Bowel Disease	0.231	
Irritable Bowel Syndrome	0.062	
Liver Disease	0.178	
Migraine	0.441	
Multiple sclerosis	0.719	Severe
Osteoporosis		
Parkinson Disease	0.575	Severe
Peripheral Vascular Disease	0.43	
Polymyalgia		
Psoriasis	0.235	
Psychosis		
Renal stone	0.294	
Rheumatoid Arthritis	0.199	
Schizophrenia	0.778	Acute
Scleroderma		
Sjogren's syndrome		
Sleep Disorder	0.1	
Stroke	0.316	Long with cognition
Systemic Lupus Erythematous	0.594	
Tuberculosis	0.333	Without HIV
Vision Problem	0.011	Presbyopia

Appendix Table 41 Elixhauser Comorbidity index

Comorbidity Domain	AHRQ Algorithm	van Walraven algorithm	Charlson Comorbidity Index	This study
Congestive heart failure	9	7	1	7
Cardiac arrhythmias	0	5		5
Valvular disease	0	-1		-1
Pulmonary circulation disorders	6	4		4
Peripheral vascular disorders	3	2	1	2
Hypertension (combined uncomplicated and complicated)	-1	0		0
Paralysis	5	7	1	7
Other neurological disorders	5	6	1	6
Chronic pulmonary disease	3	3	1	3
Diabetes, uncomplicated	0	0	1	0
Diabetes, complicated	-3	0	2	0
Hypothyroidism	0	0		0
Renal failure	6	5	2	5
Liver disease	4	11	1	11
Peptic ulcer disease, excluding bleeding	0	0	1	0
AIDS/HIV	0	0	6	0
Lymphoma	6	9	2	8
Metastatic cancer	14	12	6	8
Solid tumour without metastasis	7	4	2	8
Rheumatoid arthritis/collagen vascular diseases	0	0	1	0
Coagulopathy	11	3		
Obesity	-5	-4		
Weight loss	9	6		
Fluid and electrolyte disorders	11	5		
Blood loss anaemia	-3	-2		-2
Deficiency anaemia	-2	-2		-2
Alcohol abuse	-1	0		
Drug abuse	-7	-7		
Psychoses	-5	0		0
Depression	-5	-3		-3
Dementia			1	
Leukaemia			2	
Severe liver disease			3	

Appendix Method 1 False Discovery Rate (FDR) test methods

Controlling the false discovery rate: Benjamini–Hochberg procedure

An approach to manage multiple testing problem is false discovery rate. This is the proportion of "discoveries" (significant results) that are false positives.

One good technique for controlling the false discovery rate was briefly mentioned by Simes (1986) and developed in detail by Benjamini and Hochberg (1995).

Steps of FDR:

1. Put the individual *P* values in order, from smallest to largest.
2. The smallest *P* value has a rank of *i*=1, then next smallest has *i*=2, etc.
3. Compare each individual *P* value to its

Benjamini-Hochberg adjusted P value = (*i*/*m*)*Q*,

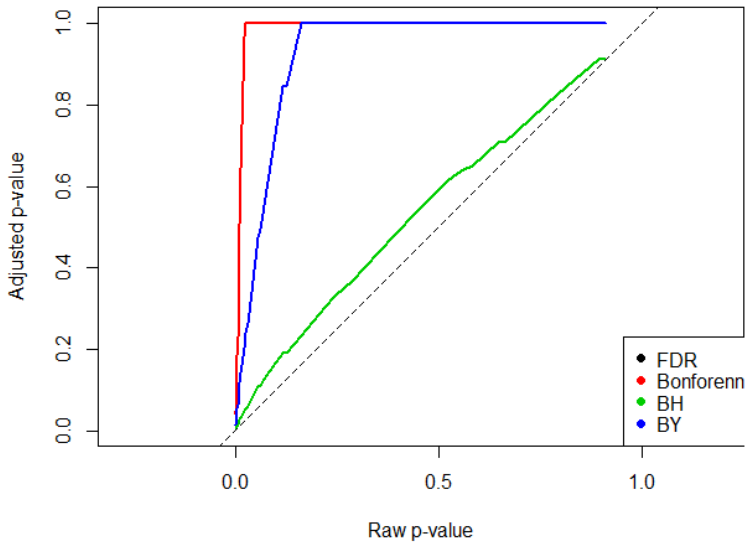
where *i* is the rank,

m is the total number of tests, and

Q is the study p value.

4. The adjusted p value less than the alpha error is considered as significant.

Sensitivity testing of different methods to estimate adjusted p value



Appendix Method 2 Model statistics for different methods for cluster analysis

	1000 random samples			2000 random samples		
	HCL	LCA	K-Mode	HCL	LCA	K-Mode
Calinski Harabasz	60.46	65.39	40.89	23.97	25.04	20.59
Tau	0.498	0.371	0.499	0.53	0.50	0.40
Silhouette	0.009	0.023	0.009	0.001	0.09	0.049

HCL -Hierarchical clustering analysis

LCA- Latent class analysis

Calinski Harabasz (1)

This is also known as the Variance Ratio Criterion. The score is defined as the ratio between the within-cluster dispersion and the between-cluster dispersion. The higher the score the better is the cluster.

Tau Index (2)

Silhouette Index (3)

1. Calinski T, Harabasz J (1974). “A Dendrite Method for Cluster Analysis.” Communications in Statistics – Theory and Methods, 3(1), 1–27.

2. Milligan GW (1981). “A Monte Carlo Study of Thirty Internal Criterion Measures for Cluster Analysis.” Psychometrika, 46(2), 187–199.

3. Rousseeuw P (1987). “Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis.” Journal of Computational and Applied Mathematics, 20, 53–65.

Arthritis Care & Research Vol. 0, No. 0, Month 2020, pp 1–10 DOI 10.1002/acr.24008

American College of Rheumatology
Empowering Rheumatology Professionals

Comorbidities in Osteoarthritis: A Systematic Review and Meta-Analysis of Observational Studies

Subhashisa Swain,[1] Aliya Sarmanova,[1] Carol Coupland,[2] Michael Doherty,[1] and Weiya Zhang[1]

Objective. Osteoarthritis (OA) is a common chronic condition in older individuals, but its association with other chronic conditions is largely unknown. This study aimed to systematically review the literature on comorbidities in individuals with OA compared to those without.

Methods. We searched 4 databases for observational studies on comorbidities in individuals with OA. Studies of OA only or in comparison with non-OA controls were included. The risk of bias and study quality were assessed using the Newcastle-Ottawa Scale. The prevalence of comorbidities in the OA group and the prevalence ratio (PR) and 95% confidence interval (95% CI) between OA and non-OA groups were calculated.

