LUND UNIVERSITY

Epidemiology of osteoarthritis in Sweden. Register and cohort studies on prevalence and mortality.

Turkiewicz, Aleksandra

2016

[Link to publication](#)

Total number of authors:
1

Epidemiology of osteoarthritis in Sweden

Register and cohort studies on prevalence and mortality

ALEKSANDRA TURKIEWICZ | FACULTY OF MEDICINE | LUND UNIVERSITY

Epidemiology of osteoarthritis in Sweden

Register and cohort studies on prevalence and mortality

Aleksandra Turkiewicz

LUND
UNIVERSITY

Organization LUND UNIVERSITY	Document name DOCTORAL DISSERTATION
	Date of disputation: May 12, 2016
Author(s) Aleksandra Turkiewicz	Sponsoring organization

Title and subtitle
Epidemiology of osteoarthritis in Sweden. Register and cohort studies on prevalence and mortality.

Abstract

Osteoarthritis (OA) is the most common form of arthritis and often causes pain and functional impairment. Despite its high burden on society, knowledge about the occurrence of OA and its associated mortality is sparse. The epidemiology of OA is difficult to determine due to the patients' fluctuating symptoms and challenges in detecting and defining the radiographic disease. The specific aims of this thesis were to estimate the current and future prevalence of OA in peripheral joints in Sweden and the mortality associated with OA using the population-based register data and data from clinical examinations.

Among the residents of the Skåne region aged ≥45 years in the year 2012, persons who have consulted a physician for OA at least once were identified using the Skåne Healthcare Register (SHR) – a mandatory register containing information on all healthcare visits made in the region. Prevalence of doctor-diagnosed OA (current and expected in the next 20 years) and all-cause mortality in persons with knee or hip OA were estimated. A subset of Skåne residents aged 56 to 84 years who participated in the Malmö Osteoarthritis study (MOA) (n=1527) underwent clinical examination including knee x-ray and knee pain questionnaire in order to assess the presence of knee OA.

Of adults aged 45 years or older and living in Skåne (n=531 254), 27% consulted a physician for OA in a peripheral joint at least once between 1999 and 2012 and this proportion is expected to increase to 30% by the year 2032. The 2012 prevalence of the doctor-diagnosed knee OA was 13.8%, of the hip OA – 5.8%, of the hand OA – 3.1% and of the OA in other joints – 12.4%. One in four of persons with the doctor-diagnosed OA consulted for the disease in more than one joint site. In Skåne, the hazard ratios of death in persons with doctor-diagnosed knee or hip OA compared to the general population were 0.87 (95% confidence interval 0.85, 0.89) and 0.90 (0.87, 0.92), respectively. In the MOA study, the prevalence of knee OA giving symptoms (with or without radiographic manifestations) was 15.4%. Of these, 53% consulted healthcare and received a knee OA diagnosis during an 8 year period. A number of sensitivity analyses, including the validation of knee OA diagnoses registered in SHR, quantification of bias and a simulation study assessing the impact of missing diagnostic codes on the derived prevalence estimates, confirmed robustness of the results.

This thesis provides a comprehensive view of clinically relevant OA as well as radiographic findings (knee) with or without symptoms in Sweden. Results will aid the healthcare planning and evaluations of the disease burden. The findings also indicate that knee or hip OA are not associated with excess mortality on the population level; possible mechanisms include relatively easy access to care and treatments or maintaining physical activity.

Key words: Epidemiology, Osteoarthritis, Knee, Hip, Radiography, Mortality, Registers, Missing data
Classification system and/or index terms (if any)

Supplementary bibliographical information	Language: English

ISSN and key title 1652-8220	ISBN 978-91-7619-273-3
Recipient's notes	Number of pages 70	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature *Aleksandra Turkiewicz*   Date 2016-04-04

Epidemiology of osteoarthritis in Sweden

Register and cohort studies on prevalence and mortality

Aleksandra Turkiewicz

LUND UNIVERSITY

For Dominik

Contents

List of papers

I.   Turkiewicz A, Petersson IF, Björk J, Hawker G, Dahlberg LE, Lohmander LS, Englund M. Current and future impact of osteoarthritis on health care: a population-based study with projections to year 2032. *Osteoarthritis Cartilage*. 2014 Nov;22(11):1826-32.

II.   Turkiewicz A, Gerhardsson de Verdier M, Engström G, Nilsson PM, Mellström C, Lohmander LS, Englund M. Prevalence of knee pain and knee OA in southern Sweden and the proportion that seeks medical care. *Rheumatology (Oxford)*. 2015 May;54(5):827-35.

III.   Turkiewicz A, Neogi T, Björk J, Peat G, Englund M. All-cause mortality in knee and hip osteoarthritis and rheumatoid arthritis. *Epidemiology*, 2016 Mar 16. [Epub ahead of print]

IV.   Turkiewicz A, Englund M, Björk J. Multiple imputation of diagnostic codes in a longitudinal healthcare database – a simulation study. *In manuscript*.

Abbreviations

ACR	American College of Rheumatology
BMI	Body Mass Index
CI	Confidence Interval
HR	Hazard Ratio
ICD	International Classification of Diseases
KL	Kellgren – Lawrence
LISA	Longitudinal integration database for health insurance and labour market studies
MAR	Missing At Random
MCAR	Missing Completely At Random
MDCS	Malmö Diet and Cancer Study
MI	Multiple Imputation
MNAR	Missing Not At Random
MOA	Malmö Osteoarthritis Study
MRI	Magnetic Resonance Imaging
OA	Osteoarthritis
RA	Rheumatoid Arthritis
SSATG	South Swedish Arthritis Treatment Group
SCB	Statistics Sweden
SHR	Skåne Healthcare Register
UK	United Kindom
US	United States

Abstract

Osteoarthritis (OA) is the most common form of arthritis and often causes pain and functional impairment. Despite its high burden on society, knowledge about the occurrence of OA and its associated mortality is sparse. The epidemiology of OA is difficult to determine due to the patients' fluctuating symptoms and challenges in detecting and defining the radiographic disease. The specific aims of this thesis were to estimate the current and future prevalence of OA in peripheral joints in Sweden and the mortality associated with OA using the population-based register data and data from clinical examinations.

Among the residents of the Skåne region aged ≥45 years in the year 2012, persons who have consulted a physician for OA at least once were identified using the Skåne Healthcare Register (SHR) – a mandatory register containing information on all healthcare visits made in the region. Prevalence of doctor-diagnosed OA (current and expected in the next 20 years) and all-cause mortality in persons with knee or hip OA were estimated. A subset of Skåne residents aged 56 to 84 years who participated in the Malmö Osteoarthritis study (MOA) (n=1527) underwent clinical examination including knee x-ray and knee pain questionnaire in order to assess the presence of knee OA.

Of adults aged 45 years or older and living in Skåne (n=531 254), 27% consulted a physician for OA in a peripheral joint at least once between 1999 and 2012 and this proportion is expected to increase to 30% by the year 2032. The 2012 prevalence of the doctor-diagnosed knee OA was 13.8%, of the hip OA – 5.8%, of the hand OA – 3.1% and of the OA in other joints – 12.4%. One in four of persons with the doctor-diagnosed OA consulted for the disease in more than one joint site. In Skåne, the hazard ratios of death in persons with doctor-diagnosed knee or hip OA compared to the general population were 0.87 (95% confidence interval 0.85, 0.89) and 0.90 (0.87, 0.92), respectively. In the MOA study, the prevalence of knee OA giving symptoms (with or without radiographic manifestations) was 15.4%. Of these, 53% consulted healthcare and received a knee OA diagnosis during an 8 year period. A number of sensitivity analyses, including the validation of knee OA diagnoses registered in SHR, quantification of bias and a simulation study assessing the impact of missing diagnostic codes on the derived prevalence estimates, confirmed robustness of the results.

This thesis provides a comprehensive view of clinically relevant OA as well as radiographic findings (knee) with or without symptoms in Sweden. Results will aid the healthcare planning and evaluations of the disease burden. The findings also indicate that knee and hip OA are not associated with excess mortality on the population level; possible mechanisms include relatively easy access to care and treatments or maintaining physical activity.

Introduction

Epidemiology is the study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to manage health problems.[1] The key paradigm of epidemiology is that disease patterns in populations may be analyzed systematically to understand their causes and to improve health. Epidemiology has been useful in yielding understanding of what causes disease in population, preventing or controlling disease, guiding healthcare policy and planning or management of diseases in individuals.[2] The surge of epidemiologic activity since the late 20th century has been accompanied by the evolution of epidemiologic concepts and methods in order to improve their validity.[3] In this thesis I attempt to study the occurrence and consequences of osteoarthritis in Sweden and to use epidemiological methods that may improve the reliability of the obtained results.

Osteoarthritis

Osteoarthritis (OA) is a progressive degenerative joint disease that affects all structures of the joint. OA can affect all joints, but the most common sites are the knee, the hip and the hand. OA has a tremendous impact on the lives of working age adults and elderly causing pain and functional impairment[4] and being the leading cause for joint replacement surgery.[5,6] The number of years lived with disability due to knee and hip OA alone increased by 64% between 1990 and 2010. OA is currently ranked eleventh in the world among the leading causes of disability.[7]

OA is a heterogeneous disease and is challenging to define, as it may involve pathological processes on the molecular level, structural changes of the joint tissues and clinical symptoms in patients. The consensus definition is as follows:

> "OA diseases are a result of both mechanical and biologic events that destabilize the normal coupling of degradation and synthesis of articular cartilage chondrocytes and extracellular matrix, and subchondral bone. Although they may be initiated by multiple factors, including genetic, developmental, metabolic and traumatic, OA diseases involve all of the tissues of the joint. Ultimately, OA diseases are manifested by morphologic, biochemical, molecular, and biomechanical changes of both cells and matrix which lead to a softening, fibrillation, ulceration, loss of articular cartilage, sclerosis and eburnation of subchondral bone, osteophytes, and subchondral cysts.

When clinically evident, OA diseases are characterized by joint pain, tenderness, limitation of movement, crepitus, occasional effusion, and variable degrees of inflammation without systemic effects."[8]

As reflected in the above definition, the concept of OA is still not fully understood and thus the research about OA is difficult and requires a number of tolls, definitions or criteria depending on the scientific question. In epidemiological studies on prevalence, risk factors for or consequences of the disease the common solution is the use of established classification criteria for defining the OA disease.

Classification criteria and diagnosis of OA

In brief, classification criteria are standardized definitions that are primarily intended to enable clinical studies to have uniform cohorts for research, which are able to define a homogenous group and where high specificity is required. In contrast, a diagnosis is the outcome of a clinical process aimed at confirming the presence (or absence) of a health disorder in the context of an encounter between a physician and a patient, in a particular healthcare setting.[9] This process requires the physician's skills, knowledge and intuition and is based on a set of signs, symptoms and tests developed for use in routine clinical work (diagnostic criteria). The usual aim is to guide the care of individual patients. Diagnostic criteria need to reflect all the possible features and severity of the disease. When the aetiology of a disease is well defined, diagnostic and classification criteria may be very similar and used interchangeably.[10] When the aetiology of the disease is not fully understood and the symptoms are not highly specific, as is the case with OA, various classification criteria and the clinical diagnosis may yield weakly overlapping groups.

The gold standard classification criteria in epidemiology of OA are based on radiographic imaging. The most commonly used is a system developed by Kellgren and Lawrence (KL), where the joint features, primarily osteophytes (outgrowth of bone) and joint space narrowing, are graded on a point 0 to 4 scale.[11] Traditionally a grade ≥2 is considered to represent radiographic OA. Radiographic OA in combination with self-reported joint pain on most days of a month is often classified as symptomatic OA. Another set of classification criteria for OA of the knee, the hip and the hand has been developed by the American College of Rheumatology (ACR).[12–14] These criteria are based on patients history, physical examination and radiographic or laboratory findings, if available. In recent years new criteria based on the findings visible on the magnetic resonance imaging (MRI), such as cartilage loss, bone marrow lesions or meniscal tear have been proposed.[15] However, defining early OA on MRI is challenging and no golden standard have been established yet.[16,17]

There are no uniform international diagnostic criteria for OA. In Sweden, the National Board of Health and Welfare's guidelines published in 2012 recommend that the diagnosis of OA is based on the anamnesis, symptoms and clinical findings

without routine use of the radiographic imaging.[18] Nevertheless, the diagnosis is based on the subjective judgment of the treating physician.

Prevalence of OA

OA is the most common form of arthritis. The prevalence of OA (i.e. the proportion of the population having the disease at a given point in time) has been studied extensively in a number of western countries. However, reliable data are lacking for large parts of the world, which has become even more evident when the latest WHOs burden of disease study has attempted to estimate the prevalence of knee and hip OA worldwide and found reports from only 27 countries and rarely covering the whole relevant age spectrum.[19]

Challenges in estimating the prevalence of OA are related to the various definitions that can be used, such as the radiographic criteria combined with reports of symptoms, the ACR criteria or a diagnosis from a health professional. The common methods of the data collection include subjects' self-report/survey data, clinical and radiographic examination or register data. The prevalence studies have focused most often on the radiographic or symptomatic OA or on the joint pain, that in persons aged 55 years or older is often attributed to OA.[20] However, the overlap between the structural changes as in OA visible on plain radiograph and clinical symptoms as in OA is usually low, and is expected to have a poor overlap with healthcare consultation for OA (Figure 1).[21–24] The prevalence of self-reported OA or OA leading to a healthcare consultation is rarely reported together with findings concerning radiographic evidence and symptoms and thus it is challenging to acquire a broader picture of the occurrence of OA in societies.[25,26]
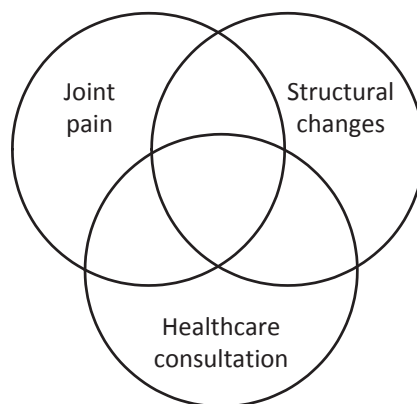

Figure 1. Selected concepts of OA.

The prevalence of radiographic OA in people ≥45 years ranges from 19.2% to 27.8% in the knee and from 19.6% to 27.0% in the hip. Typically, the prevalence of the symptomatic disease is lower with reported estimates from 6.7% to 15.9% in the knee and from 1.6% to 9.2% in the hip.[23,24,27–30] Radiographic hand OA is highly prevalent, especially in women, and present in up to 80% of older adults, but only a minority of persons with radiographic hand OA have pain in these joints.[27,31] There are only a few studies reporting prevalence of OA for all joints combined. The 2001 prevalence of OA in at least one joint in Canada, based on data from a healthcare register for persons aged 18 or older, was 14.3%,[25] while in Australia almost 15% of persons 15 years old or older reported having OA.[32] In the United States (US), in 2005 the estimated prevalence of symptomatic OA was 15.3% in persons aged 25 to 74.[27]

In Sweden, the prevalence of joint complaints in knees in persons aged 79 in mid 1980s was 25% in women and 11% in men,[33] while the prevalence of radiographic OA in right knees was 53% in women and 43% in men.[34] Forty one per cent of persons reported joint complaints and had radiographic OA of the hands. More recently, the prevalence of symptomatic knee OA in persons aged 35 to 54 was estimated to be 2.4%.[35] Up-to-date population-based data for a wider age range are lacking, as well as estimates of OA in other joints.

The prevalence of OA is expected to increase due to the anticipated changes in the age-structure of the population and the obesity epidemic.[36,37] In the US, the prevalence of self-reported doctor-diagnosed arthritis is projected to increase from 22% in 2003 to 25% in 2030.[38] Possible changes in the disease occurrence may require changes in planning and resource prioritization within healthcare and more information on the future burden of OA on healthcare is warranted.

Risk factors for OA

Systemic factors

The pathogenesis of OA is likely multifactorial. Genetic factors may account for at least 50% of cases of OA in the hands and hips and smaller percentage in the knees.[39] OA is more common in women than men and in women the prevalence increases dramatically after menopause. Thus, hormonal status has been suggested to play a role, but the evidence on a protective effect of oestrogen is weak. Nutritional factors have been proposed, where vitamin D has been reported to protect against both incident and progressive hip OA.[40] The prevalence of OA increases with increasing age in all joints. The incidence of radiographic or symptomatic OA is rare in persons under 45 years of age and typically peaks at ages 70 to 79.

Local biomechanical factors

Being overweight is a risk factor for both incidence and progression of OA, especially in the knee, and weight loss can reduce the risk of OA.[40] Joint dysplasia, or joint

injuries, e.g., fractures of articular surfaces and tears of menisci and ligaments that increase instability of the joint precede the development of OA in a high percentage of affected joints.[41,42] Muscle weakness in the quadriceps is a known consequence of knee OA, but has also been reported to be a risk factor for radiographic or symptomatic knee OA.[40,43]

Treatment of OA

There is lack of disease-modifying pharmacological treatments for OA. Available treatments are focused on alleviating symptoms and maintaining physical fitness. The core treatments, appropriate for all individuals, are education to enhance understanding of the condition and its management as well as, for knee and hip OA, non-pharmacological treatments such as exercise and manual therapy, weight loss and bracing/joint support when needed. Pharmacological management of joint pain includes the use of paracetamol and non-steroid anti-inflammatory drugs or intra-articular corticosteroid injections for moderate to severe pain.[44,45] A joint replacement surgery is the treatment of the end-stage OA and may be offered primarily to persons with knee or hip OA that do not obtain an adequate pain relief and functional improvement from a combination of non-pharmacological and pharmacological treatments.[46] It is estimated that 10-20% of persons with knee OA require eventually a joint replacement surgery.

Rheumatoid arthritis

In contrast to OA, rheumatoid arthritis (RA) is an autoimmune chronic inflammatory disease with a pronounced systemic involvement. RA affects synovium, leading to joint damage and bone destruction that in turn cause severe disability. Pain, swelling and redness are common manifestations and multiple joints are often affected. Persons with RA have decreased quality of life, restrictions in daily living and often a reduced work capacity.[47,48] Studying RA has not been the primary aim of this thesis, but I have included it in one study in order to gain information about the internal validity of the used data sources.

The diagnosis of RA relies on clinical judgement of the treating physician, but the ACR classification criteria for RA[49] are well known and often applied by rheumatologists in the routine care. Thus, the overlap between RA diagnosed within healthcare and RA identified using the ACR criteria is expected to be high. The prevalence of doctor-diagnosed RA in Sweden in 2008 was 0.66%, 0.37% in men and 0.94% in women.[50] The main risk factors for the disease include genes and smoking, but the aetiology is unknown.[51] Early recognition and treatment with disease-modifying anti-rheumatic drugs is essential in achieving control of the

disease.[52] Since the late 1990's its treatment has been much improved by the introduction of new disease modifying anti-rheumatic drugs (biological treatment), e.g., TNF-blockers.

Mortality in OA and RA

Prevention of premature death may be viewed as a primary goal of medical care and mortality has often been an extensively studied outcome in epidemiology. However, rheumatic diseases have not been commonly perceived as 'fatal'. The scientific evidence on mortality in rheumatic diseases is very modest when compared to cardiovascular diseases or cancer.[53] Increased mortality rates have been reported for a number of rheumatic diseases, such as gout, systemic lupus erythematosus, psoriatic arthritis, ankylosing spondylitis and others, including OA and RA, but the strength of evidence and the number of studies are generally limited.[53]

One exception is RA. There is solid evidence of increased mortality in patients with RA compared to the general population with the reported hazard ratios in the range of 1.6-1.7.[54] The inflammatory process (and its treatments) and higher levels of infectious, pulmonary, cardiovascular and renal comorbidities contribute.[54]

The relationship between OA and mortality is not clear. There is a rationale for a possible association of OA with excess mortality. The major mediating factors considered are immobility due to persistent symptoms, weight increase and cardiovascular diseases.[55,56] However, the evidence in the literature is not consistent.[57,58] The comprehensive systematic review by Hochberg in 2008 concluded in a best evidence synthesis of 6 studies that there was moderate evidence of somewhat increased all-cause mortality among persons with OA compared with the general population, particularly for knee OA and generalized OA.[57] Since then, Nuesch et al. have reported that persons with symptomatic knee or hip OA had a standardized mortality ratio of 1.55 as compared to the general population.[55] Other investigators have reported lower mortality rates for OA as compared with other chronic conditions or,[58] for hand OA, with the general population.[59] Several of the studies suffered from potential bias, misclassification or had short follow-up times. In addition, variation in OA management between different countries and health care systems, underlying health of the populations or life style factors may have impact on all-cause mortality because mortality in OA has been suggested to be related to the actual physical disability rather than biological consequences of the disease.

The Skåne region

The research of this thesis is based within the Skåne region. Skåne (Scania) is the southernmost region of Sweden with an area of 11 000 km^2 and 1.3 million inhabitants (Dec 2014) (Figure 2).
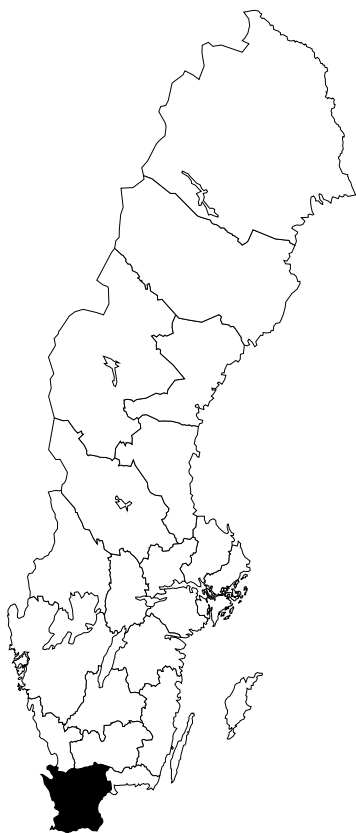
Figure 2. The Scania region (in black) in Sweden.

The region contains a large proportion of Sweden's best farm land and is strong within farming industry which employs 13% of the region's population. However, the majority of inhabitants are working within service, where healthcare is one of the major branches. One in three persons between 25 and 64 years of age has higher education. The largest city is Malmö with a population of nearly 320 000. The Skåne region is characterized by true variety of landscape, from modern urban environment in Malmö, through industry regions to small towns and nature reserves. Its sociodemographic structure is similar to that of the whole Sweden (Table 1).

Table 1.

Demographics of the Skåne and Sweden populations aged 45 years or older*.

	Skåne, N=551 592	Sweden, N=4 254 234
Age in years, mean	63.7	63.6
% women	52	52
Education, %		
<9 years	17	17
9-11 years	41	42
12-14 years	25	25
15+ years	17	17
Annual income, median in SEK†	241 525	251 557
Overweight††, %	42	41
Obese††, %	18	15

*Data source: Statistics Sweden per year 2012; †Swedish crown, euqals 0.1 Euro; ††Sweden in 2014 ages 45+ years, Skåne in 2012 based on Folkhälsoenkät Skåne 2012, ages 45-80 years,[60] overweight defined as body mass index≥25, obese defined as body mass index≥30

Epidemiological cohorts in Skåne

The region has been a suitable location for large scale epidemiological studies due to the presence of the Lund University and the Skåne University Hospital. One example is the Malmö Diet and Cancer study (MDCS). The study was designed to answer questions considering the associations of dietary factors and certain cancers.[61] All men and women living in Malmö who were born between 1923 and 1945 (men) or 1923 and 1950 (women) were invited by letters and advertisements in newspapers to take part in the study. Between March 1991 and September 1996, the respondents participated in clinical examinations at the screening centre and filled in a self-administered questionnaire.[62] There were slightly over 28 000 participants among 70 000 eligible persons. They had similar socio-demographic structure to the non-participants, but the mortality was higher in the non-participants both during and following the recruitment period.[63] Over the course of the years the MDCS has become a base for recruitment of participants to other epidemiological studies, including the Malmö OA (MOA) study used in this thesis. The aims of the MOA study were to characterize and quantify different categories of knee OA with respect to symptoms and findings from physical and radiographic examinations in a population-based cohort, including persons that have previously not sought medical care. Thus, among the MDCS responders, who were still alive and residing in Malmö in 2007, a random sample of 10 000 individuals was selected to participate in the MOA study. The invited persons answered a postal questionnaire concerning the knee pain, and a subset of responders underwent a clinical examination including x-ray of both knees and knee symptoms assessment.

Registers in research

In recent years epidemiologic research has started to utilize the data from the existing longitudinal population-based registers extensively. In Sweden, social, demographic and health data have been collected for many years and have been available for research after obtaining sufficient ethical approvals and after complying with the secrecy and confidentiality laws. Using the unique personal identification number assigned to every person living in Sweden, these data can be linked to regional registers as well as ad-hoc research cohorts.[64,65]

Utilization of existing population-based register data has several strengths. It is usually cost-effective, maximizes the public value of the already collected data and enables working with large cohorts. The data collection process, which is independent of the study question, and inclusion of the whole population of a particular geographical region result in minimizing the selection bias and often prevent the differential misclassification of the exposure, confounders and outcome.[66]

However, register data have in general important limitations that need to be recognized. Data quality is defined by the register and not by the researcher and may be unknown with respect to validity or completeness. Information on important confounders is often not available or the data have high percentage of missing values. Data are registered using definitions that are relevant for the 'administrator', which may not be the same as those relevant for the research question, and the coding practice may vary between geographical areas or year of data collection. The data are usually truncated by the start of the registration period which may be a source of misclassification. The existence and completeness of register metadata (description of data and variables, methodology and quality) are crucial for an informed usage of registers for research.[67]

This thesis uses register data from several sources. The Swedish population register is maintained by the Tax Agency and it is where the population of Sweden is registered.[68] It is a source of information for many other population-based registers and for a number of official authorities. The Longitudinal integration database for health insurance and labour market studies (LISA by Swedish acronym) is maintained by Statistics Sweden (SCB) - the provider of the official statistics in Sweden. It contains, among others, the official data related to education, employment and income that are continuously updated with the best available information.[69] The metadata for these registers is provided by the authorities and is extensive.

The data concerning the healthcare utilization in the Skåne region are stored in the Skåne Healthcare Register (SHR). The healthcare in the region is provided on primary, specialist and in-patient level, including highly specialized care given at the Skåne University Hospital. The out-patient care is given by both private and public providers and financed through a common tax-based system. The patients' co-pay is relatively low and similar irrespective of the provider. It is common that patients seek care from, or are referred to, both types of providers. The information on every

healthcare contact made in the region from the year 1998 and onwards is registered in the SHR. The SHR contains information on the healthcare provider, the profession (physician, physical therapist, nurse etc.), type of contact (e.g., primary/ specialist care, in- or out-patient visit, clinic etc.), contact date and patients' unique personal identification number. Furthermore, the register contains the publicly practicing physicians' diagnostic codes according to the International Classification of Diseases (ICD) 10 system. These codes are assigned by the doctors themselves and are retrieved from the electronic medical records to the register. The codes from the private healthcare providers, responsible for around 30% of all doctor visits, had not been forwarded to the SHR, but a substantial part of the providers started to register their diagnosis in the year 2013. The registered data have been used, among others, for reimbursement purposes. The medical diagnoses registered in the SHR have been validated against medical journals or other data sources for a number of diseases and showed generally high positive predictive values or sensitivity.[50,70–72] The completeness of the registered administrative data (such as the provider, the clinic, type of contact etc. or patients' unique personal identification number) has been nearly 100%. However, the completeness of the medical diagnoses in the public out-patient care has been increasing gradually over the years, from only 15% in the year 1998 to over 90% in the year 2009. The data from SHR are provided to the researchers by Region Skåne after obtaining required permission and approval from the ethical review board.

Aims

General aim

To increase knowledge concerning OA epidemiology in Sweden and epidemiological methods that can be used in research based on routinely collected healthcare data.

Specific aims

- To estimate the prevalence of OA in the peripheral joints that leads to a doctor consultation, today and in 20 years, and to validate the diagnosis of knee OA registered in SHR with respect to a gold standard epidemiological definition. (paper I)
- To estimate the prevalence of knee pain, radiographic, symptomatic and clinically defined knee OA and what proportion of persons with knee OA consults healthcare. (paper II)
- To assess the mortality associated with doctor-diagnosed OA of the knee or hip as compared to the general population consulting healthcare. (paper III)
- To evaluate multiple imputation as a method for handling missing diagnostic codes in a longitudinal healthcare database in order to estimate the prevalence of a chronic disease in the population. (paper IV)

Methods

Data sources

The population register

The Swedish population register contains information about vital events such as births, deaths, and changes in residential addresses for all inhabitants in Sweden and is based on the unique personal identification number. This number is assigned to every person registered as inhabitant in Sweden (i.e., not only for persons with Swedish citizenship, but all persons living in Sweden) and is used for identification purposes. The number contains information on the date of birth and sex. The population register is continuously updated by the Swedish Tax Agency. I have used the register to identify persons for inclusion in the papers I, II and III. (Table 2)

The Skåne Healthcare Register

The SHR contains information about every healthcare contact made in the Skåne region from the year 1998 and onwards. The register includes individual level information including the unique personal identification number, which enables linking of the register data with the population register. For each visit to a physician the data on the healthcare provider, the profession, type of contact (e.g., primary/ specialist care, in- or out-patient visit etc.), the clinic and the contact date are registered. I have used the SHR data to identify persons that consulted healthcare due to any reason (paper III), or for a particular disease (papers I, II and III) and as a base for data simulation in paper IV (Table 2).

Table 2.

The study cohorts and OA/RA definitions used in papers I-IV.

Paper	Inclusion criteria	Age range	N	OA/RA definitions used
I	Residents in Skåne at 2012-12-31	45+	531 254	Doctor-diagnosed OA, Clinical and radiographic knee OA (ACR*) Radiographic knee OA
II	Residents in Malmö in 2007, a random sample from the MDCS cohort†	56-84	10 000	Radiographic knee OA, Symptomatic knee OA, Clinical knee OA (ACR*), Doctor-diagnosed knee OA
III	Residents in Skåne that consulted a physician at least once between 1998-01-01 and 2012-12-31	45+	524 136	Doctor-diagnosed knee or hip OA Doctor-diagnosed RA
IV	Simulated artificial population	NA	1000	Doctor-diagnosed disease

*ACR – American College of Rheumatology, †MDCS – Malmö Diet and Cancer Study

Statistics Sweden

SCB is an administrative agency with the main task to coordinate the Swedish system for official statistics. The individual level data from the LISA databases, such as income, education and marital status, including the unique personal identification number that enables linkage with the other register data has been provided by SCB for the whole Skåne population. Additionally, I have retrieved the aggregated level longitudinal data on the prevalence of overweight and obesity in Sweden and the age and sex specific population projection for Sweden until year 2032 from SCBs publications. The data from SCB were used in papers I and III.