Results. In all, 42 studies from 16 countries (27 case-only and 15 comparative studies) met the inclusion criteria. The mean age of participants varied from 51 to 76 years. The pooled prevalence of any comorbidity was 67% (95% CI 57–74) in individuals with OA versus 56% (95% CI 44–68) in individuals without OA. The pooled PR for any comorbidity was 1.21 (95% CI 1.02–1.45). The PR increased from 0.73 (95% CI 0.43–1.25) for 1 comorbidity to 1.58 (95% CI 1.03–2.42) for 2, and to 1.94 (95% CI 1.45–2.59) for ≥3 comorbidities. The key comorbidities associated with OA were stroke (PR 2.61 [95% CI 2.13–3.21]), peptic ulcer (PR 2.36 [95% CI 1.71–3.27]), and metabolic syndrome (PR 1.94 [95% CI 1.21–3.12]).

Conclusion. Individuals with OA are more likely to have other chronic conditions. The association is dose-dependent in terms of the number of comorbidities, suggesting multimorbidities. Further studies on the causality of this association and clinical implications are needed.

INTRODUCTION

Osteoarthritis (OA) is by far the most common form of arthritis and is a major cause of pain and disability in older individuals (1). It is a common, complex disorder with multiple genetic, constitutional, and environmental risk factors (2). The presence of multiple chronic conditions in a single individual causes higher mortality, increased hospitalization, impaired physical and mental health, worse disease outcome, and poorer quality of life (3,4). The coexistence of chronic conditions with OA is also very common, especially in the later decades of life (5,6). For example, according to the Centers for Disease Control and Prevention, >30% of individuals with diabetes mellitus and heart disease have OA (7).

Most literature on OA comorbidity was published in the last 3 years. The review articles focused on the distribution and impact of individual chronic conditions such as cardiovascular diseases, diabetes mellitus, and depression in OA (8–11). Even though comorbidity was discussed as a concept in the 1960s, only in 1996 was a distinct definition first suggested to differentiate comorbidity (implying an index disease with mechanistically linked additional conditions) and multimorbidity (implying any co-occurrence of medical conditions) within an individual (12). Comorbidity research in OA is still at a preliminary stage, and the evidence is yet to be accumulated.

A systematic review on OA reported worsening of pain and a decline in functional activities among individuals due to the presence of other chronic conditions (13). Clinically, comorbidities in OA create greater challenges for management. The number and pattern of different comorbid conditions determine the severity and burden in patients with multimorbidities (14,15). However, except for shared risk factors such as aging and obesity, little is known about biologic plausibility to explain the concurrence of OA and associated comorbidities (16,17). According to the European League Against Rheumatism (EULAR) and the National Institute

PROSPERO: CRD42016038484.

Supported by Versus Arthritis (grants 20777 and 21595), formerly Arthritis Research UK, the University of Nottingham Vice-Chancellor's Scholarship, and a Beijing Joint Care Foundation Scholarship.

[1]Subhashisa Swain, MPH, Aliya Sarmanova, PhD, Michael Doherty, MD, Weiya Zhang, PhD: University of Nottingham, Nottingham, UK; [2]Carol Coupland, PhD: School of Medicine, University of Nottingham, Nottingham, UK.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Weiya Zhang, PhD, Academic Rheumatology, Division of Rheumatology, Orthopaedics and Dermatology, The University of Nottingham, Nottingham, UK. E-mail: Weiya.Zhang@nottingham.ac.uk.

Submitted for publication October 11, 2018; accepted in revised form June 11, 2019.

SIGNIFICANCE & INNOVATIONS

- This is the first systematic review of the current literature on comorbidities in osteoarthritis (OA), with an extensive list of the conditions.
- In total, 67% of individuals with OA have at least 1 other chronic condition, which is 20% higher than for those without OA.
- There was a graded effect in terms of the risk of having 1, 2, and ≥3 comorbidities in individuals with OA compared to those without.
- In individuals with OA, the systems most likely to be affected by comorbidities are upper gastrointestinal, psychological, cardiovascular, and endocrine.
- Stroke, peptic ulcer, and metabolic syndrome are the most common comorbidities.

for Health and Care Excellence, the diagnosis and management of specific comorbidities and understanding their pattern in OA are important and are recommended for best practice (18,19). An Arthritis Research UK report on multimorbidity in OA also highlighted the importance of understanding the presence of multiple comorbidities with OA for formulating a patient-centered management plan (20). This study aimed to systematically review the current literature on the comorbidities in OA, specifically, the risk (prevalence or incidence) of comorbidities in individuals with OA compared to those without OA.

MATERIALS AND METHODS

Search methods and sources. A protocol adhering to the Preferred Reporting Items for Systematic Review and Meta-Analysis 2015 statement was designed and registered online with PROSPERO. Medline, PubMed, Embase, and Scopus databases were used to identify studies conducted in any country between January 1, 1995 and December 31, 2017. Additionally, "comorbidity in OA" was searched in the Google Scholar search engine, and the first 1,000 articles were screened for inclusion. The complete search consisted of searches for OA (any joint), searches for comorbidities, and searches for observational studies. The 3 search strategies were then combined using AND to generate citations. The details of the search strategies can be seen in Supplementary Appendix A, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.24008/abstract. In addition, websites of international societies dealing with arthritis such as EULAR, the American College of Rheumatology, and Osteoarthritis Research Society International were searched (18,21,22). References within systematic reviews and review articles were also read for additional relevant original articles.

Selection criteria. All types of observational studies (with or without a non-OA control) documenting prevalence or incidence and the risk ratio of OA comorbidity were included in the study. We defined comorbidity as the presence of any concurrent chronic condition in individuals with OA (as an index disease).

Studies included individuals with OA diagnosed by a physician through physical examination or radiographic findings. OA was the primary exposure, and outcomes were the presence of any comorbidities. Other comparisons included studies comparing the prevalence/incidence of comorbidities in OA with non-OA controls (comparative) and studies of comorbidities in individuals with OA (case-only).

According to the above criteria, all studies identified by title and abstract were gathered, and duplicates were removed. Potentially relevant articles were selected through initial title and abstract screening by 2 authors (SS and AS) independently. Any disagreement was discussed with a third author (WZ). The full text copies of these relevant articles were then retrieved. We retained articles that studied the prevalence of other chronic conditions in individuals with OA. Full texts of potentially suitable articles were further screened for inclusion (SS). Disagreement in the screening of full texts was resolved by a third reviewer (WZ). There was no language limitation. We used Endnote for screening of articles, and data extraction was done without using any software.

Quality assessment. One reviewer (SS) independently assessed study quality based on items in the Newcastle-Ottawa Scale (NOS) checklist (23). Any concern on quality scoring was decided in consultation with another reviewer (AS or WZ). The NOS tool has a scoring scale under 3 sections: participant selection and representativeness, comparability of study groups, and assessment of outcome or exposure. The quality score is based on a star system (range 0–9 stars for case–control and cohort studies and 0–10 for cross-sectional studies), with a higher score representing better methodologic quality.