The MOA study

The MOA study was carried out between 2007 and 2008. The first part of the study consisted of a knee pain questionnaire sent to a random sample of 10 000 participants aged 56 to 84 years from the MDCS cohort who were still alive and resident in the Malmö area at the beginning of 2007.(Figure 4)
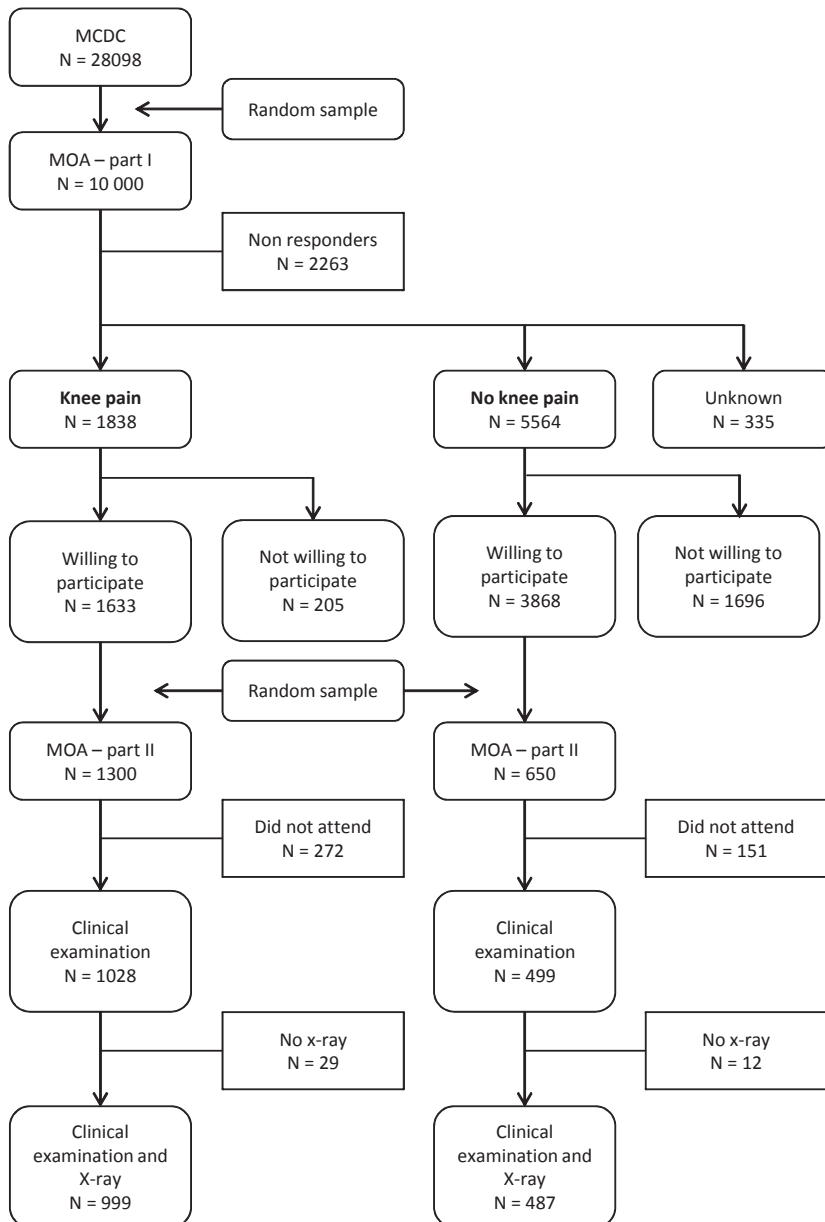
Figure 4. The flow-chart of MOA study sample.

Respondents answered a question about whether they had knee pain during the previous 12 months and its duration (<1 week, 1-4 weeks, 1-3 months, >3 months). Persons with pain in one or both knees in the past 12 months and duration of more than 4 weeks were classified as having frequent knee pain. In the second part of the study a random sample of 1300 persons with frequent knee pain and 650 persons without were invited to a clinical visit and radiographic examination. At the clinical visit the trained study nurse measured weight and height and the body mass index (BMI) was calculated. The participants were asked if they have had previous knee arthroplasty and completed a questionnaire assessing, among other things, pain in the whole body and its location. Data on birth date, sex, education and BMI from the MDCS examination in 1991-96 were available for the whole 10 000 study sample (Table 3). I have used the data from the MOA study in papers I and II.

Table 3.
Demographics of the MOA study sample (N=10 000).

Age in years, mean (SD), range	70 (7.6), 56-84
% women	62
Education, %	
≤9 years	40
10-12 years	36
13-14 years	9
15+ years	15
Overweight*, %	40
Obese*, %	14

*At the MDCS examination, overweight defined as BMI≥25, obese defined as BMI ≥30

Ethical considerations

Relevant ethical approvals from the ethical review board in Lund were obtained for studies included in this thesis. An informed consent was obtained from all participants of the MOA study in accordance with the Declaration of Helsinki. Data were processed using the *encrypted* unique personal identification number without access to the personal information (such as name, residential address or unique personal identification number) of the study participants. All results were presented on the aggregated level preventing any individual from being identified.

Definitions of OA and RA

Doctor-diagnosed OA

The main definition of OA (in knee, hip, hand and other joints) used in this thesis is the *doctor-diagnosed OA*. A person (aged 45 years or older) was defined as an OA case if having a diagnosis of OA registered in SHR, at a healthcare visit to a physician, at least once (Table 4). I assumed that a person has OA from the earliest diagnosis date encountered in the register and onwards (i.e., one cannot be cured from OA). I classified the OA locations into knee, hip, hand and other (including elbow, shoulder, jaw and foot), and I referred to OA in at least one peripheral joint as 'any OA'.

Table 4.

The ICD 10 system of OA.

Location	ICD 10 code	Diagnosis
Knee	M17	Gonarthrosis [arthrosis of knee]
Hip	M16	Coxarthrosis [arthrosis of hip]
Hand/wrist	M18	Arthrosis of first carpometacarpal joint
	M15.1	Heberden's nodes (with arthropathy)
	M15.2	Bouchard's nodes (with arthropathy)
	M19.0D	Primary arthrosis of other joints, site: wrist/hand
	M19.1D	Post-traumatic arthrosis of other joints, site: wrist/hand
	M19.2D	Other secondary arthrosis, site: forearm
Other	M15 (other than M15.1, M15.2)	Polyarthrosis (excluding Heberden's/Bouchard's nodes)
	M19(other than M19.0D, M19.1D, M19.2D)	Other arthrosis (excluding arthrosis in hand/wrist)

Knee OA according to ACR criteria

Two sets of ACR classification criteria could be applied to determine the presence or absence of knee OA in the participants of the MOA study. The *clinically defined knee OA* according to the recursive positioning method was defined as knee pain for most days of the prior month in combination with one of the three groups of symptoms: 1) crepitus on active joint motion and morning stiffness <30 minutes in duration and age > 38 years; 2) crepitus on active joint motion and bony enlargement of the knee on examination; 3) bony enlargement of the knee on examination.

The *clinical and radiographic knee OA* was defined as knee pain and osteophytes in combination with at least one of the following three criteria: 1) age > 50 years; 2) stiffness <30 minutes or 3) crepitus. I used this definition in the validation of the diagnoses registered in the SHR due to its relatively high sensitivity and specificity with respect to a clinical diagnosis (91% and 86%, respectively).[14]

Radiographic and symptomatic knee OA

As a part of the clinical examination in the second part of the MOA study the participants underwent x-rays of both knees in a weight-bearing and semi-flexed position (knees in 10-15 degrees of flexion) using a posterior-anterior beam direction (film focus distance 110 cm, 60 kV and 10 mA) with the aid of fluoroscopy to optimally align the tibia plateau. Patella axial images were obtained with knees in 30-40 degrees of flexion. An independent senior radiologist specialized in musculoskeletal conditions who was blinded to the clinical data assessed joint space narrowing and osteophytes according to the atlas from the Osteoarthritis Research Society International.[73] The *radiographic knee OA* was present if one or more of the following criteria were fulfilled in either the medial tibiofemoral, lateral tibiofemoral or patellofemoral compartment: joint space narrowing grade 2 or worse, the sum of marginal osteophyte grades in the same compartment 2 or worse, joint space narrowing grade 1 and osteophyte grade 1 in the same compartment (approximating a KL grade ≥2) (Figure 5).[42]

I considered those with radiographic knee OA and frequent knee pain (pain in one or both knees in the past 12 months and duration at least 4 weeks) as having *symptomatic knee OA*.
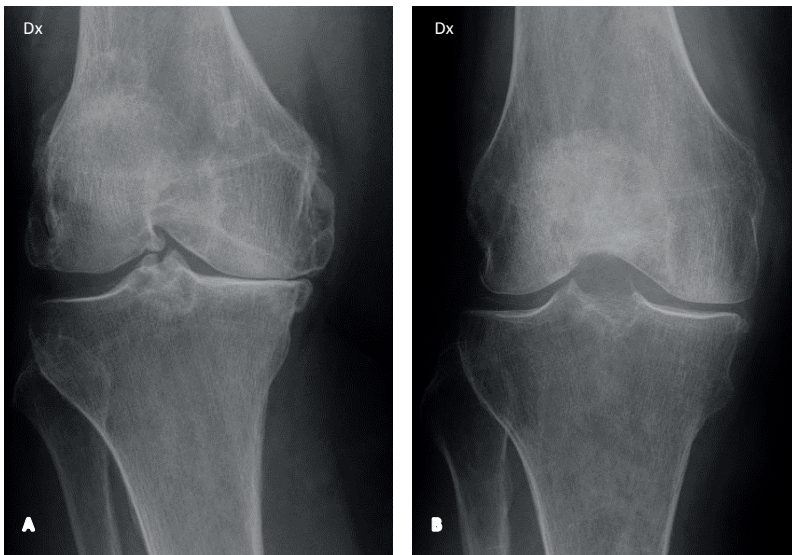

Figure 5. Frontal radiographic knee images from two 80 years old MOA participants.
A. A right knee with radiographic medial osteoarthritis. Joint space narrowing and osteophytes on both tibia and femur are visible in the medial compartment. **B.** A right knee without any signs of medial radiographic osteoarthritis. Both persons suffered from frequent knee pain.

Doctor-diagnosed RA

To be considered an RA case an ICD 10 code M05 or M06 assigned and registered in the SHR at least twice, with at least one being from a specialist (or a physician under specialty training) in rheumatology or internal medicine was required. This definition was previously validated against a registry held by the South Swedish Arthritis Treatment Group (SSATG). Of all RA patients registered in SSATG (98% fulfilling ACR classification criteria for RA) aged ≥20 years, 93% were captured in SHR using the above definition.[50]

Epidemiological methods

Prevalence

Prevalence is an epidemiological measure of the occurrence of the disease that focuses on existing states. The point prevalence of a disease is the proportion of the population with the disease at a specific time point.[74]

Studying the point prevalence of OA using longitudinal healthcare data requires several assumptions. First, the target population needs to be defined. In the paper I, I estimated the point prevalence of OA at 31st Dec 2012 among residents of Skåne aged 45 years or older. Thus, all persons in these ages, alive and being residents in the region at the end of the 2012 were included in the denominator. Second, I assumed that a person has OA from the date of diagnosis and onwards. Further, the number of years with healthcare history included needed to be sufficient to capture all prevalent cases of OA that had consulted a physician for the disease. In the case of OA the sufficient number of years have been reported to lie between 10 and 15.[25,75] I included 14 years of healthcare visits, from 1999 to 2012. Fourth, the positive predictive value of the OA diagnosis in the register needs to be high. In order to verify this assumption, I estimated the positive predictive value of an OA diagnosis in the SHR using the clinical data from the MOA study. Finally, all persons in the target population and having doctor-diagnosed OA were included in the numerator (Figure 6).
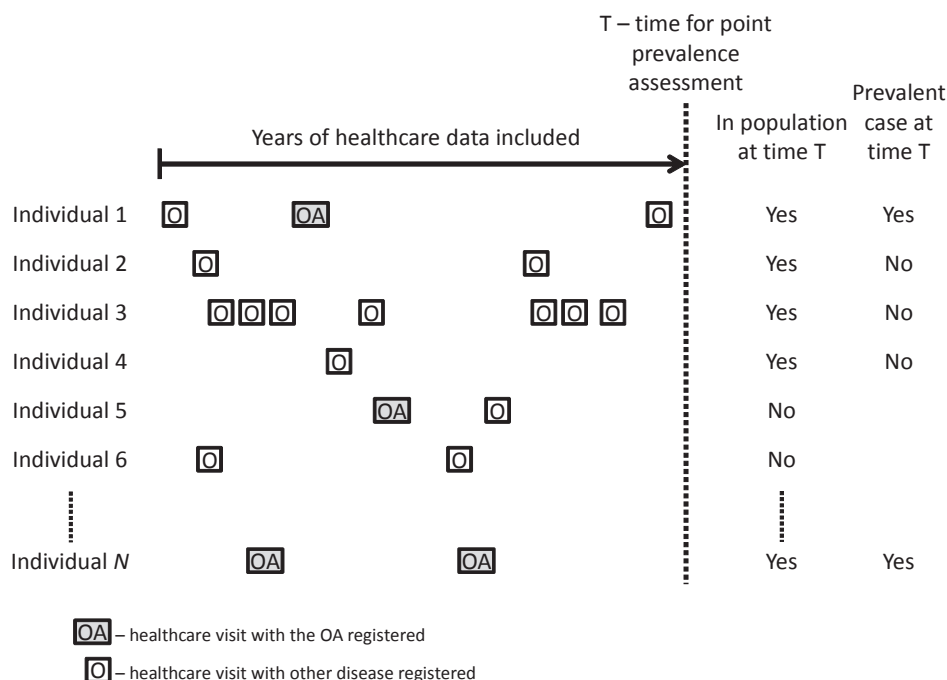
Figure 6. Estimation of the prevalence of OA using healthcare data for a population of size *N*.

Estimating the prevalence of OA from cross-sectional data, such as data from the MOA study, is more straightforward, as the disease status is assessed directly at the time of the completion of the questionnaire and/or clinical examination. If the survey sample is not a simple random sample, the sampling scheme should be taken into account when estimating population parameters. In the MOA study, the sampling was stratified on the knee pain status (frequent knee pain present or not) with different sampling probabilities in each stratum. I have used weighting to adjust for this.

Future prevalence

The projected future age and sex structure of the Swedish population and the anticipated increase in the prevalence of overweight and obesity in Sweden together with the prevalence estimates of the doctor-diagnosed OA in Skåne were used to project the future prevalence of OA in Sweden, up to the year 2032. First the sex and age specific (in age categories: 45-49, 50-59, 60-69, 70-79, 80+ years) population projection from SCB and sex and age specific prevalence of doctor-diagnosed OA were used to project the future prevalence of OA due to changes in the age and sex structure of the population.

Further, I used the observed sex and age specific prevalence of overweight and obesity in Sweden measured by SCB in 1988, 2008 and 2010 (three measurements) to assess the impact of obesity on the future prevalence of OA. I assumed that the increase observed between 1988 and 2010 was linear and would continue until 2032. Based on the previously published results from meta-analyses I assumed that an increase in BMI of five units (which represents moving one "full step" from the normal weight category to overweight, or from overweight to obese) increases the risk for incident OA by 1.35 for the knee (1.22 for men and 1.38 for women) and 1.11 for hip and hand (for both men and women).[76,77] The mean of the estimates for knee and hip was used for 'any OA' because 50% of all 2012 prevalent OA cases had knee OA. I used the following formula linking the prevalence proportions and incidence rates:

$$(number\ of\ prevalent\ OA\ cases) * (1/disease\ duration) = incidence *$$
$$(number\ of\ people\ in\ population - number\ of\ prevalent\ OA\ cases) .$$

The disease duration was assumed to be equal to the life expectancy at the mean age of persons with OA. The years of life expectancy for those aged 65 in the year 2012 were retrieved from SCB publications.

To assess sensitivity of the projections with respect to assumptions about the BMI increase and its impact on risk of OA I modified the risk rates for incident OA to be equal with their lower or upper confidence limits, i.e., 1.23 (lower limit) or 1.54 (upper limit) for knee OA in women, 1.19 or 1.25 for knee OA in men and 1.07 or 1.16 for hip and hand OA. Further, I modified the projected prevalence of overweight and obesity and assumed that the increase would be 10% lower or 10% higher than the one observed between 1988 and 2010.

Handling of missing data

Missing data are inevitably ubiquitous in observational epidemiological research. There are several known methods of dealing with missing data, such as weighting, random effects models incorporating partially observed data, maximum likelihood estimation or Bayesian methods including multiple imputation (MI).[78] All these methods relay on the assumption that the data are missing completely at random (MCAR) or missing at random (MAR). The data are MCAR if the missingness mechanism is unrelated to any inference we wish to draw from the data. The data are MAR if the probability that an observation is missing depends on the observed data, but not on the unobserved data. If the data are neither MCAR nor MAR they are called missing not at random (MNAR). It cannot be statistically tested if the data are MAR rather than MNAR, but the judgment of the plausibility of the MAR assumption is nevertheless an important step of the analysis. When the data are MNAR, other approaches, such pattern-mixture models, can be used under some assumptions concerning the missingness patterns. A discussion about such models falls beyond the scope of this thesis.

Missing diagnostic codes in the SHR

In the SHR 1998-2012 data the completeness was high with respect to the administrative information regarding healthcare visits (such as the personal identification number, the date, the clinic or the healthcare provider), however the ICD 10 diagnostic codes were missing for some visits. First, the diagnostic codes set in the private care were not forwarded to the register, while the other details of the visits were. Second, the percentage of visits with a diagnostic code varied depending on the healthcare level, with almost 100% completeness in in-patient care, about 90% completeness in specialist care but around 15% missing values in public primary care. Third, the registration of codes in out-patient care has been continuously increasing over the years, and the completeness was higher from year the 2004 and onwards, when the registration of codes was made obligatory for the public care. This missingness may have different impact on the inferences depending on the study question. In situations when a high sensitivity is required, such as prevalence studies, ignoring missing data may lead to severe underestimation of the prevalence. However, it may have smaller impact in studies of risk factors or disease consequences where a high specificity may be of importance.

In the SHR, the substantial part of the missingness in diagnostic codes could be linked to visit-specific factors (such as healthcare level, provider or year of visit). Thus, to make the MAR assumption more plausible, these factors needed to be taken into account when attempting to estimate the prevalence of OA in the population. One of the established methods suitable for this purpose was MI.

MI was introduced by Rubin and conceived as a two-stage Bayesian approach to estimation of the parameters in the presence of missing data but the estimates can have very good frequentist properties.[79,80] A joint model is fitted where the partially observed data are the response and a number of complete datasets are created by drawing the missing data from their conditional distribution given the observed data. These complete datasets can be analysed with the standard methods and the results are combined taking into account the statistical uncertainty of the imputation process.[81]

The MI methods for handling unbalanced longitudinal data, where the observations coming from the same individual are expected to be correlated, are limited. In the paper I and II I used the multivariate normal model for longitudinal data developed by Schafer to multiply impute missing diagnostic codes.[82] It has been showed that imputing binary data with a multivariate normal model can be done without introducing bias, especially if the imputed values don't need to be rounded or when a proper rounding procedure is used.[83,84] I used the calibration method introduced by Yucel to derive the cut-off for rounding of the imputed values.[85] This method has been shown to provide unbiased marginal probabilities even if the proportions of interest were low, as is the case with the OA diagnosis at a healthcare visit. To evaluate this approach I performed a simulation study (paper IV). In brief, in a generated fully observed datasets, I created missing diagnostic codes according to the MCAR or MAR mechanisms. I then derived the point prevalence estimates using

34

the MI method. Additionally, I estimated the prevalence using a naïve method that simply ignored missing data and compared the two approaches. The technical details of the imputation procedure and the analysis can be found in the papers I and IV.

In the paper III, where I studied the relative mortality in doctor-diagnosed OA and RA as compared to the general population, I did not adjust for the missing diagnostic codes explicitly. It was unclear if the MI approach would be sufficient to preserve the relationships between OA and other variables of interest, such as confounders, death and time to death. The positive predictive value of an OA diagnosis in the SHR was high (88%) and thus the OA group defined through diagnoses registered in SHR could be expected to have high specificity. I chose to perform sensitivity analyses in order to assess the robustness of the results to the possible misclassification of the OA status (where the missing diagnostic codes may have contributed).

Non-response in the MOA study

In the MOA study, the missing data were mainly a consequence of the non-response in the postal questionnaire and non-attendance to the clinical assessment and/or x-ray examination. However, data on age, sex, education and BMI measured at the MDCS examination were available for the whole study sample. The response rate in the mailed survey was 77% and 72% of responders were willing to further participate in the study. Among persons sampled to undergo the clinical examination 78% attended the visit. In the paper II, I have used weighting to adjust for the possible selection bias due to the non-response.[86] A logistic regression model with sex, age and BMI as well as the highest education level as covariates was used to estimate the probability of response in the mailed survey, and the reciprocal was used as a weight. Covariates in models for willingness to participate as well as for attendance at the clinical examination included additionally the knee pain status (from part I). The sampling weights (the reciprocal of the sampling probability for those with and without frequent knee pain) were multiplied by the weights for non-response and willingness to participate to construct the final weights used in the analyses.

Survival analysis

In the survival analysis the average time to the occurrence of an event, starting from some designated zero time, is modelled. Usually not all subjects will experience the event of interest during the study period and thus the information about their time to event is incomplete (censored). The available information for censored subjects is that they have not experienced the event of interest during the observation period. The proportional hazards model proposed by Cox has become commonly used for analysis of censored data.[87] The hazard function at time t gives an information of the instantaneous potential per unit time for the event to occur, given that the individual has not experienced the event up to time t.[88] In the Cox model the hazard function is

related to a set of explanatory variables through a regression equation which enables estimation of the hazard ratios (HR) for these variables.

In the paper III, I used the Cox model to estimate the HR of death for persons with doctor-diagnosed RA or OA as compared with the general population seeking healthcare, while controlling for a number of other factors. The first healthcare visits with any diagnosis registered in the SHR between the year 1998 and 2012 indicated the beginning of the follow-up time for each included person (being resident in Skåne and 45 years of age or older). Using the information from the population register all subjects were followed until death, relocation outside the Skåne region or the end of the study at the 31st Dec 2013, whichever occurred first. To avoid immortal time bias, the follow-up time from the first healthcare visit until the date at which the doctor-diagnosed RA or OA definition was fulfilled was treated as unexposed, while the follow-up time from the date of fulfilling the disease definition and onwards was treated as exposed.

When analysing mortality in a cohort study, one of the important choices is that of a time scale. Commonly, the time since entry to the study is used. However, in epidemiological studies of chronic diseases, where age often exerts substantial confounding and when the time of inclusion to the study is arbitrary, the time since birth (age) can be a more suitable option as it provides unbiased estimates of HRs.[89] I used the attained age as a time scale because both RA/OA and mortality are heavily dependent on the person's age. An important assumption of the Cox model is that hazards are proportional. That is, with time-fixed covariates (i.e., covariates that do not change over time), the relative hazard for any two individuals is independent of time and this holds for each variable in the model individually. I verified this assumption using plots of Schoenfeld residuals.[90]

Controlling for confounding

Confounding is a central concept in epidemiological research. It is a process that can result in biased results when examining the causal association between exposure and outcome. Confounding is said to occur when there are differences in outcome in the exposed and unexposed populations that are not due to the exposure, but are due to other variables. Such a variable must be a cause of the outcome (or a proxy for a cause) and must not be an effect of the exposure.[91]

The most common way of adjusting for confounders is regression modelling. When including the sufficient set of correctly measured confounders in the regression model one is able to estimate the unbiased association between the exposure and outcome. However, if the confounding is time-dependent then the standard regression techniques may still lead to biased estimates.[92] The aim of the paper III was to estimate the relative mortality in persons with OA or RA as compared with the general population. There is evidence suggesting that cardiovascular disease may be a risk factor for OA. At the same time, both cardiovascular disease and other serious

comorbidities such as cancer or diabetes may have influenced the registration of OA at the healthcare visit. On the other hand, cardiovascular disease may be a consequence of OA.[93,94] (Figure 7) Similar constraints apply for RA.
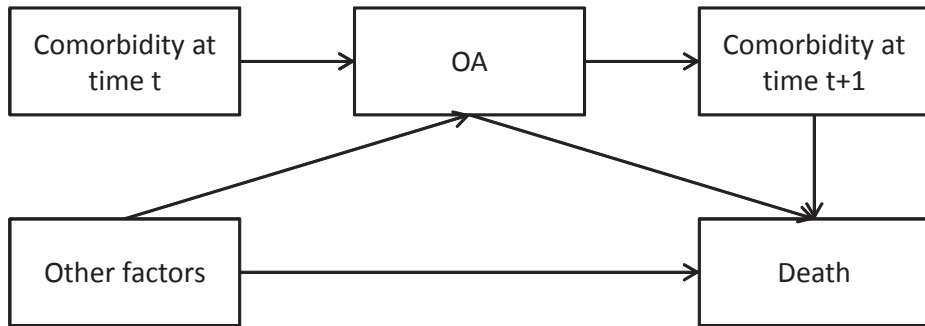

Figure 7. Time-dependent confounding of the OA-death association by a comorbidity.
Flow-chart adapted from Fewell et al.[95]

Thus, to adjust for the comorbidities that were possible confounders of the exposure (OA or RA)-death association I used the inverse probability weighting approach.[96] The model was fitted in a two-stage process in which:

1. Each subject's probability of having their own exposure history was estimated and used to derive *inverse-probability-of-exposure weights* (or *inverse-probability-of-censoring weights*)

2. The exposure-outcome association was estimated in a regression model weighted by the *inverse-probability-of-exposure/censoring weights*

I used the logistic regression both to derive the exposure (and censoring) weights and to estimate the final hazard ratios of death. Using the SHR data, I adjusted for a number of chronic comorbidities that are common causes of death or associated with risk factors for RA or OA (smoking and overweight): ischaemic heart diseases (ICD 10 code: I20 to I25), cerebrovascular disease (I60-I69), diabetes mellitus (E10-E14) and other malignant neoplasms (C, other than C34), chronic obstructive pulmonary disease (J44) and malignant neoplasm of bronchus and lung (C34). I also adjusted for a number of baseline covariates (assumed to be constant during the follow-up time), i.e., sex, residential area, income, highest level of achieved education, marital status and the year of the first healthcare visit. The aim of this process was to estimate the average causal effect of the OA/RA on the mortality. In the observational study setting any causal inference requires strong assumptions of no unmeasured confounding and no measurement error. These assumptions cannot be statistically tested and are seldom fully fulfilled. Sensitivity analyses may be a tool to assess the robustness of the results in such situations.

Quantifying bias

When using register data, substudies are seldom performed in order to examine misclassification of exposure, endpoints or confounders. Likewise, obtaining additional data on unmeasured confounders is often not feasible. Thus, the only alternative may be to present sensitivity analyses, including quantification of bias. Such analyses have traditionally been overlooked in favour of a preoccupation with random error.[67] In recent years the concept of quantifying bias, instead of only discussing it, have been given greater attention and several methods to quantify bias due to unmeasured confounding, misclassification or selection bias are available.[97]

There were several sources of bias that may have influenced the results in the paper III. First, physical activity (that may lead to joint injury) is a known important confounder of the OA-death association, which I was not able to adjust for in the analysis model.[62] I used the rule-out approach to estimate how common and how strongly associated with mortality an unmeasured binary confounder would need to be to explain the results if doctor-diagnosed knee and hip OA truly had an increased mortality risk (HR=1.5) or had no association with mortality (HR=1).[98]

Second, based on the data on the healthcare consultation in knee OA patients from paper II, I assumed that as much as 40% of all true OA cases were misclassified when using the healthcare data to identify the OA cases. I sampled additional OA patients (separately for the knee and the hip) from those not fulfilling the doctor-diagnosed OA case definition so that their mortality rate ratio would be 1.5 (i.e., I sampled proportionally more individuals that died during the study period) and then repeated the Cox regression analysis adjusted for sex and other baseline covariates. I performed this procedure 1000 times and reported estimates with bootstrap 95% confidence intervals (CI).

Results and Discussion

Prevalence of OA

Current prevalence

Doctor-diagnosed OA

I found that among the 531 254 adult residents of Skåne aged ≥45 years in the year 2012 26.6% (95% CI 26.5, 26.8) had a doctor-diagnosed OA in at least one peripheral joint.(paper I) The most common location was the knee, with the prevalence of 13.8%, followed by the other joints (12.8%), the hip (5.8%) and the hand (3.1%). The prevalence was higher in women than men and typically increased with increasing age.(Figure 8) Of all persons with OA, 32% were younger than 65 years. The prevalence estimate of 'any OA' is in line with previously published data based on healthcare consultations in Canada, where prevalence of OA in any joint was reported to be 24.7% in persons ≥45.[25] The prevalence in the knee and the hip, the two most common affected sites, were in the middle of the range reported previously in a number of western countries and could be expected if not all symptomatic patients consult healthcare.[23,27,28] The prevalence of doctor-diagnosed hand OA was 3.1% and decreased after the age of 75 years, especially in women. This prevalence was lower and the pattern was different than in studies of the elderly in the US or United Kingdom (UK) based on clinical examinations, where the prevalence was stable across age groups[99] or increased with increasing age[100,101]. This discrepancy may be explained by the healthcare consultation pattern, if the hand problems are down prioritized in the presence of other comorbidities or if the elderly patients see the joint symptoms in hand as a part of aging and thus delay seeking medical care.
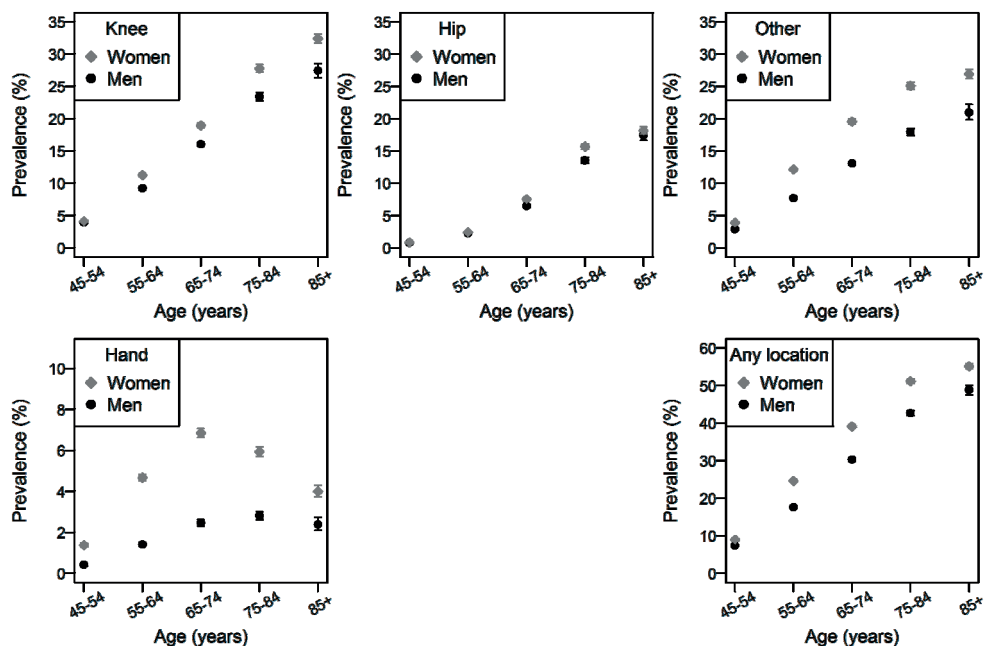
Figure 8. The 2012 point prevalence (with 95% confidence intervals) of doctor-diagnosed OA in the Skåne population aged ≥45 years.

The group with the second highest prevalence was OA in other joints where I included peripheral joints typically less often studied, such as foot, elbow, shoulder and jaw. The combined prevalence of doctor-diagnosed OA in these joints was 9.6% in men and 14.9% in women, only slightly lower than for the knee. This group of patients causes a substantial burden on the healthcare system and should receive more attention in the healthcare planning, health economic evaluations and studies about the burden of OA.