Data extraction. A customized data extraction form was used to extract data from each study. For each included study, we collected the following information: authors and publication year, title and journal, study country and location (urban or rural), study design, sampling method (random or nonrandom), sample size, sample characteristics such as age and sex, the number of conditions included, methods of comorbidity measurement, and prevalence (overall and group specific for each comorbidity).

Outcome. The primary outcome was the risk (prevalence) of comorbidities in individuals with OA (cases) versus those without OA (controls), and secondary outcomes included the types of comorbidities associated with OA. The risk of having the comorbidity between OA and non-OA controls was estimated through the prevalence ratio (PR), separately for all comparative studies and for age- and sex-matched/adjusted comparative studies. For cohort studies, the prevalence of comorbidities reported at baseline was included for the estimation because in these studies comorbidity was not reported as the outcome.

Statistical analysis. Descriptive characteristics of the studies are expressed as means/medians and/or frequencies, as appropriate, depending on the variables. For comorbidity count, we used the median because of wider variation in the list of the diseases across the studies. Heterogeneity between studies was measured using the I^2 (%) and Q test (*P* value) (24,25). Publication bias was assessed using funnel plots and Egger's test, with statistical significance being conferred to a *P* value less than 0.05 (26). For heterogeneity, I^2 above 75% was considered as wider heterogeneity, demanding careful interpretation of the findings (25). Prevalence and PR and 95% confidence intervals (95% CIs) were calculated wherever possible for each comorbidity. The PR was chosen over odds ratio (OR) because we had prevalence data for both OA cases (exposure) and non-OA controls (non-exposure). In this scenario, PR is recommended over OR to minimize the overestimation of the relative risk (27). For prevalence estimation, subgroup analysis was done according to the study design (cross-sectional, case–control, and cohort). For PR, however, only 1 article had a different study design from the others, thus not allowing us to perform the subgroup analysis as per the study design. Therefore, for the PR estimation, subgroup analysis was done as per the NOS. We used the median NOS score of 8 as a cutoff for grouping the studies. To remove the impact of age and sex, the association of disease-specific and system-specific comorbidities analysis was done for all the comparative studies and for age-/sex-matched control comparative studies. The results across different studies were pooled using the random effects meta-analysis Metaprop module (28), an additional function of Stata software, version 15 (29), and Revman software, version 5.3 (30).

RESULTS

Search results and study qualities. The initial search yielded 70,014 articles from 4 databases. After removal of duplicates, 48,661 remained, of which 1,091 appeared relevant after title screening. Abstract reading confirmed 56 relevant articles and full-text articles that were fully assessed. In all, 42 articles met the inclusion criteria (Figure 1). On the quality assessment scale (maximum of 10) for cross-sectional studies (n = 33), the aver-

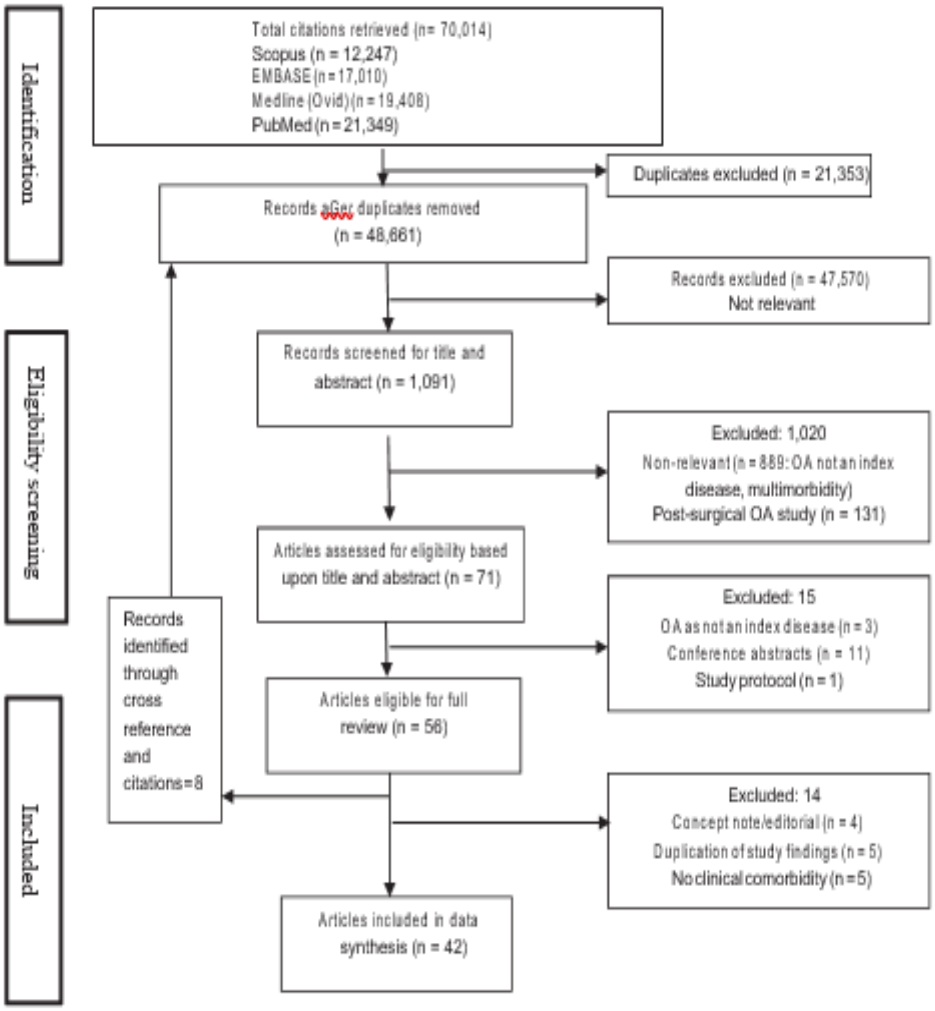


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart of study selection. OA = osteoarthritis.

age score was 5.44 (median 6), and of those, 22 studies had ≥5 stars (31–52). Five case–control and 4 cohort studies had an average score of 5.22 (range 0–9), with 6 studies (6,33,53–56) having ≥5 stars (see Supplementary Table 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.24008/abstract). References 51–92 are in Supplementary Appendix B, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.24008/abstract.