Of all doctor-diagnosed OA cases, 26.8% were diagnosed with OA in at least two locations.(Figure 9) This number was lower than among 500 consecutive patients (age 45+ years) in a rheumatology clinic in UK, where half of all examined patients had OA in at least two sites.[102] However, the percentage of persons with hip OA among those with knee OA was similar (12% in UK and 13% in SHR). The co-occurrence of the OA disease in different joint sites has been observed previously,[103,104] but the phenomenon is still not fully understood and the magnitude of effects vary with the definitions of the diseases used and study samples. It is important to acknowledge that my results represent not only the co-occurrence of the OA diseases in different joints, but also the healthcare seeking pattern, when persons already diagnosed with OA in one location may have higher propensity for getting diagnosed in another site.
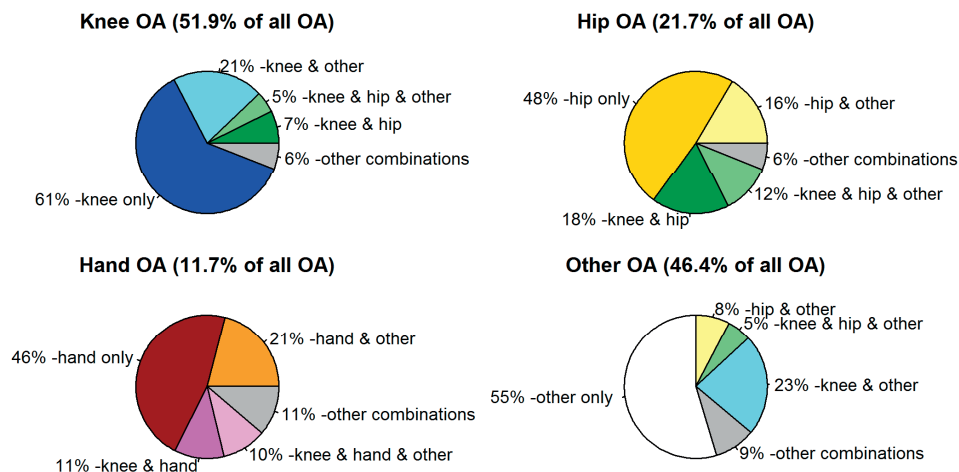
Knee OA (51.9% of all OA)

21% -knee & other
5% -knee & hip & other
7% -knee & hip
6% -other combinations
61% -knee only

Hip OA (21.7% of all OA)

48% -hip only
16% -hip & other
6% -other combinations
12% -knee & hip & other
18% -knee & hip

Hand OA (11.7% of all OA)

46% -hand only
21% -hand & other
11% -other combinations
10% -knee & hand & other
11% -knee & hand

Other OA (46.4% of all OA)

8% -hip & other
5% -knee & hip & other
23% -knee & other
9% -other combinations
55% -other only

Figure 9. The 2012 site distribution of OA that burdens health care and its overlap in adults aged ≥45. Other - includes OA of joints other than knee, hip and hand such as shoulder, elbow, ankle and foot, jaw and polyarthrosis (excluding Bouchard's and Heberden's nodes as those are included in the hand). Other combinations - includes all combinations of knee, hand, hip and other, not already listed in a pie. Spine is not included.

The prevalence of OA based on the healthcare register data may depend on the definition used. I required one visit to a physician with the OA diagnosis registered. Several different definitions have been used in previous studies; for example, at least two visits with a diagnosis, diagnosis from a specialist or a combination of diagnosis with medication prescriptions.[25,75] Generally, the obtained prevalence estimates were higher when a less stringent definition was used. The validity of OA diagnoses may be expected to vary depending on the data source used, such as a visit to a physician, any contact with healthcare, inpatient stay or medication prescriptions.[105] Additionally, some of the validity measures, such as sensitivity to capture all prevalent cases, may vary with the number of years of healthcare history included even if the positive predictive value of the particular case definition is fixed. Using the MOA data and a gold-standard definition of knee OA (ACR clinical and radiographic criteria or KL≥2) I could estimate that the positive predictive value of the knee OA diagnosis in the SHR was 88%. I considered this a sufficient justification for using one diagnostic OA code to define an OA case. This approach has also been advocated previously in similar settings.[106] However, the positive predictive values of OA diagnoses in joints other than knee have not been evaluated in SHR.

Radiographic, symptomatic and clinically defined knee OA

In the MOA study, the 2008 prevalence of radiographic knee OA among persons aged 56 to 84 was 25.6% (95%CI 22.7, 28.6) and was similar to the prevalence of frequent knee pain - 25.1% (95%CI 24.1, 26.1).(paper II) One in ten persons could be classified as having symptomatic knee OA and 9.6% had clinically defined knee

OA, but the overlap between the groups was low.(Figure 10) The prevalence of painful knee OA (either symptomatic or clinically defined knee OA) was 15.4%.

While the prevalence of radiographic knee OA in this study wass in line with numbers from a Dutch study in a population aged 55-79 years, it was lower than the prevalence in the US.[23,107–109] One of the explanations may be the lower prevalence of overweight and obesity in Sweden than in many western countries, However, the prevalence of symptomatic knee OA in the MOA study sample was similar to this in US, Canada or Western Europe,[23,27,28] while the prevalence of knee pain was similar or higher.[20,110,111]

Clinically defined knee OA
Frequent knee pain
0.7%
0.9%
13.5%
4.3%
3.1%
6.5%
11.1%
Radiographic knee OA
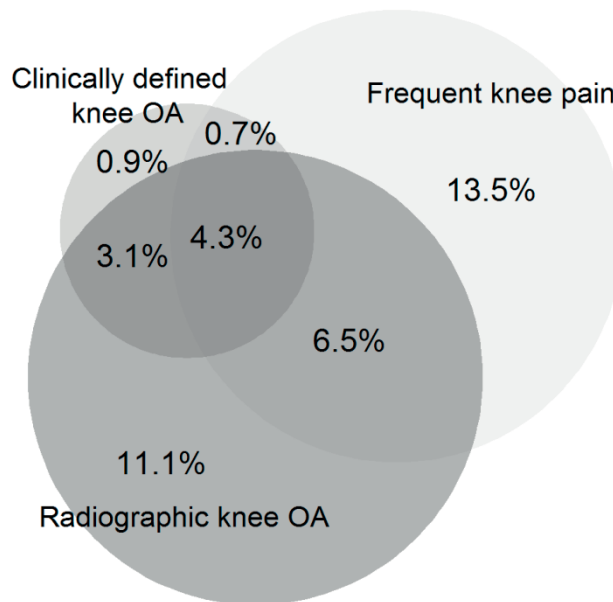
s

Figure 10. The 2008 prevalence and overlap od frequent knee pain and knee OA in Swedish adults 56-84 years old.

As also seen in previous works, the overlap between knee pain and radiographic signs of knee OA was low.(Figure 10) This should not be seen as evidence of a weak causal association, as pain is a subjective experience unique to each person and may be strongly associated with knee OA even if it is not a good predictor of the latter. A previous study has shown that when analysing matched knees within individual participants with discordant pain status, the radiographic features are strongly associated with knee pain.[112] Similarly, the overlap between the symptomatic knee OA and clinically defined knee OA (according to the ACR criteria) was low, a finding that is in line with previous work from England.[21] These results emphasize the

importance of incorporating several OA definitions in order to obtain a broader picture of the occurrence of the disease. In the MOA study sample the prevalence of knee OA giving symptoms (symptomatic or clinically defined) was 50% higher than the prevalence of symptomatic knee OA alone (or clinically defined knee OA alone). It is a substantial difference in terms of the disease burden that would be missed out when using only one definition.

Knee OA and healthcare consultations

Between the year 2004 and 2011 half of all MOA participants with either symptomatic or clinically defined knee OA were diagnosed with knee OA by a physician in Skåne, while 69% received either knee OA or pain in joint diagnosis.(paper II) This indicates a potential unmet need among persons with knee OA. The relatively low proportion of persons with knee OA consulting healthcare has been previously reported from England.[21] Older people may often view the chronic joint pain as a part of normal aging, delay seeking medical care and adhere poorly to the prescribed treatment, either in the form of medical drugs or recommended physical activity.[113,114] On the other hand, self-coping strategies, over-the-counter pain treatments and relatively easy access to physiotherapy may be other explanations. For example, during the years 2004-2011 84% of the persons in the MOA study with either symptomatic or clinically defined knee OA visited a physiotherapist at least once compared to 66% of those without.

The major strength of the results from papers I and II is the inclusion of a number of relevant knee OA definitions within the same population. The SHR data provided information important for healthcare planning from a large study sample with low risk of selection bias and high statistical precision of the estimates. However, the data were left-truncated (available from year 1998 and onwards) and thus the definition of doctor-diagnosed OA may be prone to misclassification if the only diagnosis was made before the year 1998. The MOA study, on the other hand, provided detailed information not covered in the SHR regarding the symptoms and radiographic severity of knee OA. These detailed data were plausible to obtain only in a smaller sample prone to selection bias due to non-response. The two data sources combined gave a complete, robust and consistent picture of the current prevalence of OA in southern Sweden. Interestingly, applying the age and sex specific estimates of doctor-diagnosed knee OA (from paper I) in the age group 55 to 84 of the Malmö population (i.e. the source population for the MOA study) as per year 2012 yielded a prevalence of 16.2%, similar to the 15.4% prevalence of 'clinically relevant' knee OA (i.e. symptomatic or clinically defined knee OA). Due to the complexity of the OA disease and no uniform diagnostic criteria available, the group of doctor-diagnosed knee OA does not have a perfect overlap with groups defined by radiographic, symptomatic or ACR criteria. Yet, it may be an important target group for epidemiological studies of occurrence and consequences of the disease.

Future prevalence

Taking into account only the projected changes in the sex-age structure of the population aged 45 years or older the prevalence of the doctor-diagnosed 'any OA', the knee OA or the hip OA, respectively, is expected to increase from 26.6% to 29.0%, from 13.8% to 15.2% and from 5.8% to 6.8% over the next two decades.(paper I) When accounting additionally for the increase in the estimated prevalence of overweight and obesity the relative increase will be 10%, 12% and 18% for 'any OA', knee OA and hip OA, respectively.(Figure 11) These estimates are lower than those of other researchers in US and Canada who reported projections of the prevalence of self-reported arthritis with relative increase that ranged from 16 % to 50% within two decades in the population aged 15 or older. The reason for this discrepancy may be that the relative shift in the future population towards higher fraction of older adults affected my results (in persons aged 45+) to a lesser extent. Nevertheless, my estimates suggest that in Sweden in the year 2032 there will be additional 350 000 persons diagnosed with doctor-diagnosed OA, while almost 200 000 of them will be diagnosed with knee OA specifically. This may pose a great challenge for the primary care, physiotherapists and orthopaedic surgeons, especially if no breakthrough in the treatment of OA will occur.[36]

OA continues to affect a large proportion of middle-aged and elderly people both in terms of personal suffering and healthcare utilization. In most western European countries the age of retirement is increasing and therefore, the future work force will include an increasing number of individuals suffering from OA. Prioritization of research on the population health strategies to reduce OA, including weight loss and joint injury prevention, and, for those affected by OA, the development of effective therapies is needed.

Naturally, projection of the future is difficult and associated with a high degree of uncertainty. There are many factors, such as the incidence of knee injuries, willingness to consult, future available treatments for OA, or the capacity of the healthcare system that will impact the future prevalence of doctor-diagnosed OA. The estimates of the future population structure and prevalence of overweight and obesity used in my projection are themselves subject to uncertainty of unknown magnitude. However, even though there are obvious difficulties in quantifying the future challenges, such a population based projections may be a step towards a better healthcare planning. These results should be put into a broader perspective, where the future occurrence of other highly prevalent diseases, such as cardiovascular diseases, diabetes or cancer, and their co-occurrence with OA is taken into account.[115] More knowledge about the comorbidity patterns in persons with OA could aid this process.
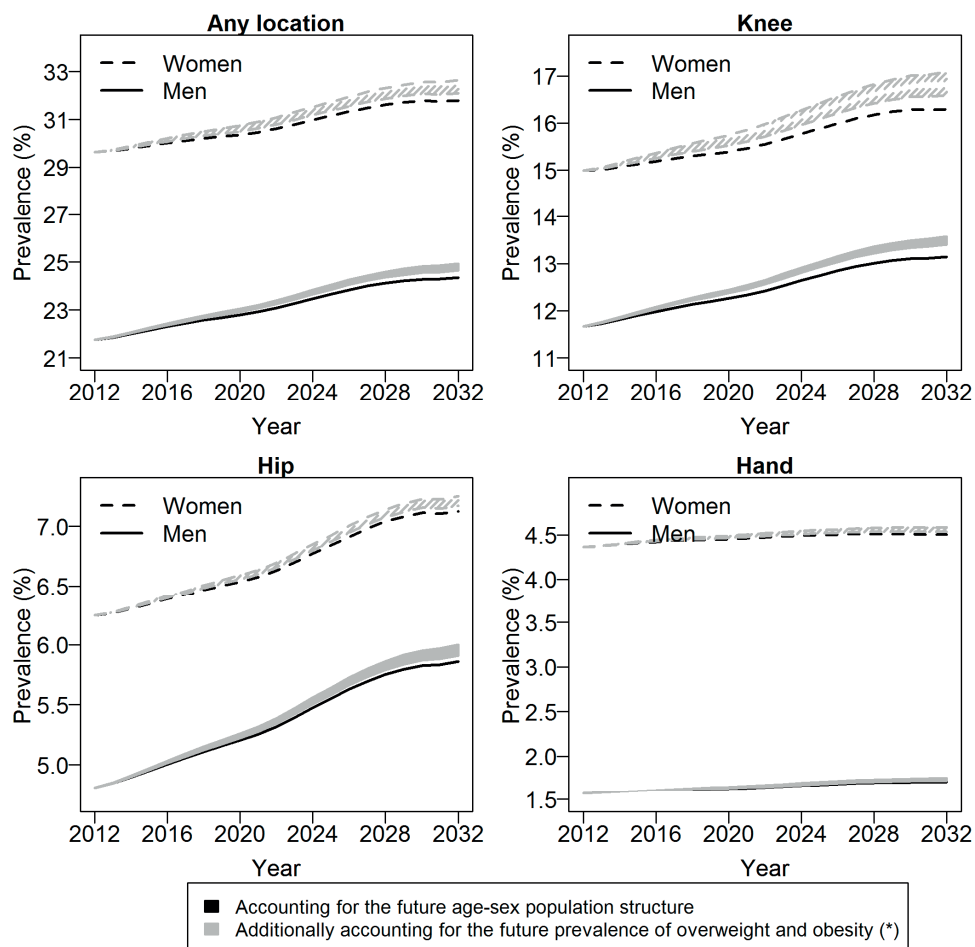
Figure 11. The 2012 prevalence and the projection for year 2013-2032 of doctor-diagnsosed OA.
*The projection is based on the predicted changes in age-structure of the Swedish population, predicted increase in the age and sex specific prevalence of overweight and obesity in Sweden and differential effects of increased BMI on different joint sites. The lower boundary of the gray area shows the projected prevalence when assuming the impact of BMI on the incidence of OA equal to the lower confidence level and the increase in the prevalence of overweight and obesity 10% lower than that observed in Sweden between 1988 and 2010. Accordingly, the upper bound of the gray area shows the projected prevalence assuming the impact of BMI equal to its upper confidence level and the increase in the prevalence of overweight and obesity 10% higher than that observed in Sweden between 1988 and 2010.

Mortality in OA and RA

In the paper III, the register based study of mortality in doctor-diagnosed OA and RA, I identified 51 939 persons with knee OA, 29 422 persons with hip OA and 8067 persons with RA, among 524 136 persons aged 45 or older and resident in Skåne who consulted a physician at least once between 1998 and 2012. The persons were followed for a median of 10 years and 157 008 of them died during the study period (Table 5). The HR of death for persons with the hip or knee OA was 0.87 (95%CI 0.85, 0.89) and 0.90 (0.87, 0.92), respectively.(Table 6) The HR of death for persons with RA was 1.6 (95%CI 1.5,1.7)

Table 5.

Descriptive characteristics of the study cohort.