Study characteristics. Of the 42 included studies, 15 compared comorbidities between individuals with OA and those without (comparative studies), whereas 27 examined comorbidities in individuals with OA only (case-only studies). These included 3 case–control studies (6,57,58), 6 cohort studies (53–56,59,60), and 33 cross-sectional studies that explored comorbidity in individuals with OA (1,5,31–52,61–70) (Table 1). We used the baseline comorbidity information from the cohort studies. Thus, we could calculate prevalence only for the comorbidities. Twelve studies were from the US (32,34–40,53,57,63,65), 9 from the Netherlands (1,33,41–44,54,55,59,60), 4 from the UK (6,45,46,66), 2 each from Finland (47,48), Japan (49,56), and Italy (61,67), and 1 each from Canada (50), Hong Kong (5), Spain (68), Australia (51), South Korea (52), Germany (31), Turkey (62), India (69), Brazil (70), Iraq (58), and Latin America (64). Twelve studies were community-based (6,40,45,47,49,51–53,55,56,65,66), 2 were based on national insurance data (32,57), and 28 were hospital-based studies. Eleven studies collected information on knee OA (5,37,49,52,53,56–58,62,68,69), 2 were on hip OA (47,63), 14 were on both knee and hip OA (1,6,33,39,41,42,45,54,55,60,66,67,71), and there was 1 each on ankle (35), hand (43), and hip/knee/hand OA (61). Of 15 comparative studies, 12

Table 1. Characteristics of included studies*

Characteristic	Comparative studies (OA versus non-OA)	Case-only studies (OA only)
Studies	15	27
Study participants	773,592	832,423
Age, mean (range) years	60.1 (50.8–76.1)	63.9 (54.1–74.0)
Women, % (95% CI)	53.0 (35.0–70.0)	63.0 (57.0–69.0)
Body mass index, mean (range)†	27.3 (24.0–29.8) (3 studies)	27.0 (22.0–31.3) (17 studies)
Obesity prevalence, % (95% CI)†	53.4 (42.7–64.1) (7 studies)	31.9 (21.6–42.3) (18 studies)
Comorbidities assessed, median (IQR)‡	6 (4–24) (15 studies)	13 (8–15) (27 studies)
OA site Knee	5	6
Hip	0	2
Ankle	1	0
Both knee and hip	3	11
Hand	0	1
Any joint	5	6
Hand, hip, and knee	0	1
Not given	1	0
Methods of comorbidity measurement		
Charlson Comorbidity Index	0	2
Chronic Illness Rating Scale	1	3
Simple count	10	16
Functional comorbidity assessment	1	0
Self-assessed comorbidity questionnaire	0	1
Three methods	0	1
Not mentioned Study settings	3	4
Community-based	7	5
Hospital-based	6	22
Insurance data	2	0
Methods of OA diagnosis		
Physician diagnosed	6	18
Self-reported	3	1
Radiographic	1	0
Physician diagnosed and radiographic	5	8

* Values are the number unless indicated otherwise. OA = osteoarthritis; 95% CI = 95% confidence interval; IQR = interquartile range.
† Information on the variable was available on the number of studies.
‡ Number of comorbidities assessed in the studies. Information on the variable was available on the number of studies.

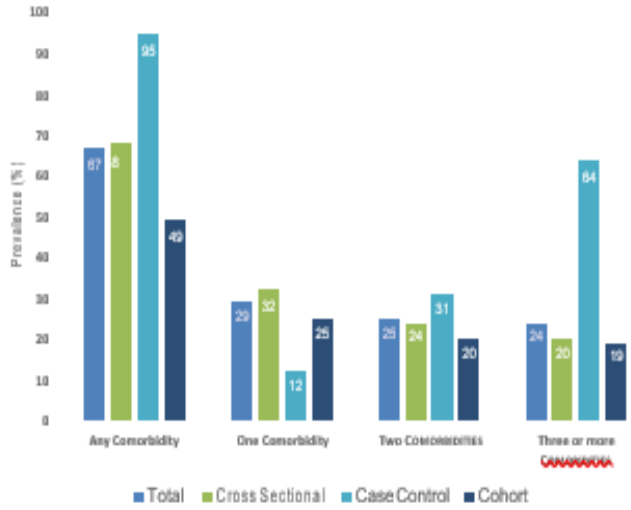


Figure 2. Prevalence of the number of any comorbidities in individuals with osteoarthritis across the study design. Number of studies in each group: Any Comorbidity: total (21), cross-sectional (16), case–control (1), cohort (4); One Comorbidity: total (18), cross-sectional (13), case–control (2), cohort (3); Two Comorbidities: total (16), cross-sectional (12), case–control (2), cohort (2); Three or more Comorbidities: total (14), cross-sectional (10), case–control (2), cohort (2).

had controls minimally matched for age and sex of OA cases. In the included studies, OA was diagnosed in the following ways: clinician assessment without radiographic findings (n = 24), clinical assessment with radiographic diagnosis (n = 13), self-reported physician-given diagnosis (n = 4) (40,45,51,65), and radiographic findings alone (n = 1) (62). Details of the study characteristics are provided in Supplementary Table 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.24008/abstract.

The mean age of the study participants varied from 50.8 years to 76.1 years across studies. The sample size of the included studies ranged from 91 to 237,172 (40,49) and included both men and women, except 1 study that had only women (56). The detailed demographic information (age, sex, body mass index, and obesity) is shown in Table 1.

Prevalence of comorbidities. Of 42 included studies, 15 case-only studies and 8 comparative studies had data on the comorbidity count for analysis. The pooled prevalence of any chronic condition in all studies among individuals with OA was 66% (95% CI 58–74). In OA cases, 29% of participants had a single comorbidity, 25% had 2, and 24% had ≥3. Further subgroup prevalence across the study design is shown in Figure 2. High heterogeneity was observed across studies. Technical details of the data extraction are provided in Supplementary Tables 3 and 4, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.24008/abstract.

The leading systems in terms of pooled prevalence in individuals with OA were cardiovascular (35%), musculoskeletal (34%), neurologic (30%), and upper gastrointestinal (19%). The leading chronic conditions reported among individuals with OA were hypertension (50% [95% CI 36–57]), dyslipidemia (48% [95% CI 14–66]), and back pain (33% [95% CI 11–37]), followed by thyroid disorder (26% [95% CI 6–68]) and depression (17% [95% CI 12–22]). The proportion of chronic conditions was reported to be higher in case–control and cross-sectional studies compared to cohort studies (Figure 3). Details of the prevalence across the study designs are given in Supplementary Tables 5 and 6, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.24008/abstract. All the included studies were cross-sectional in nature except for 2 studies.

Comparison between individuals with and with- out OA. Forest plots for PR and 95% CI between OA and the number of chronic conditions in comparative studies are shown in Figure 4. Eight studies reported the prevalence of comorbidities in individuals with OA and in age- and sex-matched controls, which were used to estimate PR for matched studies (6,34,44,45,49,56,64,65).