	RA n=8067	Knee OA n=51 939	Hip OA n=29 442	All n=524 136
Women, %	69.6	59.8	58.4	52.8
Age*, mean years (SD)	67.6 (10.6)	69.5 (10.7)	71.5 (10.1)	63.3 (12.2)
Person time, years	61 985	358 803	203 955	5 379 060
Deaths, n	3105	11 852	8128	157 008
Crude mortality rate per 10 000 person-years	500.9	330.3	398.5	291.9
Married, %	60	63	61	59
Highest level of achieved education, %				
Primary and lower secondary education, up to 9 years	49	45	46	41
Upper secondary education	36	37	37	37
Post-secondary education, less than 3 years	7	8	8	9
Post-secondary education, 3 years or longer, or postgraduate education	8	10	10	13
Income, %				
1st quartile	32	27	28	25
2nd quartile	30	27	29	25
3rd quartile	24	26	25	25
4th quartile	14	20	18	25
Number of visits to a physiotherapist per one follow-up year, median[interquartile range]	0.5 [0-3.4]	0.9 [0-3.8]	0.7 [0-3.3]	0.1 [0-1.5]

*Age at inclusion for All, age at the date of fulfilling the RA/OA definition for those with RA/OA; RA=rheumatoid arthritis, OA=osteoarthritis.

The estimates for RA were in line with previously reported data suggesting that SHR data may be a valid source for studying mortality in rheumatic diseases.[54] The estimates for OA indicated mortality not higher than in general population. In order to aid the interpretation of these results I have performed a number of sensitivity analyses.

Table 6.

Mortality in doctor-diagnosed RA, knee OA and hip OA estimated with the Cox proportional regression model with attained age as time scale.

		Adjusted for sex*		Adjusted for sex* and baseline confounders†		Adjusted for sex*, baseline confounders† and comorbidities‡	
RA	Men	1.59	(1.50, 1.70)	1.53	(1.43, 1.63)	1.59	(1.47, 1.72)
	Women	1.98	(1.90, 2.07)	1.92	(1.83, 2.00)	2.03	(1.93, 2.13)
	All	1.83	(1.77, 1.90)	1.76	(1.70, 1.83)	1.86	(1.78, 1.94)
Knee OA	Men	0.82	(0.80, 0.85)	0.83	(0.81, 0.86)	0.89	(0.86, 0.92)
	Women	0.81	(0.79, 0.83)	0.81	(0.79, 0.83)	0.85	(0.83, 0.87)
	All	0.82	(0.80, 0.83)	0.82	(0.80, 0.84)	0.87	(0.85, 0.89)
Hip OA	Men	0.86	(0.83, 0.89)	0.86	(0.83, 0.89)	0.90	(0.87, 0.94)
	Women	0.84	(0.81, 0.86)	0.85	(0.82, 0.88)	0.90	(0.86, 0.93)
	All	0.84	(0.83, 0.86)	0.86	(0.84, 0.88)	0.90	(0.87, 0.92)

Results are presented as hazard ratios with 95% confidence intervals. RA=rheumatoid arthritis, OA=osteoarthritis *In the analysis of both sexes combined; †confounders measured at inclusion: income, highest level of achieved education, marital status, residential area and year of first healthcare visit; ‡adjusted for ischaemic heart diseases (ICD 10 code: I20 to I25), cerebrovascular disease (I60-I69), diabetes mellitus (E10-E14) and other malignant neoplasms (C, other than C34), chronic obstructive pulmonary disease (J44), malignant neoplasm of bronchus and lung (C34) registered in Skåne Healthcare Register using the inverse probability weighting approach.

There were two major confounding factors of the OA-death association that were not adjusted for in the analysis model. First was the physical activity. It is associated with sports injuries, which are a strong risk factor for OA, and inversely with the risk of mortality.[116] In the sensitivity analysis of an unmeasured binary confounder (physical activity 'sufficient for sustaining health') I found that it could explain my result of HR equal to 0.9 if the true HR was 1, but very unlikely if the true HR was 1.5. Second, the data on BMI were not available for the included individuals. However, persons with OA are expected to have on average higher BMI than persons without OA. If so, an adjustment would drive the estimates farther from the null. To additionally address this issue I have used the sample of 422 persons with symptomatic knee OA from the MOA study. The proportion of overweight individuals that had a knee OA diagnosis registered in the SHR at least once between 1998 and 2012 was 65%. Among those who did not have the knee OA diagnosis registered the corresponding proportion was 56%. This suggests that overweight persons with knee OA were not less likely to consult for knee OA.

The misclassification of the OA status can be expected in the left-truncated healthcare data if the only diagnosis is made before the start of the registration period. Additionally, not all symptomatic cases seek healthcare or get the diagnosis registered. As a consequence, one cannot be sure if those with a registered OA diagnosis are representative of all 'true' OA cases. Those consulting for OA would not be expected to be on average healthier than non-consulters. Previous research has indicated that those with severe or chronic pain are more likely to consult, while the number of comorbidities was similar.[117–119] On the other hand, it has been reported that only a fraction of persons with symptomatic disease seek care.[20] In the sensitivity analysis I assumed that 40% of true OA cases were falsely classified as non-OA and those 40%

had a HR of death of 1.5. Even in this extreme scenario the corrected HR would not exceed 1.15.

When including the entire general population, without the requirement of a consultation with a physician, the fully adjusted HR of death was 2.05 (95% CI 1.97, 2.13) for RA, 0.92 (95% CI 0.90, 0.94) for knee OA and 0.95 (95% CI 0.93, 0.97) for hip OA.

The conclusion from the analyses is that in Sweden knee and hip OA are not associated with increased mortality as compared to the general population. Despite the point estimates being lower than 1, I would refrain from claiming a 'protective' effect due to the results of the sensitivity analyses and due to the doubtful value of such a statement from the practical perspective. The finding of no increased mortality is of importance for the public health and clinical practice. It is of interest to replicate these results even for OA patients not diagnosed within healthcare to better understand the mechanisms behind this finding. My results are in contrast with results reported previously in England and US,[55,120] but similar to those based on the healthcare data from other countries.[58,121] I cannot exclude that persons with OA may be at higher risk of death from vascular causes specifically and this is an important question for future research.

The generalizability of the results to other countries is not straightforward. The underlying health of the population and the treatment of OA received when consulting a physician may have important implications. In Sweden, the prevalence of obesity is lower,[122,123] while the proportion of individuals involved in physical activity is higher than in most European countries.[122] Relatively easy access to physiotherapy and joint replacement may be effective in preventing physical inactivity, weight gain or lower the need for the use of nonsteroidal anti-inflammatory drugs.[124–126] A recent publication on physical activity levels among persons with and without knee OA in 6 European countries reported that persons with OA have generally lower activity level than those without OA. However, such a difference could not be found in Sweden, where those with and without OA reported similar physical activity levels.[127] A health economic study based on the data from the MOA cohort found that the mean EQ-5D scores among persons with knee OA in the Malmö area was higher than for comparable patients in Singapore or UK.[128] There is a clear need for up-to-date studies of mortality in OA in countries other than Sweden in order to improve our understanding of the prerequisites for preventing potential excess mortality in OA on the population level.

Multiple imputation for handling of missing diagnostic codes in a longitudinal healthcare register

I performed a simulation study to assess the usefulness of the MI of missing diagnostic codes in a longitudinal healthcare register in order to estimate the prevalence of a chronic disease D in a population and to compare MI with the naïve method that ignores the missing data. I found that the MI approach had lower magnitude of bias than when ignoring the missing data (naïve method), but still was not unbiased in some data generation scenarios.(Figure 12) The naïve estimates consistently underestimated the true prevalence, with mean relative bias of −25% when 50% of the diagnoses were missing and of −10% when 25% of diagnoses were missing. The more visits with the disease of interest per person, the lower the bias of the naïve method. Thus, the naïve method could have a better performance for diseases that require frequent visits to a physician (not considered in this simulation study). One could speculate that including more years of the healthcare history than actually needed to identify all the prevalent cases could reduce the bias of the naïve method, but this approach could only be helpful in identifying persons with long disease duration rather than recent incident cases.
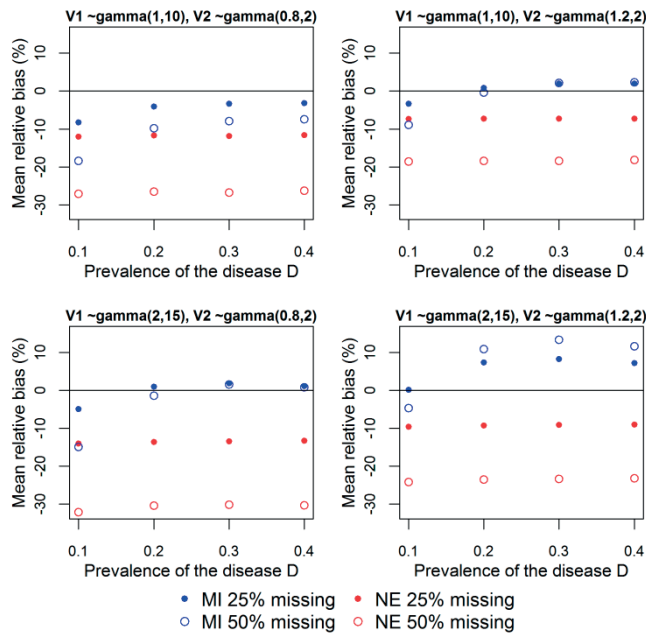

Figure 12. The mean relative bias from 1000 simulations for different distributions of number of healthcare visits (V1) and number of visits with the disease D registered per person (V2) (missing data MAR).
NE – naive method ignoring the missing data. The *gamma(a,b)* distribution with a shape parameter *a* and a scale parameter *b* is a two-parameter continious probability distribution, where a special case is for example exponential distribution.

The mean relative bias of the MI prevalence estimates across all scenarios was 1% when 25% of data were missing and −1% when 50% of data were missing. In most scenarios the MI prevalence estimates were close to the nominal values. However, when the true prevalence of the disease was low and there were relatively few healthcare visits per person, both concerning the disease of interest and in total, the prevalence could be underestimated by −18% when 50% of the diagnoses were missing. Thus, the estimated prevalence would be 0.08 when the true prevalence was 0.1. On the other hand, with the true disease prevalence of 0.2 or higher and many healthcare visits per person (both in total and due to the disease of interest) the prevalence could be overestimated by 15%. For example, the estimated prevalence would be 0.34 when the true prevalence was 0.3 The coverage of the 95%CI was greater than or equal to 95% for scenarios with low bias, but was <95% when relative bias was greater than 6%.

The performance of both methods, the naïve method and the MI, depended on the distributions of the total number of visits and the number of visits with the diagnosis of interest registered. Thus, one can expect that the same percentage of missing diagnostic codes can have different impact on the prevalence estimates depending on the disease, the population studied or the number of years of healthcare contacts included.

The simulation study provided a means of validation of the results in the paper I. Additionally it gave information about a possible bias, and its magnitude, in the prevalence estimates derived from the longitudinal healthcare databases with missing diagnostic codes when using MI or when ignoring the missing data. The crucial conclusion is that the efficiency of the MI approach will depend on the amount of missing data and on the characteristics of the studied disease and the data source. The knowledge on the data generating processes and the healthcare seeking pattern as well as the routines for setting a diagnosis are crucial for obtaining reliable prevalence estimates. In OA, which is often diagnosed in primary care and where the diagnosis is not set very often, ignoring the missing diagnostic codes could lead to a severe underestimation of the prevalence. However, the MI approach used in this study was not optimal in all data generation scenarios and thus this method should be modified in order to provide unbiased estimates in all situations. In particular, an improved calibration procedure for rounding the imputed values may be a solution.

A new implementation of the MI method for longitudinal data has recently become available. In this method, accessible through package *jomo* in R,[129] the ordinal and categorical variables can be imputed using a correct model while accounting for the dependence between the observations coming from the same individual. This may be a superior alternative to the multivariate normal model used in my study and should be evaluated in the future.

Application to the prevalence of doctor-diagnosed OA

The results from the simulation study suggest that the impact of missing data on the estimates of the prevalence of doctor-diagnosed OA in the paper I could vary between

the age groups or joints studied. The number of consultations with the OA diagnosis registered was on average higher in persons with the disease in the knee or the hip than in those with the disease in the hand or other joints. The total number of healthcare visits was on average highest in the oldest age group. Based on the results of the simulation study I could expect that the age specific prevalence in paper I was overestimated for the knee in persons aged 65 years or older by relatively 12%, while it was almost unbiased for hip and other joints, and underestimated for the hand by 15%. After applying the age and location specific corrections (based on the simulation results) to the prevalence estimates from paper I, the corrected prevalence of OA would be 13.0% in the knee, 6.0% in the hip, 3.7% in the hand and 13.1% in the other joints.

On the other hand, if the prevalence estimates in the paper I were based only on the observed diagnostic codes for OA, I would obtain an estimate of 8.7% in the knee, 4.1% in the hip, 2.7% in the hand and 7.0% in the other joints. Based on the simulation results, this would be an underestimation of the true prevalence by 20% for the knee, 26% for the hip and over 30% for the hand and the other joints. Similarly, these naïve estimates can be corrected using the results of the simulations study. The corrected estimates from the naïve method were not fully consistent with the corrected estimates from MI (the differences in corrected prevalences were 2% for the knee and 3% for the hip between the two methods). One of the possible explanations is that in the SHR there may be a component of data MNAR with respect to the information included in the imputation model.

The above corrections are crude and based on the assumption that the SHR data were truly MAR and that the results of the simulation study are directly applicable to the SHR data – a data set more complicated that the one used in the simulations. In some age groups and locations, the prevalence of OA was considerably smaller than 0.1 (the smallest prevalence considered in the simulation study) and it is unclear if the simulation results hold. Nevertheless, the MI approach seems to provide nearly unbiased estimates for most of the locations and age groups when estimating the prevalence of OA in middle-aged and older persons in Skåne. However, there is a risk of overestimating the prevalence of the knee OA in older adults.

Conclusions

In this thesis I have studied epidemiology of OA in Skåne as well as epidemiological methods that can be used in research based on routinely collected healthcare data. I have used population based healthcare register data and data from clinical examinations including assessment of knee symptoms and radiographic findings in knees. The studies have led to the following conclusions:

- In Sweden, one in four adults aged 45 years or older is diagnosed with OA in at least one peripheral joint and 15.4% of middle aged and elderly persons have knee OA with frequent knee pain. The OA disease is thus one of the most common chronic diseases among Swedish adults and is associated with substantial healthcare consumption and personal suffering.
- Half of the persons with knee OA giving symptoms consulted a physician and were diagnosed with knee OA during an 8 year time period. There is a potential unmet need among patients with painful disease, as many do not consult a physician for the disease.
- Due to the ageing of the population and obesity epidemic the prevalence of OA may increase to 30% within the next 20 years. Preventive strategies against obesity and joint injury are needed together with comprehensive healthcare utilization strategies for the handling of OA.
- Due to the complexity of the OA disease and the poor overlap between groups defined using the common epidemiological definitions, persons with OA diagnosed within healthcare system may be an important group to study. They represent a clinically relevant disease cohort that only partly overlap with the cohorts defined using the common OA classification criteria.
- Doctor-diagnosed knee or hip OA in Sweden is not associated with increased all-cause mortality as compared to the general population. Possible mechanisms include relatively easy access to care or maintaining physical activity despite the disease.
- Epidemiological methods for the handling of missing data and possible sources of bias may give additional information on the robustness of the results derived using register data. In particular, missing data in longitudinal healthcare databases is a challenging problem that requires further development of the existing methods.