The pooled PR for any comorbidity in studies matched for age and sex was 1.21 (95% CI 1.02–1.45; $I^2 = 100\%$; $P < 0.001$) (Figure 4). The PR increased from 0.73 (95% CI 0.43–1.25) for 1 comorbidity to 1.58 (95% CI 1.03–2.42) for 2, and to 1.94 (95% CI 1.45–2.59) for ≥3 comorbidities in OA compared with individuals without OA (Supplementary Figure 1, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.24008/abstract). Subgroup analysis was done for the studies according to the NOS score (Figure 4). Funnel plots for the studies are given in Supplementary Figure 2, available on the *Arthritis Care & Research* web site at http://onlinelibrary.wiley.com/doi/10.1002/acr.24008/abstract, and Egger's test reported nonsignificant publication bias ($P = 0.72$).

The risks for having system-specific comorbidities in age- and sex-matched/adjusted studies among individuals with OA

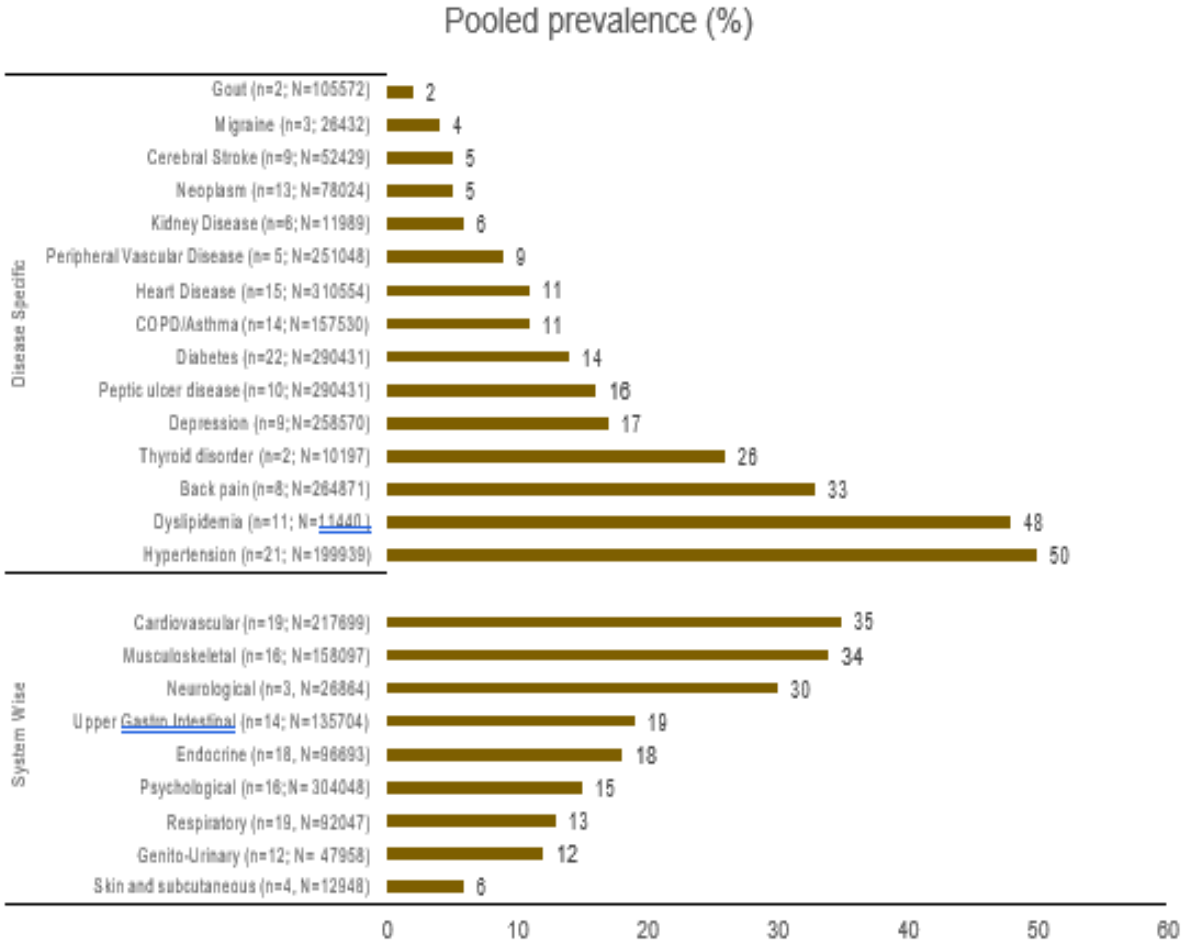


Figure 3. Prevalence (%) of comorbidities in individuals with osteoarthritis (disease and system specific). n = number of studies; N = number of participants; COPD = chronic obstructive pulmonary disease.

were significantly high for upper gastrointestinal disorder (PR 2.36 [95% CI 2.31–2.41]), psychological conditions (PR 1.75 [95% CI 1.20–2.54]), and cardiovascular disease (PR 1.56 [95% CI 1.34–1.86]) compared to individuals without OA. For specific diseases, the risk of stroke was 2.61 (95% CI 2.13–3.21) times higher among individuals with OA compared to those without OA, followed by peptic ulcer (PR 2.36 [95% CI 1.71–3.27]) and metabolic syndrome (PR 1.94 [95% CI 1.21–3.12]) (Table 2).

DISCUSSION

To our knowledge, this is the first systematic review of the literature to examine the evidence of an extensive list of comorbidities in OA. A total of 42 studies from 16 countries were included. The key findings are: 1) 67% of individuals with OA had at least 1 other chronic condition, a level 20% higher than for those without OA, 2) there was a graded effect in terms of the risk of having 1, 2, and ≥3 comorbidities in individuals with OA compared to those without, 3) the systems most likely to be affected by comorbidities in individuals with OA were upper gastrointestinal, psychological, cardiovascular, and endocrine, and 4) stroke, peptic ulcer, and metabolic syndrome were the most common comorbidities in OA.

Studies on multimorbidity from both the developed and developing countries reported OA as a leading chronic condition (14,15,72,73). The risk of having any comorbidities in OA was reported to be 2.35 times higher in the UK general practices population (46), and the risk for multimorbidity was 3 times higher in the Australian population compared to a non-OA group (51). The stronger association of the number of comorbidities in OA indicates the existence of the problem of multimorbidity among these individuals. Besides the number, a pattern of chronic conditions in OA influences management decisions. Comorbidities increase the complexity of care through increased exposure to medication and other chronic conditions. However, the relationship of these factors with the comorbidities is yet to be discovered. The association with multiple chronic conditions requires further research to explore the pattern and causality of comorbidities in OA.

However, the risk for patients with OA of developing comorbidities and the biologic plausibility of such comorbidities is not well investigated. Of the 42 studies included, only 12 primarily

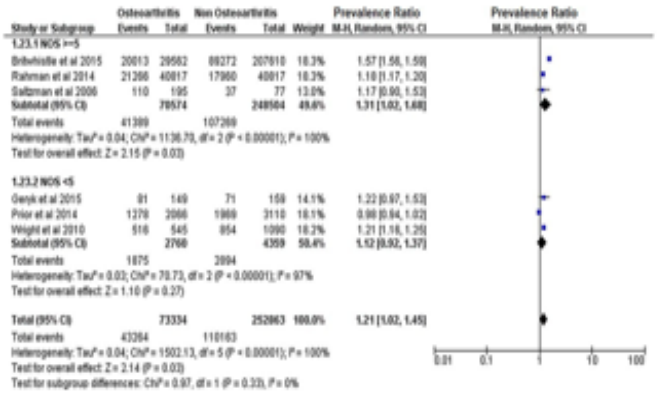


Figure 4. Risk of having comorbidities among individuals with osteoarthritis compared to individuals without osteoarthritis. Only information on systemic comorbidities has been used from the study by Saltzman et al (35) in all estimates. M-H = Mantel-Haenszel; 95% CI = 95% confidence interval; NOS = Newcastle-Ottawa Scale.

examined the comorbidity in OA, 15 had a comparative group, and 27 were published in the years 2010–2017. This summary indicates the quality of the evidence and growing interest in OA comorbidity. Evidence on the risk of having disease-specific comorbidities is not well documented, except for hypertension, diabetes mellitus, and heart diseases, and these comorbidities are further less reported according to the system (61). Few studies are available to explain the association. For example, a meta-analysis done by Wang et al (9) on the association of OA with cardiovascular diseases reported an association with a risk ratio of 1.24, which is less than in our study. Strong associations with other generalized and localized musculoskeletal conditions appeared evident (74,75), but whether coexistence with respiratory diseases was independent or related was considered inconclusive (76), in contrast to our result. According to Parkinson et al (77), individuals with OA are at a 1.41 times higher risk of getting diabetes mellitus. Nearly one-fifth of OA patients have depression (11,78), but previous systematic reviews have been inconclusive about the extent of an association (11). We report risks of 11 comorbidities among patients with OA, which is more comprehensive than any previous study to date.

Exploring factors for comorbidities can be difficult because OA might share different common risk factors with different diseases. The presence of multiple comorbidities could be explained by aging, an important risk factor for OA and other chronic conditions, but we found positive associations in age-matched comparative studies. Associations of OA with gastrointestinal diseases are well documented and usually attributed to long-term use of analgesics, particularly nonsteroidal antiinflammatory drugs (NSAIDs) (79,80). We found nonuniform recording of symptomatic gastrointestinal disorders by the studies, which necessitates correct diagnosis and reporting among patients with OA. The coexistence of cardiovascular comorbidities could be due to shared risk factors such as obesity and metabolic syndrome (81,82). Besides these, NSAIDs and impaired physical activities in OA have been reported to increase risks of developing cardiovascular disease (82–85). Nevertheless, the causal association between OA and cardiovascular disease is not well understood and could in part be attributable to a genetic linkage (86-88). For the association of OA with depression, we hypothesized that the chronicity of the disease, pain, repeated health care utilization, health expenditure, and functional limitation could be the drivers of depression among individuals having OA, and depression can also influence pain experience (78). Endocrine disorders such as hypothyroidism and diabetes mellitus could have an association with OA at specific joint sites (89), but a lack of joint-specific information and endocrine conditions in many studies limits our findings. We did not find fair evidence for musculoskeletal comorbidities in OA, even though we found reports of similar age-related changes in other joints (90) or muscle weakness or injury causing biomechanical derangement leading to pain (91). The increased reporting of back pain and migraine among patients with both symptomatic OA and asymptomatic OA might reflect multiple regional pain resulting from altered pain physiology and central pain mechanisms (92).

Although we estimated the pooled prevalence, it needs careful interpretation owing to the large heterogeneity. However, this is only a systematic review of the current literature, and the purpose of the review is to identify a signal for future research, not to confirm the prevalence and the risk ratio. Prevalence reported in epidemiologic studies is determined by the study design, sample size, case definition, and diagnostic method. The reported prevalence as per the system and disease indicates the existing burden of other chronic conditions in OA, which might affect care. Most of the chronic conditions are age related, and thus understanding their coexistence across the age groups could have been helpful. However, because of the limited articles available, we could not perform such subgroup analysis, and we limited our discussion to the association only. The heterogeneity of the studies and the limited research highlight the need for better-quality comorbidity research in OA.

Table 2. Prevalence ratio of comorbidities associated with osteoarthritis (comparative studies)*

	All studies					Age-/sex- matched studies				
	Studies	Participants (OA)	Participants (non- OA)	PR (95% CI)	I^2, % (P_{val})	Studies	Participants (OA)	Participants (non- OA)	PR (95% CI)	I^2, % (P_{val})
Systems involved										
Upper gastrointestinal	3	127,943	130,021	2.35 (2.31–2.40)†	100 (<0.00001)	2	124,326	124,731	2.36 (2.31–2.41)†	100 (<0.00001)
Psychological condition	4	129,817	139,895	1.67 (1.23–2.29)†	98 (<0.00001)	2	124,326	124,731	1.45 (1.01–2.34)†	99 (<0.00001)
Cardiovascular	9	177,056	342,858	1.57 (1.35–1.82)†	99 (<0.00001)	5	167,274	168,723	1.42 (1.41–1.43)†	100 (<0.00001)
Endocrine	5	56,125	58,496	1.26 (1.14–1.39)†	76 (0.002)	3	52,257	52,662	1.18 (1.13–1.23)†	16 (0.30)
Genitourinary	2	14,992	17,070	1.43 (0.91–2.25)	96 (<0.00001)	1	11,375	11,780	1.14 (1.07–1.22)†	NA
Musculoskeletal	3	124,521	124,826	2.20 (0.83–5.80)	100 (<0.00001)	3	124,521	124,826	2.20 (0.83–5.80)	100 (<0.00001)
Respiratory diseases	3	55,809	57,887	1.11 (1.00–1.24)	85 (0.001)	2	52,192	52,597	1.05 (0.96–1.15)	76 (0.04)
Disease specific										
Stroke	2	5,491	15,164	1.52 (1.30–1.79)†	98 (<0.00001)	1	1,874	9,874	2.61 (2.13–3.21)	NA
Peptic ulcer disease	2	124,326	124,371	2.36 (1.71–3.27)†	88 (0.004)	2	124,326	124,371	2.36 (1.71–3.27)†	88 (0.004)
Metabolic syndrome	2	316	597	1.60 (1.20–2.13)†	1 (0.31)	1	65	65	1.94 (1.21–3.12)†	NA
Peripheral vascular disease	2	124,326	124,731	1.76 (1.04–2.97)†	96 (<0.0001)	2	124,326	124,731	1.76 (1.04–2.97)†	96 (<0.0001)
Depression	4	129,817	139,895	1.94 (1.62–2.32)†	84 (0.0003)	2	124,326	124,731	1.70 (1.29–2.24)†	90 (0.001)
Dyslipidemia	5	120,924	277,277	1.45 (1.15–1.84)†	97 (<0.0001)	2	113,016	113,016	1.57 (1.55–1.58)†	0 (0.93)
Hypertension	8	165,681	331,078	1.76 (1.44–2.17)†	100 (<0.0001)	4	155,899	156,943	1.55 (1.26–2.07)†	100 (<0.0001)
COPD/asthma	3	55,809	57,887	1.45 (1.21–1.74)†	85 (0.001)	2	52,192	52,597	1.35 (1.10–1.66)†	89 (0.003)
Back pain	2	124,326	124,731	1.92 (1.00–3.66)	99 (<0.0001)	2	124,326	124,731	1.92 (1.00–3.66)	99 (<0.0001)
Coronary heart disease	5	131,883	143,005	1.27 (0.69–2.33)	99 (<0.0001)	3	123,692	127,841	0.98 (0.39–2.44)	100 (<0.0001)
Diabetes mellitus	4	44,750	46,716	1.17 (1.13–1.21)†	0 (0.55)	2	40,882	40,882	1.25 (0.87–1.78)	26 (0.25)
Neoplasm	2	14,992	17,070	2.08 (0.47–9.18)	99 (<0.0001)	1	11,375	11,780	0.98 (0.87–1.10)	NA