Future perspectives

The OA disease is highly prevalent in Sweden and impacts individuals and the healthcare system. The results of this thesis do not only provide epidemiological information on the occurrence of OA but may also aid the healthcare planning for the future. A more in-depth understanding of the healthcare utilization by OA patients, including presence of comorbidities and use of physiotherapy are important questions for future research. In OA, a complex disease that is challenging to define, the longitudinal healthcare registers may be a promising source of clinically relevant data about the occurrence or consequences of the disease, if the data quality is sufficient. One of the important topics that require further investigation is the association between OA and mortality. Replication of the results from this thesis using other definitions of OA and in other countries is warranted. The specific-cause mortality is of interest, especially with respect to the cardiovascular diseases that have been reported to be a consequence of OA.

Epidemiological methods for quantification of bias are important in assessing the robustness of the results derived from the longitudinal healthcare registers, where the validity of the data may be unknown or when data on important confounders may not be available. This type of analysis can show the likely effects of anticipated sources of bias and should be used more often in register based research. It requires some analytical effort but may be effective and informative if relevant substudies are not feasible.

Populärvetenskaplig sammanfattning

Artros är vår vanligaste kroniska ledsjukdom. Sjukdomen omfattar i princip alla strukturer men kännetecknas typiskt av nedbrytning av ledbrosk. Artros orsakar ofta ledsmärta och leder till nedsatt ledfunktion. Trots att artros är en folksjukdom så finns det förvånansvärt lite kunskap om förekomst av artros i Sverige, om hur många som har symptom eller strukturella förändringar i leder, om hur många som söker vård, samt om sambandet mellan artros och dödlighet.

Syftena med denna avhandling var primärt att studera förekomst av artros i Sverige idag, förutse förekomst av artros om 20 år, samt att få ny kunskap om sambandet mellan artros och dödlighet.

Alla läkarbesök som görs i Skåne registreras i Region Skånes Vårddatabaser. Bland en halv miljon skåningar i åldern 45 år eller äldre har jag med avidentifierade data studerat de som fick en artrosdiagnos hos en läkare minst en gång de senaste 14 åren (1999-2012). En mindre grupp av personer i åldrarna 56 till 84 år, boende i Malmö, har deltagit i en klinisk studie – MOA studie. De har svarat på ett frågeformulär om knäbesvär och fått bägge knäna röntgade för att undersöka om de hade förändringar i knäleden så som vid artros. Jag har använt dessa olika datakällor för att skatta förekomst av artros i Skåne samt för att studera dödligheten hos personer med och utan artros. I avhandlingen har jag använt ett antal olika statistiska metoder och s.k. sensitivitetsanalyser för att utreda om resultaten är tillförlitliga.

Jag fann att en av fyra skåningar 45 år eller äldre hade fått diagnosen artros av en läkare i minst en led: knä, höft, hand eller andra leder så som fot-, armbågs-, skulder- eller käkled. Knäledsartros var diagnosticerad hos 13.8% av befolkningen i denna åldersgrupp, höftledsartros hos 5.8%, artros i fingrarna hos 3.1% och 12.4% hade artros i andra leder. År 2032 kan vi förvänta oss att ca 30% av befolkningen i Sverige över 45 år kommer att ha sökt läkare och fått diagnosen artros. Detta främst beroende på vår åldrande befolkning men även på ökad förekomst av övervikt och fetma. I MOA studien fann jag att 25% av personerna i åldern 56 till 84 år hade knäsmärta och lika många hade röntgenologiska fynd som vid artros. Dessa två grupper överlappar dock inte helt. Femton procent hade symptomgivande knäledsartros d.v.s. uppfyllde kliniska kriterier för knäledsartros eller hade både knäsmärta och artrostecken på röntgen. Under en 8 års period hade hälften av dessa fått diagnosen knäledsartros hos en läkare.

Jag fann att personer med knä- eller höftledsartros inte löper större risk att dö jämfört med bakgrundsbefolkningen. En av förklaringarna kan vara att omhändertagandet av personer med artros är förhållandevis bra i Sverige, med relativt

enkel tillgång till sjukgymnastik, ledproteskirurgi och (eller) att personer med artros i Sverige upprätthåller fysisk aktivitet oavsett sjukdomen.

Denna avhandling tillhandahåller ny kunskap om kliniskt relevant artros i Skåne. Resultaten kan bistå i bättre planering inom sjukvården samt till att utvärdera sjukdomsbördan i samhället.

Acknowledgements

This work would not be possible without the help and engagement of many wonderful persons that I have had the privilege to work with and learn from. I would like to thank you all!

Special thanks to:

Martin Englund, my main supervisor. You have inspired me to start the PhD project, shaped me as a researcher, always shared your broad knowledge and time, and given me the freedom to work as I wanted to. I am grateful to you.

Jonas Björk, my co-supervisor. You have always met my anxious questions with calm and balanced answers; you have also inspired me to take the methodological courses that I truly needed.

My co-authors: *Stefan Lohmander, Ingemar Petersson, Gillian Hawker, Leif Dahlberg, Maria Gerhardsson de Verdier, Gunnar Engström, Peter Nilsson, Carl Mellström, Tuhina Neogi* and *George Peat*. You have shared your knowledge and shown me points of view that I have not been able to see myself. I have benefitted greatly from the cooperation with you.

My co-workers at the Clinical Epidemiology Unit and at the ERC South for fruitful discussions and a lot of fun at work, especially to *Jonas Ranstam, Rebecca Rylance* and *Philippe Wagner* for all the statistical (and other) disputes; and to *Charlotte Bergknut* for the superb data management support.

Astra Zeneca, *staff* and *participants* from the *Malmö Osteoarthritis Study* for giving me the opportunity to work with this study. It has become a very important part of this thesis.

To my parents, *Roma* and *Tomasz*, for always believing in me and to my brother *Tomasz* for all the babysitting he has done.

To my beloved husband *Dominik* for being a true supporter and an inspiring discussion partner and to my kids, *Jakub* and *Maria*, for putting everything in the right perspective.

References

1. Last JM, Spasoff RA, Harris SS, Thuriaux MC, Byron Anderson J. A Dictionary of Epidemiology. Oxford University Press; 2000.

2. Bhopal R. Concepts of Epidemiology. Oxford University Press; 2008.

3. Rothman KJ, Greenland S, Lash TL. Introduction. In: Modern Epidemiology. Rothman, K.J., Greenland, S., Lash, T.L. Philadelphia, PA: Lippincott, Williams & Wilkins; 2008.

4. Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, et al. The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. Am J Public Health. 1994 Mar;84(3):351–8.

5. Losina E, Thornhill TS, Rome BN, Wright J, Katz JN. The Dramatic Increase in Total Knee Replacement Utilization Rates in the United States Cannot Be Fully Explained by Growth in Population Size and the Obesity Epidemic. J Bone Joint Surg Am. 2012 Feb 1;94(3):201–7.

6. Robertsson O, Bizjajeva S, Fenstad AM, Furnes O, Lidgren L, Mehnert F, et al. Knee arthroplasty in Denmark, Norway and Sweden. Acta Orthopaedica. 2010 Feb 1;81(1):82–9.

7. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012 Dec 15;380(9859):2163–96.

8. Sharma L, Kapoor D. Epidemiology of Osteoarthritis. In: Osteoarthritis: Diagnosis and Medical/Surgical Management. Fourth edition. Philadelphia, PA: Lippincott Williams & Wilkins; 2007.

9. Taylor WJ, Fransen J. Distinctions between diagnostic and classification criteria: Comment on the article by Aggawal et al. Arthritis Care & Research. 2015 Jul 1;n/a – n/a.

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

11. Kellgren JH, Lawrence JS. Radiological Assessment of Osteo-Arthrosis. Ann Rheum Dis. 1957 Dec;16(4):494–502.

12. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum. 1991 May;34(5):505–14.

13. Altman R, Alarcón G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. Arthritis Rheum. 1990 Nov;33(11):1601–10.

14. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. 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 Rheum. 1986 Aug;29(8):1039–49.

15. Hunter DJ, Arden N, Conaghan PG, Eckstein F, Gold G, Grainger A, et al. Definition of osteoarthritis on MRI: results of a Delphi exercise. Osteoarthr Cartil. 2011 Aug;19(8):963–9.

16. Menashe L, Hirko K, Losina E, Kloppenburg M, Zhang W, Li L, et al. The diagnostic performance of MRI in osteoarthritis: a systematic review and meta-analysis. Osteoarthritis Cartilage. 2012 Jan;20(1):13–21.

17. Ding C, Cicuttini F, Jones G. How important is MRI for detecting early osteoarthritis? Nature Clinical Practice Rheumatology. 2008;4(1):4–5.

18. Nationella riktlinjer för rörelseorganens sjukdomar 2012 [Internet]. Socialstyrelsen; 2012 [cited 2015 Oct 13]. Available from: http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/18665/2012-5-1.pdf

19. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis. 2014 Jul;73(7):1323–30.

20. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. Ann Rheum Dis. 2001 Feb 1;60(2):91–7.

21. Peat G, Thomas E, Duncan R, Wood L, Hay E, Croft P. Clinical classification criteria for knee osteoarthritis: performance in the general population and primary care. Ann Rheum Dis. 2006 Oct;65(10):1363–7.

22. Duncan RC, Hay EM, Saklatvala J, Croft PR. Prevalence of radiographic osteoarthritis—it all depends on your point of view. Rheumatology. 2006 Jun 1;45(6):757–60.

23. Jordan JM, Helmick CG, Renner JB, Luta G, Dragomir AD, Woodard J, et al. Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol. 2007 Jan;34(1):172–80.

24. Kim C, Linsenmeyer KD, Vlad SC, Guermazi A, Clancy MM, Niu J, et al. Prevalence of radiographic and symptomatic hip osteoarthritis in an urban United States community: the Framingham osteoarthritis study. Arthritis & Rheumatology (Hoboken, NJ). 2014 Nov;66(11):3013–7.

25. Kopec JA, Rahman MM, Berthelot J-M, Le Petit C, Aghajanian J, Sayre EC, et al. Descriptive epidemiology of osteoarthritis in British Columbia, Canada. J Rheumatol. 2007 Feb;34(2):386–93.

26. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Prevalence and burden of osteoarthritis: results from a population survey in Norway. J Rheumatol. 2008 Apr;35(4):677–84.

27. Lawrence RC, Felson DT, Helmick CG, Arnold LM, Choi H, Deyo RA, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum. 2008 Jan;58(1):26–35.

28. Andrianakos AA, Kontelis LK, Karamitsos DG, Aslanidis SI, Georgountzos AI, Kaziolas GO, et al. Prevalence of symptomatic knee, hand, and hip osteoarthritis in Greece. The ESORDIG study. J Rheumatol. 2006 Dec;33(12):2507–13.

29. Musculoskeletal Problems and Functional Limitation [Internet]. EUROPEAN COMMISSION,DIRECTORATE-GENERAL HEALTH & CONSUMER PROTECTION DIRECTORATE GENERAL - PUBLIC HEALTH, University of Oslo; 2003 [cited 2015 Oct 13]. Available from: http://ec.europa.eu/health/ph_projects/2000/monitoring/fp_monitoring_2000_frep_01_en.pdf

30. Carmona L, Ballina J, Gabriel R, Laffon A. The burden of musculoskeletal diseases in the general population of Spain: results from a national survey. Ann Rheum Dis. 2001 Nov 1;60(11):1040–5.

31. Dahaghin S, Bierma-Zeinstra SMA, Ginai AZ, Pols H a. P, Hazes JMW, Koes BW. Prevalence and pattern of radiographic hand osteoarthritis and association with pain and disability (the Rotterdam study). Ann Rheum Dis. 2005 May;64(5):682–7.

32. March LM, Bagga H. Epidemiology of osteoarthritis in Australia. Med J Aust. 2004 Mar 1;180(5 Suppl):S6–10.

33. Bagge E, Bjelle A, Edén S, Svanborg A. A longitudinal study of the occurrence of joint complaints in elderly people. Age Ageing. 1992 May;21(3):160–7.

34. Bagge E, Bjelle A, Valkenburg HA, Svanborg A. Prevalence of radiographic osteoarthritis in two elderly European populations. Rheumatol Int. 1992;12(1):33–8.

35. Petersson IF, Boegård T, Saxne T, Silman AJ, Svensson B. Radiographic osteoarthritis of the knee classified by the Ahlbäck and Kellgren & Lawrence systems for the tibiofemoral joint in people aged 35-54 years with chronic knee pain. Ann Rheum Dis. 1997 Aug;56(8):493–6.

36. Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of Primary and Revision Hip and Knee Arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am. 2007 Apr 1;89(4):780–5.

37. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. JAMA. 2010 Jan 20;303(3):235–41.

38. Hootman JM, Helmick CG. Projections of US prevalence of arthritis and associated activity limitations. Arthritis Rheum. 2006 Jan;54(1):226–9.

39. Spector TD, Cicuttini F, Baker J, Loughlin J, Hart D. Genetic influences on osteoarthritis in women: a twin study. BMJ. 1996 Apr 13;312(7036):940–3.

40. Felson DT, Lawrence RC, Dieppe PA, Hirsch R, Helmick CG, Jordan JM, et al. Osteoarthritis: New Insights. Part 1: The Disease and Its Risk Factors. Ann Intern Med. 2000 Oct 17;133(8):635–46.

41. Buckwalter JA, Lane NE. Athletics and osteoarthritis. Am J Sports Med. 1997 Dec;25(6):873–81.

42. Englund M, Roos EM, Lohmander LS. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. Arthritis Rheum. 2003 Aug;48(8):2178–87.

43. Øiestad BE, Juhl CB, Eizen I, Thorlund JB. Knee extensor muscle weakness increases the risk of knee osteoarthritis. a systematic review and meta-analysis. Osteoarthritis and Cartilage. 2014 Apr 1;22:S336.

44. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthr Cartil. 2014 Mar;22(3):363–88.

45. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res. 2012;64(4):465–74.

46. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, et al. OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. Osteoarthr Cartil. 2008 Feb;16(2):137–62.

47. Albers JM, Kuper HH, Riel PL van, Prevoo ML, Hof MA van 't, Gestel AM van, et al. Socio-economic consequences of rheumatoid arthritis in the first years of the disease. Rheumatology. 1999 May 1;38(5):423–30.

48. O'Dell JR. Therapeutic Strategies for Rheumatoid Arthritis. New England Journal of Medicine. 2004 Jun 17;350(25):2591–602.

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

50. Englund M, Jöud A, Geborek P, Felson DT, Jacobsson LT, Petersson IF. Prevalence and incidence of rheumatoid arthritis in southern Sweden 2008 and their relation to prescribed biologics. Rheumatology (Oxford). 2010 Aug;49(8):1563–9.

51. Oliver JE, Silman AJ. Risk factors for the development of rheumatoid arthritis. Scandinavian Journal of Rheumatology. 2006 May;35(3):169–74.

52. Singh JA, Saag KG, Bridges SL, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. Arthritis & Rheumatology (Hoboken, NJ). 2016 Jan;68(1):1–26.

53. Pincus T, Gibson KA, Block JA. Premature Mortality: A Neglected Outcome in Rheumatic Diseases? Arthritis Care Res (Hoboken). 2015 Aug;67(8):1043–6.

54. Sokka T, Abelson B, Pincus T. Mortality in rheumatoid arthritis: 2008 update. Clin Exp Rheumatol. 2008 Oct;26(5 Suppl 51):S35–61.

55. Nuesch E, Dieppe P, Reichenbach S, Williams S, Iff S, Juni P. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. BMJ (Clinical research ed). 2011;342:d1165.

56. Hawker GA, Croxford R, Bierman AS, Harvey PJ, Ravi B, Stanaitis I, et al. All-Cause Mortality and Serious Cardiovascular Events in People with Hip and Knee Osteoarthritis: A Population Based Cohort Study. PLoS ONE. 2014 Mar 7;9(3):e91286.

57. Hochberg MC. Mortality in osteoarthritis. Clin Exp Rheumatol. 2008 Oct;26(5 Suppl 51):S120–4.

58. Lee TA, Pickard AS, Bartle B, Weiss KB. Osteoarthritis: a comorbid marker for longer life? Ann Epidemiol. 2007 May;17(5):380–4.