* Values are the number unless indicated otherwise. OA = osteoarthritis; PR = prevalence ratio; 95% CI = 95% confidence interval; NA = not applicable; COBB = chronic obstructive pulmonary disease.
† $P < 0.05$; P_{val} = P for heterogeneity test.

There are several limitations to this study. First, since multimorbidity/comorbidity in OA is not well indexed in literature databases, we may have omitted some studies. Second, heterogeneity in the prevalence estimates observed in our review, stemming from diversity of methodologies, may have caused uncertainty of the results. Third, there was ambiguity in disease definitions, which creates uncertainty, for example over whether peptic ulcer, gastritis, and acidity should be considered separate entities. Fourth, suboptimal information about OA reported in studies made it difficult to differentiate between structural OA and symptomatic OA and to determine whether associations were linked primarily with structural OA or with pain experience. Similarly, the count of chronic conditions and the definition used varied considerably between studies and may have influenced the estimates (93). Our comparative groups included any non-OA cases, so the comorbidity pattern might have been different because of the selection of comparative/control groups, which needs to be interpreted with caution. Furthermore, the unavailability of joint-specific OA within comparative studies limited the estimation of joint-specific comorbidities. The study also compiles data from different study designs and thus has limitations for understanding the time sequences of OA with comorbidities. Unfortunately, there were not enough studies in each subgroup (only 1 in the cohort design) in comparative studies to perform subgroup analysis as per the study design.

In conclusion, individuals with OA are 1.2 times more likely to have any comorbidity than non-OA controls and 2.5 times more likely to have ≥3 comorbidities. The comorbidities with the highest increase in risk are stroke, peptic ulcer, hypertension, and depression. Further research is needed to determine the causality between OA and these common comorbidities to optimize treatment and develop preventive strategies.

ACKNOWLEDGMENT

The authors thank Soumya Suravita Nanda, who was involved in the initial stage of the protocol preparation.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Mr. Swain had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study conception and design. Swain, Doherty, Zhang.
Acquisition of data. Swain, Sarmanova.
Analysis and interpretation of data. Swain, Coupland, Zhang.

REFERENCES

1. Van Dijk GM, Veenhof C, Lankhorst GJ, Dekker J. Limitations in activities in patients with osteoarthritis of the hip or knee: the relationship with body functions, comorbidity and cognitive functioning. Disabil Rehabil 2009;31:1685–91.
2. Litwic A, Edwards MH, Dennison EM, Cooper C. Epidemiology and burden of osteoarthritis. Br Med Bull 2013;105:185–99.
3. Ryan A, Wallace E, O'Hara P, Smith SM. Multimorbidity and functional decline in community-dwelling adults: a systematic review. Health Qual Life Outcomes 2015;13:168.
4. Fortin M, Lapointe L, Hudon C, Vanasse A, Ntetu AL, Maltais D. Multimorbidity and quality of life in primary care: a systematic review. Health Qual Life Outcomes 2004;2:51.
5. Chan KW, Ngai HY, Ip KK, Lam KH, Lai WW. Co-morbidities of patients with knee osteoarthritis. Hong Kong Med J 2009;15:168–72.
6. Kadam U, Jordan K, Croft P. Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales. Ann Rheum Dis 2004;63:408–14.
7. Centers for Disease Control and Prevention. Comorbidities: what does "comorbidity" mean? 2018. URL: https://www.cdc.gov/arthritis/data_statistics/comorbidities.htm.
8. Hall AJ, Stubbs B, Mamas MA, Myint PK, Smith TO. Association between osteoarthritis and cardiovascular disease: systematic review and meta-analysis. Eur J Prev Cardiol 2016;23:938–46.
9. Wang H, Bai J, He B, Hu X, Liu D. Osteoarthritis and the risk of cardiovascular disease: a meta-analysis of observational studies. Sci Rep 2016;6:39672.
10. Louati K, Vidal C, Berenbaum F, Sellam J. Association between diabetes mellitus and osteoarthritis: systematic literature review and meta-analysis. RMD Open 2015;1:e000077.
11. Stubbs B, Aluko Y, Myint PK, Smith TO. Prevalence of depressive symptoms and anxiety in osteoarthritis: a systematic review and meta-analysis. Age Ageing 2016;45:228–35.
12. Van den Akker M, Buntinx F, Knottnerus JA. Comorbidity or multimorbidity. Eur J Gen Pract 1996;2:65–70.
13. Calders P, Van Ginckel A. Presence of comorbidities and prognosis of clinical symptoms in knee and/or hip osteoarthritis: a systematic review and meta-analysis. Semin Arthritis Rheum 2018;47:805–13.
14. Freund T, Kunz CU, Ose D, Szecsenyi J, Peters-Klimm F. Patterns of multimorbidity in primary care patients at high risk of future hospitalization. Popul Health Manag 2012;15:119–24.
15. Garin N, Koyanagi A, Chatterji S, Tyrovolas S, Olaya B, Leonardi M, et al. Global multimorbidity patterns: a cross-sectional, population-based, multi-country study. J Gerontol A Biol Sci Med Sci 2016;71:205–14.
16. Prados-Torres A, Calderón-Larrañaga A, Hancco-Saavedra J, Poblador-Plou B, van den Akker M. Multimorbidity patterns: a systematic review. J Clin Epidemiol 2014;67:254–66.
17. Ricci-Cabello I, Stevens S, Kontopantelis E, Dalton AR, Griffiths RI, Campbell JL, et al. Impact of the prevalence of concordant and discordant conditions on the quality of diabetes care in family practices in England. Ann Fam Med 2015;13:514–22.
18. European League Against Rheumatism. EULAR. URL: https://www.eular.org/index.cfm.
19. National Institute for Health and Care Excellence. Osteoarthritis: care and management. URL: https://www.nice.org.uk/guidance/cg177/chapter/1-recommendations.
20. Arthritis Research UK. Musculoskeletal conditions and multimorbidities report. 2018. URL: https://www.arthritisresearchuk.org/policy-and-public-affairs/policy-reports/multimorbidity.aspx.
21. Goldberg VM, Buckwalter J, Halpin M, Jiranek W, Mihalko W, Pinzur M, et al. Recommendations of the OARSI FDA Osteoarthritis Devices Working Group. Osteoarthritis Cartilage 2011;19:509–14.
22. American College of Rheumatology. Osteoarthritis. URL: https://www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines/Osteoarthritis.

23. Wells G, Shea B, O'Connell D. Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomized studies in meta-analysis. Ottawa (ON): The Ottawa Hospital. URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. 2007.
24. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004;23:1663–82.
25. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.
26. Van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang MM. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a meta-epidemiological study. BMC Med Res Methodol 2014;14:70.
27. Tamhane AR, Westfall AO, Burkholder GA, Cutter GR. Prevalence odds ratio versus prevalence ratio: choice comes with consequences. Stat Med 2016;35:5730–5.
28. Hunter JE, Schmidt FL. Methods of meta-analysis. Thousand Oaks (CA): SAGE Publications; 2004.
29. Nyaga VN, Arbyn M, Aerts M. Metaprop: a Stata command to perform meta-analysis of binomial data. Arch Public Health 2014;72:39.
30. Review Manager (RevMan). Copenhagen: The Cochrane Collaboration; 2014.
31. Rosemann T, Laux G, Szecsenyi J, Wensing M, Grol R. Pain and osteoarthritis in primary care: factors associated with pain perception in a sample of 1,021 patients. Pain Med 2008;9:903–10.
32. Gore M, Tai KS, Sadosky A, Leslie D, Stacey BR. Clinical comorbidities, treatment patterns, and direct medical costs of patients with osteoarthritis in usual care: a retrospective claims database analysis. J Med Econ 2011;14:497–507.
33. Van Dijk GM, Veenhof C, Schellevis F, Hulsmans H, Bakker JP, Arwert H, et al. Comorbidity, limitations in activities and pain in patients with osteoarthritis of the hip or knee. BMC Musculoskelet Disord 2008;9:95.
34. Singh G, Miller JD, Lee FH, Pettitt D, Russell MW. Prevalence of cardiovascular disease risk factors among US adults with self-reported osteoarthritis: data from the third National Health and Nutrition Examination Survey. Am J Manag Care 2002;8:S383–91.
35. Saltzman CL, Zimmerman MB, O'Rourke M, Brown TD, Buckwalter JA, Johnston R. Impact of comorbidities on the measurement of health in patients with ankle osteoarthritis. J Bone Joint Surg Am 2006;88:2366–72.
36. Dominick KL, Dudley TK, Coffman CJ, Bosworth HB. Comparison of three comorbidity measures for predicting health service use in patients with osteoarthritis. Arthritis Rheum 2005;53:866–72.
37. Inneh IA. The combined influence of sociodemographic, preoperative comorbid and intraoperative factors on longer length of stay after elective primary total knee arthroplasty. J Arthroplasty 2015; 30: 1883–6.
38. Zullig LL, Bosworth HB, Jeffreys AS, Corsino L, Coffman CJ, Oddone EZ, et al. The association of comorbid conditions with patient-reported outcomes in veterans with hip and knee osteoarthritis. Clin Rheumatol 2015;34:1435–41.
39. Gad BV, Higuera CA, Klika AK, Elsharkawy KA, Barsoum WK. Validity of patient-reported comorbidities before total knee and hip arthroplasty in patients older than 65 years. J Arthroplasty 2012;27:1750–6.
40. Ong KL, Wu BJ, Cheung BM, Barter PJ, Rye KA. Arthritis: its prevalence, risk factors, and association with cardiovascular diseases in the United States, 1999 to 2008. Ann Epidemiol 2013;23:80–6.
41. Hoogeboom TJ, den Broeder AA, Swierstra BA, de Bie RA, van den Ende CH. Joint-pain comorbidity, health status, and medication use in hip and knee osteoarthritis: a cross-sectional study. Arthritis Care Res (Hoboken) 2012;64:54–8.
42. Reeuwijk KG, de Rooij M, van Dijk GM, Veenhof C, Steultjens MP, Dekker J. Osteoarthritis of the hip or knee: which coexisting disorders are disabling? Clin Rheumatol 2010;29:739–47.
43. Damman W, Liu R, Kroon FP, Reijnierse M, Huizinga TW, Rosendaal FR, et al. Do Comorbidities play a role in hand osteoarthritis disease burden? Data from the Hand Osteoarthritis in Secondary Care cohort. J Rheumatol 2017;44:1659–66.
44. Nielen MM, van Sijl AM, Peters MJ, Verheij RA, Schellevis FG, Nurmohamed MT. Cardiovascular disease prevalence in patients with inflammatory arthritis, diabetes mellitus and osteoarthritis: a cross-sectional study in primary care. BMC Musculoskelet Disord 2012;13:150.
45. Rahman MM, Kopec JA, Cibere J, Goldsmith CH, Anis AH. The relationship between osteoarthritis and cardiovascular disease in a population health survey: a cross-sectional study. BMJ Open 2013;3:e002624.
46. Kadam UT, Croft PR. Clinical comorbidity in osteoarthritis: associations with physical function in older patients in family practice. J Rheumatol 2007;34:1899–904.
47. Juhakoski R, Tenhonen S, Anttonen T, Kautiainen T, Arokoski JP. Factors affecting self-reported pain and physical function in patients with hip osteoarthritis. Arch Phys Med Rehabil 2008;89:1066–73.
48. Tuominen U, Blom M, Hirvonen J, Seitsalo S, Lehto M, Paavolainen P, et al. The effect of co-morbidities on health-related quality of life in patients placed on the waiting list for total joint replacement. Health Qual Life Outcomes 2007;5:16.
49. Inoue R, Ishibashi Y, Tsuda E, Yamamoto Y, Toh S, Matsuzaka M, et al. Medical problems and risk factors of metabolic syndrome among radiographic knee osteoarthritis patients in the Japanese general population. J Orthop Sci 2011;16:704–9.
50. Birtwhistle R, Morkem R, Peat G, Williamson T, Green ME, Khan S, et al. Prevalence and management of osteoarthritis in primary care: an epidemiologic cohort study from the Canadian Primary Care Sentinel Surveillance Network. CMAJ Open 2015;3:E270–5.