59. Haugen IK, Ramachandran VS, Misra D, Neogi T, Niu J, Yang T, et al. Hand osteoarthritis in relation to mortality and incidence of cardiovascular disease: data from the Framingham heart study. Ann Rheum Dis. 2015 Jan;74(1):74–81.

60. Fridh M, Moden B, Lindström M, Grahn M, Rosvall M. Folkhälsorapport Skåne 2013. Region Skåne; 2013.

61. Berglund G, Elmstähl S, Janzon L, Larsson SA. The Malmo Diet and Cancer Study. Design and feasibility. J Intern Med. 1993 Jan;233(1):45–51.

62. Lohmander LS, Gerhardsson de Verdier M, Rollof J, Nilsson PM, Engström G. Incidence of severe knee and hip osteoarthritis in relation to different measures of body mass: a population-based prospective cohort study. Ann Rheum Dis. 2009 Apr;68(4):490–6.

63. Manjer J, Carlsson S, Elmståhl S, Gullberg B, Janzon L, Lindström M, et al. The Malmö Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. Eur J Cancer Prev. 2001 Dec;10(6):489–99.

64. Olsen J. Register-based research: some methodological considerations. Scand J Public Health. 2011 May;39(3):225–9.

65. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. Eur J Epidemiol. 2009 Nov;24(11):659–67.

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

67. Olsen J. Using Secondary Data. In: Modern Epidemiology. Rothman, K.J., Greenland, S., Lash, T.L. Philadelphia, PA: Lippincott, Williams & Wilkins; 2008.

68. Population registration in Sweden, SKV 717B edition 6 - 717B06.pdf [Internet]. [cited 2015 Nov 25]. Available from: https://www.skatteverket.se/download/18.8dcbbe4142d38302d74be9/1387372677650/717B06.pdf

69. Longitudinell intergrationsdatabas för Sjukdomförsäkrings- och Arbetsmarkandsstudier (LISA) 1990-2009 [Internet]. Statistiska Centralbyrån; 2011 [cited 2015 Nov 25]. Available from: http://www.scb.se/statistik/_publikationer/AM9901_1990I09_BR_AM76BR1104.pdf

70. Löfvendahl S, Theander E, Svensson Å, Carlsson KS, Englund M, Petersson IF. Validity of diagnostic codes and prevalence of physician-diagnosed psoriasis and psoriatic arthritis in southern Sweden--a population-based register study. PLoS ONE. 2014;9(5):e98024.

71. Haglund E, Bremander AB, Petersson IF, Strömbeck B, Bergman S, Jacobsson LTH, et al. Prevalence of spondyloarthritis and its subtypes in southern Sweden. Ann Rheum Dis. 2011 Jun;70(6):943–8.

72. Rosengren BE, Karlsson M, Petersson I, Englund M. The 21st-century landscape of adult fractures: cohort study of a complete adult regional population. J Bone Miner Res. 2015 Mar;30(3):535–42.

73. Altman RD, Hochberg M, Murphy WA, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. Osteoarthr Cartil. 1995 Sep;3 Suppl A:3–70.

74. Greenland S, Rothman KJ. Measure of Occurence. In: Modern Epidemiology. Rothman, K.J., Greenland, S., Lash, T.L. Philadelphia, PA: Lippincott, Williams & Wilkins; 2008.

75. Marshall DA, Vanderby S, Barnabe C, MacDonald KV, Maxwell C, Mosher D, et al. Estimating the Burden of Osteoarthritis to Plan for the Future. Arthritis Care Res (Hoboken). 2015 Oct;67(10):1379–86.

76. Jiang L, Rong J, Wang Y, Hu F, Bao C, Li X, et al. The relationship between body mass index and hip osteoarthritis: a systematic review and meta-analysis. Joint Bone Spine. 2011 Mar;78(2):150–5.

77. Jiang L, Tian W, Wang Y, Rong J, Bao C, Liu Y, et al. Body mass index and susceptibility to knee osteoarthritis: a systematic review and meta-analysis. Joint Bone Spine. 2012 May;79(3):291–7.

78. Allison PD. Missing Data [Internet]. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc.; 2002 [cited 2016 Feb 9]. Available from: http://srmo.sagepub.com/view/missing-data/SAGE.xml

79. Rubin DB. Introduction. In: Multiple Imputation for Nonresponse in Surveys [Internet]. John Wiley & Sons, Inc.; 1987 [cited 2014 Sep 24]. p. 1–26. Available from: http://onlinelibrary.wiley.com/doi/10.1002/9780470316696.ch1/summary

80. Carpenter JR, Goldstein H, Kenward MG. Statistical Modelling of Partially Observed Data Using Multiple Imputation: Principles and Practice. In: Tu Y-K, Greenwood DC, editors. Modern Methods for Epidemiology [Internet]. Dordrecht: Springer Netherlands; 2012 [cited 2015 Oct 12]. p. 15–31. Available from: http://www.springerlink.com/index/10.1007/978-94-007-3024-3_2

81. Spratt M, Carpenter J, Sterne JAC, Carlin JB, Heron J, Henderson J, et al. Strategies for multiple imputation in longitudinal studies. Am J Epidemiol. 2010 Aug 15;172(4):478–87.

82. Schafer JL, Yucel RM. Computational Strategies for Multivariate Linear Mixed-Effects Models With Missing Values. Journal of Computational and Graphical Statistics. 2002 Jun 1;11(2):437–57.

83. Schafer JL. Analysis of Incomplete Multivariate Data [Internet]. Chapman & Hall/CRC; 1997 [cited 2014 Sep 24]. Available from: http://books.google.se/books/about/Analysis_of_Incomplete_Multivariate_Data.html?id=3TFWRjn1f-oC&redir_esc=y

84. Bernaards CA, Belin TR, Schafer JL. Robustness of a multivariate normal approximation for imputation of incomplete binary data. Statist Med. 2007 Mar 15;26(6):1368–82.

85. Yucel RM, He Y, Zaslavsky AM. Using Calibration to Improve Rounding in Imputation. The American Statistician. 2008 May 1;62(2):125–9.

86. Little RJ, Vartivarian S. On weighting the rates in non-response weights. Stat Med. 2003 May 15;22(9):1589–99.

87. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological). 1972;34(2):187–220.

88. Kleinbaum DG, Klein M. Introduction to Survival Analysis. In: Survival Analysis [Internet]. New York, NY: Springer New York; 2012 [cited 2016 Feb 17]. p. 1–54. Available from: http://link.springer.com/10.1007/978-1-4419-6646-9_1

89. Thiébaut ACM, Bénichou J. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. Stat Med. 2004 Dec 30;23(24):3803–20.

90. Therneau TM, Grambsch PM. Testing Proportional Hazards. In: Modeling Survival Data: Extending the Cox Model [Internet]. New York, NY: Springer New York; 2000 [cited 2015 Nov 9]. (Dietz K, Gail M, Krickeberg K, Samet J, Tsiatis A, editors. Statistics for Biology and Health). Available from: http://link.springer.com/10.1007/978-1-4757-3294-8

91. Law GR, Green R, Ellison GTH. Confounding and Causal Path Diagrams. In: Tu Y-K, Greenwood DC, editors. Modern Methods for Epidemiology [Internet]. Dordrecht: Springer Netherlands; 2012 [cited 2015 Oct 12]. p. 1–13. Available from: http://www.springerlink.com/index/10.1007/978-94-007-3024-3_1

92. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000 Sep;11(5):550–60.

93. Rahman MM, Kopec JA, Anis AH, Cibere J, Goldsmith CH. Risk of cardiovascular disease in patients with osteoarthritis: a prospective longitudinal study. Arthritis Care Res (Hoboken). 2013 Dec;65(12):1951–8.

94. Ong KL, Wu BJ, Cheung BMY, Barter PJ, Rye K-A. Arthritis: its prevalence, risk factors, and association with cardiovascular diseases in the United States, 1999 to 2008. Ann Epidemiol. 2013 Feb;23(2):80–6.

95. Fewell Z, Hernán MA, Wolfe F, Tilling K, Choi H, Sterne JAC. Controlling for time-dependent confounding using marginal structural models. Stata Journal. 2004;4(4):402–20.

96. Hernán MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. Statist Med. 2002 Jun 30;21(12):1689–709.

97. Lash TL, Fox MP, Fink AK. Applying Quantitative Bias Analysis to Epidemiologic Data [Internet]. New York, NY: Springer New York; 2009 [cited 2015 Nov 4]. (Statistics for Biology and Health). Available from: http://link.springer.com/10.1007/978-0-387-87959-8

98. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiol Drug Saf. 2006 May;15(5):291–303.

99. Zhang Y, Niu J, Kelly-Hayes M, Chaisson CE, Aliabadi P, Felson DT. Prevalence of Symptomatic Hand Osteoarthritis and Its Impact on Functional Status among the Elderly The Framingham Study. Am J Epidemiol. 2002 Dec 1;156(11):1021–7.

100. Dillon CF, Hirsch R, Rasch EK, Gu Q. Symptomatic hand osteoarthritis in the United States: prevalence and functional impairment estimates from the third U.S. National Health and Nutrition Examination Survey, 1991-1994. Am J Phys Med Rehabil. 2007 Jan;86(1):12–21.

101. Marshall M, Peat G, Nicholls E, van der Windt D, Myers H, Dziedzic K. Subsets of symptomatic hand osteoarthritis in community-dwelling older adults in the United Kingdom: prevalence, inter-relationships, risk factor profiles and clinical characteristics at baseline and 3-years. Osteoarthritis and Cartilage. 2013 Nov;21(11):1674–84.

102. Cushnaghan J, Dieppe P. Study of 500 patients with limb joint osteoarthritis. I. Analysis by age, sex, and distribution of symptomatic joint sites. Ann Rheum Dis. 1991 Jan 1;50(1):8–13.

103. Croft P, Cooper C, Wickham C, Coggon D. Is the Hip Involved in Generalized Osteoarthritis? Rheumatology. 1992 May 1;31(5):325–8.

104. Prieto-Alhambra D, Judge A, Javaid MK, Cooper C, Diez-Perez A, Arden NK. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints. Ann Rheum Dis. 2014 Sep;73(9):1659–64.

105. Widdifield J, Labrecque J, Lix L, Paterson JM, Bernatsky S, Tu K, et al. Systematic Review and Critical Appraisal of Validation Studies to Identify Rheumatic Diseases in Health Administrative Databases. Arthritis Care & Research. 2013 Sep 1;65(9):1490–503.

106. Yu D, Peat G, Bedson J, Jordan KP. Annual consultation incidence of osteoarthritis estimated from population-based health care data in England. Rheumatology. 2015 Jul 9;kev231.

107. van Saase JL, van Romunde LK, Cats A, Vandenbroucke JP, Valkenburg HA. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. Ann Rheum Dis. 1989 Apr;48(4):271–80.

108. Dillon CF, Rasch EK, Gu Q, Hirsch R. Prevalence of knee osteoarthritis in the United States: arthritis data from the Third National Health and Nutrition Examination Survey 1991-94. J Rheumatol. 2006 Nov;33(11):2271–9.

109. Felson DT, Naimark A, Anderson J, Kazis L, Castelli W, Meenan RF. The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. Arthritis Rheum. 1987 Aug;30(8):914–8.

110. Odding E, Valkenburg HA, Algra D, Vandenouweland FA, Grobbee DE, Hofman A. Associations of radiological osteoarthritis of the hip and knee with locomotor disability in the Rotterdam Study. Ann Rheum Dis. 1998 Apr;57(4):203–8.

111. Anderson JJ, Felson DT. Factors associated with osteoarthritis of the knee in the first national Health and Nutrition Examination Survey (HANES I). Evidence for an association with overweight, race, and physical demands of work. Am J Epidemiol. 1988 Jul;128(1):179–89.

112. Neogi T, Felson D, Niu J, Nevitt M, Lewis CE, Aliabadi P, et al. Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. BMJ [Internet]. 2009 [cited 2015 Nov 23];339. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2730438/

113. Wallis JA, Webster KE, Levinger P, Taylor NF. What proportion of people with hip and knee osteoarthritis meet physical activity guidelines? A systematic review and meta-analysis. Osteoarthr Cartil. 2013 Nov;21(11):1648–59.

114. Woolf AD, Zeidler H, Haglund U, Carr AJ, Chaussade S, Cucinotta D, et al. Musculoskeletal pain in Europe: its impact and a comparison of population and medical perceptions of treatment in eight European countries. Ann Rheum Dis. 2004 Apr;63(4):342–7.

115. Andersson T, Ahlbom A, Carlsson S. Diabetes Prevalence in Sweden at Present and Projections for Year 2050. PLoS One [Internet]. 2015 Nov 30 [cited 2016 Jan 26];10(11). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4664416/

116. Blair SN, Kohl HW, Paffenbarger RS, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality. A prospective study of healthy men and women. JAMA. 1989 Nov 3;262(17):2395–401.

117. Bedson J, Mottram S, Thomas E, Peat G. Knee pain and osteoarthritis in the general population: what influences patients to consult? Fam Pract. 2007 Oct;24(5):443–53.

118. Jordan K, Jinks C, Croft P. A prospective study of the consulting behaviour of older people with knee pain. Br J Gen Pract. 2006 Apr 1;56(525):269–76.

119. Paskins Z, Sanders T, Hassell AB. What influences patients with Osteoarthritis to consult their GP about their symptoms? A narrative review. BMC Family Practice. 2013 Dec 20;14(1):195.

120. Lawrence RC, Everett DF, Hochberg MC. Arthritis. In: Joan C. Cornoni-Huntley, Robert R. Huntley and Jacob J. Feldman. New York: Oxford University Press; 1989. p. 136–51.

121. Watson DJ, Rhodes T, Guess HA. All-cause mortality and vascular events among patients with rheumatoid arthritis, osteoarthritis, or no arthritis in the UK General Practice Research Database. J Rheumatol. 2003 Jun;30(6):1196–202.

122. Vaz de Almeida MD, Graça P, Afonso C, D'Amicis A, Lappalainen R, Damkjaer S. Physical activity levels and body weight in a nationally representative sample in the European Union. Public Health Nutrition. 1999 Jan;2(Supplement 1a):105–14.

123. Neovius M, Janson A, Rössner S. Prevalence of Obesity in Sweden. Obesity Reviews. 2006 Feb 1;7(1):1–3.

124. Jansen MJ, Viechtbauer W, Lenssen AF, Hendriks EJM, de Bie RA. Strength training alone, exercise therapy alone, and exercise therapy with passive manual mobilisation each reduce pain and disability in people with knee osteoarthritis: a systematic review. Journal of Physiotherapy. 2011;57(1):11–20.

125. Fransen M, McConnell S, Hernandez-Molina G, Reichenbach S. Does land-based exercise reduce pain and disability associated with hip osteoarthritis? A meta-analysis of randomized controlled trials. Osteoarthritis and Cartilage. 2010 May;18(5):613–20.

126. Walsh M, Woodhouse LJ, Thomas SG, Finch E. Physical Impairments and Functional Limitations: A Comparison of Individuals 1 Year After Total Knee Arthroplasty With Control Subjects. PHYS THER. 1998 Mar 1;78(3):248–58.

127. Herbolsheimer F, Schaap LA, Edwards MH, Maggi S, Otero Á, Timmermans EJ, et al. Physical Activity Patterns Among Older Adults With and Without Knee Osteoarthritis in Six European Countries. Arthritis Care & Research. 2016 Feb 1;68(2):228–36.

128. Kiadaliri AA, Lamm CJ, Gerhardsson de Verdier M, Engström G, Turkiewicz A, Lohmander LS, et al. Association of knee pain and different definitions of knee osteoarthritis with health-related quality of life: a population- based cohort study in southern Sweden. submitted.

129. Quartagno M, Carpenter J. Multilevel Joint Modelling Multiple Imputation [Internet]. 2015 [cited 2015 Dec 1]. Available from: https://cran.r-project.org/web/packages/jomo/jomo.pdf

LUND UNIVERSITY
Faculty of Medicine